

RICCARDO RICCI

LEZIONI DI SISTEMI DINAMICI

a cura di

LUIGI BARLETTI

ANGIOLO FARINA

LORENZO FUSI

FEDERICO TALAMUCCI



MANUALI
SCIENZE

- 8 -

RICCARDO RICCI

Lezioni di Sistemi Dinamici

a cura di

LUIGI BARLETTI

ANGIOLO FARINA

LORENZO FUSI

FEDERICO TALAMUCCI

Firenze University Press

2016

Lezioni di Sistemi Dinamici / Riccardo Ricci ; a cura di Luigi Barletti, Angiolo Farina, Lorenzo Fusi, Federico Talamucci. – Firenze : Firenze University Press, 2016. (Manuali . Scienze ; 8)

<http://digital.casalini.it/9788864534015>

ISBN 978-88-6453-400-8 (print)

ISBN 978-88-6453-401-5 (online)

Immagine di copertina: La figura di copertina, gentilmente concessa dalla famiglia Ricci, rappresenta due trottole in movimento sopra gli appunti del professor Ricci. Il moto delle trottole ha un'elegante descrizione matematica che viene trattata nel capitolo 7 di questo libro.

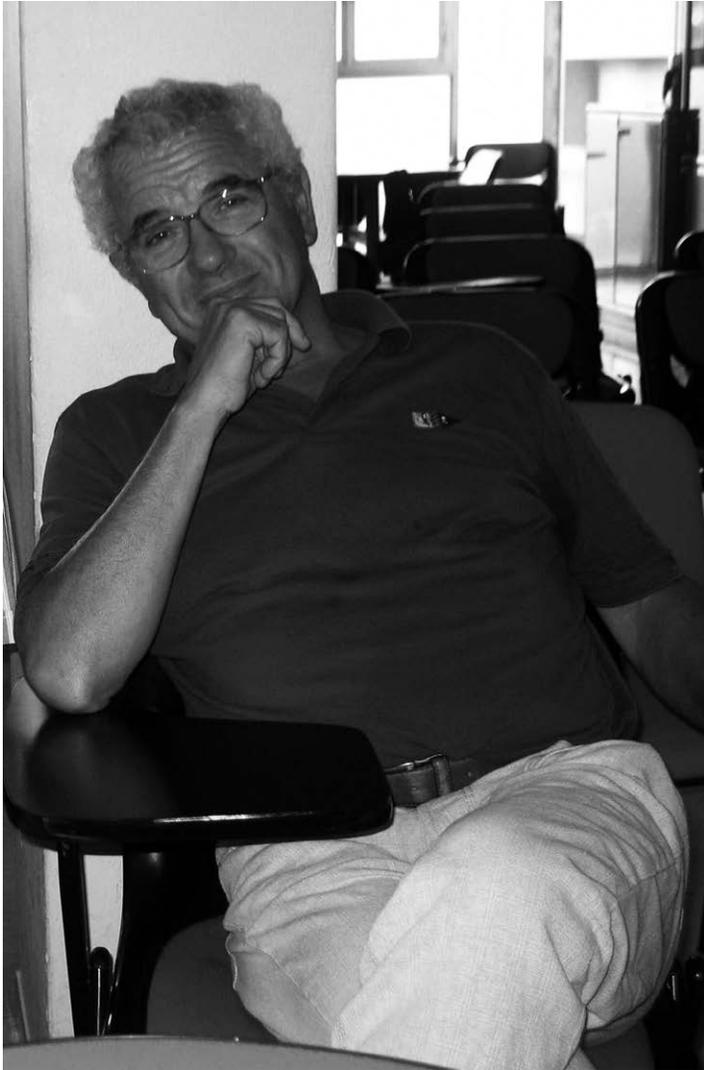
Certificazione scientifica delle Opere

Tutti i volumi pubblicati sono soggetti ad un processo di referaggio esterno di cui sono responsabili il Consiglio editoriale della FUP e i Consigli scientifici delle singole collane. Le opere pubblicate nel catalogo della FUP sono valutate e approvate dal Consiglio editoriale della casa editrice. Per una descrizione più analitica del processo di referaggio si rimanda ai documenti ufficiali pubblicati sul catalogo on-line della casa editrice (www.fupress.com).

Consiglio editoriale Firenze University Press

G. Nigro (Coordinatore), M.T. Bartoli, M. Boddi, R. Casalbuoni, C. Ciappei, R. Del Punta, A. Dolfi, V. Fargion, S. Ferrone, M. Garzaniti, P. Guarnieri, A. Mariani, M. Marini, A. Novelli, M.C. Torricelli, M. Verga, A. Zorzi.

© 2016 Firenze University Press
Università degli Studi di Firenze
Firenze University Press
via Cittadella, 7, 50144 Firenze, Italy
www.fupress.com
Printed in Italy



RICCARDO RICCI (1953-2013) è stato professore ordinario di Fisica Matematica nell'Università di Firenze. La sua attività di ricerca si è svolta prevalentemente nell'ambito dei problemi a frontiera libera e delle equazioni alle derivate parziali. È stato titolare per numerosi anni della cattedra di Meccanica Razionale e di Sistemi Dinamici. Per tutta la sua carriera accademica si è dedicato attivamente alla divulgazione e alla didattica della matematica, anche attraverso dispense e conferenze, sempre molto apprezzate dagli studenti.

Indice

Introduzione	1
1 Spazio affine e vettori applicati	3
1.1 Spazi vettoriali su \mathbb{R}	3
1.2 Spazi vettoriali euclidei	7
1.3 Trasformazioni di basi ortonormali	14
1.4 Spazio affine e coordinate curvilinee	16
1.5 Prodotto vettoriale e sistemi di vettori applicati	21
1.5.1 Il prodotto vettoriale	21
1.5.2 Vettori applicati e momento rispetto ad un polo	24
1.5.3 Sistemi equivalenti di vettori applicati	27
1.5.4 Sistemi di vettori paralleli	29
2 Equazioni differenziali	31
2.1 Introduzione	31
2.2 Considerazioni generali	32
2.3 Il problema di Cauchy	34
2.3.1 Equazioni autonome	35
2.3.2 Equazioni reversibili	36
2.3.3 Equazioni integrabili	37
2.4 Equazioni del primo ordine in forma normale	37
2.5 Equazioni del secondo ordine in forma normale del tipo $\ddot{q} = f(\dot{q})$	39
2.6 Equazioni del secondo ordine del tipo $\ddot{q} = f(q)$: caso conservativo	41
2.6.1 Analisi qualitativa nel caso conservativo	43
2.6.2 Periodo delle oscillazioni in vicinanza di punti di equilibrio stabile	54
2.6.3 Generalizzazione del caso conservativo	55
2.7 Il piano delle fasi	56
2.8 Punti di equilibrio, stabilità	58
2.8.1 Il criterio di Lyapunov	59
2.8.2 Asintotica stabilità	60
2.8.3 Punti di equilibrio per sistemi conservativi: il criterio di Dirichlet	62
2.9 I potenziali isocroni	63
2.10 Sistemi lineari bidimensionali	67
2.11 Moto armonico smorzato	78
2.12 Moto armonico smorzato con forzante esterna	80

2.13	Sistemi lineari n -dimensionali	83
3	Equazioni di Lagrange	87
3.1	Cinematica del punto	87
3.2	Forze conservative	92
3.2.1	Campi scalari e gradiente	92
3.2.2	Campi vettoriali e campi gradiente	94
3.2.3	Forza posizionale e forza conservativa	97
3.3	Equazioni di Lagrange per un punto materiale	102
3.3.1	Conservazione dell'energia, variabili cicliche e funzione di Routh	103
3.4	Il moto centrale	108
3.4.1	L'equazione per r	112
3.4.2	Il problema di Keplero	113
3.4.3	L'orbita del problema di Keplero	118
3.4.4	La terza legge	121
4	Sistemi vincolati e coordinate lagrangiane	123
4.1	Sistemi olonomi	123
4.1.1	Spostamenti virtuali in funzione delle coordinate lagrangiane .	133
4.1.2	Velocità ed energia cinetica	141
4.1.3	Punto vincolato sulla superficie	145
5	Le equazioni di moto per sistemi vincolati	151
5.1	Dinamica di un punto vincolato sulla superficie.	152
5.2	L'equazione simbolica della dinamica	155
5.3	Le equazioni di Lagrange	160
5.3.1	Equazioni di Lagrange e statica del punto materiale sulla superficie	162
5.3.2	Equazioni di Lagrange di prima specie	172
5.4	Risolubilità delle equazioni di Lagrange	174
5.5	Invarianza delle equazioni di Lagrange	175
5.6	Coordinate cicliche	178
5.7	La funzione Hamiltoniana e la conservazione dell'energia	182
5.8	Il teorema di Noether	190
5.9	Equilibrio	195
5.9.1	Stabilità	198
5.10	Piccole Oscillazioni	199
5.10.1	Sistemi con un solo grado di libertà	200
5.10.2	Sistemi con l gradi di libertà	201
5.10.3	Esempio: la catena di oscillatori	211
6	Cinematica dei Sistemi Rigidi	215
6.1	Introduzione	215
6.2	Moti rigidi	215
6.2.1	Primo caso: $\Omega \equiv O$	217
6.2.2	Secondo caso: $\Omega \neq O$	224
6.2.3	Gradi di libertà di un corpo rigido	224
6.3	Formula fondamentale del moto rigido	227

6.4	Angoli di Eulero	238
6.5	Asse istantaneo di moto, rigate del moto	242
6.6	Cinematica relativa: composizione delle velocità	247
6.7	Formula di Poisson	249
6.8	Composizione di moti rigidi	251
6.9	Cinematica relativa: l'accelerazione	253
7	Dinamica Sistemi Rigidi	255
7.1	Il centro di massa	256
7.2	Il momento angolare	259
7.3	Geometria delle masse	259
7.3.1	Momenti d'inerzia	260
7.3.2	Omografia d'inerzia, matrice d'inerzia e terna principale d'inerzia	263
7.3.3	Ellissoide d'inerzia	269
7.3.4	Determinazione della terna principale d'inerzia nel caso di sistemi piani	270
7.3.5	Esempi e complementi	274
7.4	Le equazioni cardinali	283
7.5	Le equazioni cardinali sono sufficienti per determinare il moto dei rigidi	286
7.6	Momento angolare, energia cinetica e seconda equazione cardinale per i sistemi rigidi	289
7.6.1	Momento angolare per un sistema rigido	290
7.6.2	Seconda equazione cardinale per i sistemi rigidi	291
7.6.3	Reazioni vincolari applicate all'asse di rotazione	297
7.6.4	L'energia cinetica	301
7.7	Le precessioni per inerzia	306
7.7.1	Le equazioni di Eulero	307
7.7.2	Risoluzione dell'equazione di Eulero nel caso di precessioni per inerzia	308
7.7.3	Il moto à la Poinsot	313
7.8	Lagrangiana del corpo rigido	314
7.9	Il giroscopio pesante	318
8	Principi variazionali	325
8.1	La brachistocrona	326
8.1.1	La trattazione moderna	328
8.2	L'equazione di Eulero-Lagrange	329
8.2.1	Un integrale primo e ritorno alla brachistocrona	334
8.3	Funzionali dipendenti da l funzioni	339
8.3.1	Massimizzazione vincolata	341
8.4	Il principio di Hamilton	346
8.4.1	Sistemi vincolati ed equazioni di Lagrange di prima specie	349
8.5	Il principio di Jacobi	351
8.5.1	Coordinate cicliche nell'ambito del principio di Hamilton	351
8.5.2	Il tempo come variabile	354
8.5.3	Formulazione del principio di Jacobi	356

9	Il Sistema Canonico	361
9.1	Il Teorema di Liouville	363
9.2	Le parentesi di Poisson	366
9.3	Derivazione variazionale delle equazioni di Hamilton	368

Prefazione

Non c'è miglior modo di onorare la memoria di un Collega prematuramente scomparso che quello di dare continuità al suo impegno didattico pubblicando i testi delle sue lezioni. E bene ha fatto la moglie di Riccardo, Annick, ad affidare il compito di curarne la pubblicazione ad alcuni dei più giovani componenti del gruppo dei cultori delle materie che nel tempo di maggior fulgore si chiamavano fisico-matematiche. Ed è stato per noi – che di quel gruppo abbiamo fatto parte per tanti anni- un onore quello di ricevere la richiesta di scrivere poche righe di presentazione di questo volume.

Un volume che rispecchia pienamente le capacità didattiche di Riccardo, caratterizzate da una intuizione sicura, chiarezza di vedute e grande comunicativa. Il testo tiene fede alle sue convinzioni in ordine alla formazione degli studenti di matematica. Pur essendo un testo destinato agli studenti della laurea triennale e che quindi non riveste le caratteristiche di un trattato specialistico di livello avanzato, la sua impostazione denota una padronanza degli argomenti di meccanica e degli strumenti - di tipo geometrico ed analitico - che hanno contraddistinto anche tutta la produzione scientifica di Riccardo, i suoi articoli di ricerca e, in generale, il suo modo di intendere la matematica applicata.

Infatti gli argomenti di meccanica costituiscono per gli studenti di matematica (ma non solo per essi) la prima occasione di confrontarsi con il “modello matematico”: un insieme di oggetti matematici, di leggi e di relazioni che entro certi limiti è in grado di fornire, proprio attraverso le proprietà matematiche, informazioni sul comportamento di oggetti che matematici non sono. Per la meccanica si tratterà di corpi, che di volta in volta saranno descritti e identificati mentalmente con punti, con sistemi indeformabili di punti, con sistemi continui e così via; ma per altri tipi di modelli si tratterà di cellule, di popolazioni, di prodotti finanziari ecc.

Costruire – e poi studiare - un modello matematico non è cosa agevole; è necessaria una profonda conoscenza degli oggetti che si vogliono modellizzare ed anche la sensibilità di capire quali sono (tutte e sole) le grandezze e i parametri che, in un determinato contesto, devono essere introdotti nel modello per ottenere una adeguata descrizione del fenomeno che si intende studiare. In altre parole, come diceva Einstein “a model should be as simple as possible . . . but not simpler!”

E' come dire che la grande sfida di fronte alla costruzione di un modello matematico per un problema complesso è la ricerca di un compromesso ottimale che colga le componenti principali e sacrifichi quelle meno significative. Ciò richiede la capacità di costruire una visione prospettica del fenomeno, cosa che Riccardo possedeva al massimo grado.

Un altro pregio di questo libro è quello di rendere chiari anche argomenti di per sé molto complessi, talvolta anche ricorrendo ad esempi di esperienza quotidiana; d'altra parte rientrava nello stile di Riccardo la ricerca di continui agganci con applicazioni anche semplici, con i casi-limite più illuminanti, con le considerazioni su possibili analogie tra situazioni apparentemente diverse.

Però sarebbe molto riduttivo confinare queste considerazioni ad aspetti puramente tecnici e non menzionare il fatto che Riccardo era una persona dai molteplici interessi e di una cultura vastissima. E soprattutto piace ricordare il suo spirito arguto e pungente, che rendeva la sua compagnia estremamente piacevole.

Abbiamo trascorso tanti anni nello stesso luogo di lavoro, siamo stati insieme a tanti congressi, abbiamo condiviso l'emozione intellettuale della ricerca e, nei momenti di svago, le escursioni nella letteratura e nella musica di cui era appassionatissimo. La perdita di una persona così lascia un segno indelebile. Siamo riconoscenti ai curatori di questa opera, che, a noi che l'abbiamo conosciuto così bene, riporta alla mente perfino la sua voce. E' bello che attraverso questo libro molti giovani possano ancora fruire della eredità intellettuale di una mente tanto brillante.

Antonio Fasano

Mario Primicerio

Introduzione

Quest'opera si basa sugli *Appunti per il Corso di Sistemi Dinamici*, scritti dal Prof. Riccardo Ricci. Nonostante la prima versione degli *Appunti* risalga al 2005, questi vennero da lui continuamente aggiornati ed ampliati sino al 2012. L'anno successivo, purtroppo, il Prof. Ricci venne a mancare. Gli *Appunti* erano, e lo sono tuttora, il testo di riferimento per gli studenti del corso annuale di Sistemi Dinamici, secondo anno Laurea Triennale in Matematica, dell'Università degli Studi di Firenze.

Gli *Appunti* riflettono non solo i contenuti del corso di Sistemi Dinamici che il Prof. Ricci ha tenuto sin dal 1999, ma anche il modo con cui egli affrontava la disciplina. Il Prof. Ricci, pur partendo dai contenuti della classica Meccanica Razionale, sviluppava un corso che metteva in particolare evidenza gli aspetti analitici, algebrici e geometrici. In tal modo gli studenti venivano continuamente chiamati ad applicare (spesso anche in modo critico) i metodi matematici che gli stessi avevano acquisito nei corsi di Analisi, Geometria e Algebra. E tale caratteristica è, in sostanza, il motivo per cui gli *Appunti* siano sempre stati molto apprezzati dagli studenti.

Alla scomparsa del Prof. Ricci, avvenuta il 19 agosto 2013, è spontaneamente nata l'idea di pubblicare gli *Appunti* (disponibili sino al 2016 come semplici dispense del corso di Sistemi Dinamici), aggiungendovi quel materiale che veniva generalmente presentato agli studenti durante le esercitazioni. Il progetto è stato subito accolto dalla vedova del Prof. Ricci, La Prof.ssa Annick Magnier, che ha messo a disposizione tutto il materiale del Prof. Ricci ai curatori di quest'opera, i quali hanno provveduto a riorganizzarlo ed ampliarlo. L'idea era quella di giungere ad un testo che fosse un manuale per il corso di Sistemi Dinamici basato sugli *Appunti* e che, soprattutto, seguisse lo "spirito" con cui il Prof. Ricci aveva scritto i suoi *Appunti*.

Partendo quindi dall'ultima versione degli *Appunti*, quella datata 10 aprile 2012, è stata intrapresa un'opera di rivisitazione ed ampliamento del testo originario che ha portato al presente manuale. In particolare, rispetto all'originale versione degli *Appunti*, è stato aggiunto il capitolo 1 dove vengono richiamati alcuni concetti di algebra lineare e geometria frequentemente utilizzati e viene introdotta la teoria dei momenti (parte che veniva usualmente illustrata nell'ambito delle esercitazioni). Il materiale del capitolo 2, pur essendo stato riorganizzato, è rimasto essenzialmente invariato. Sono stati aggiunti gli esempi e le sezioni 2.6.3 e 2.10. Nel capitolo 3 è stato aggiunto il paragrafo *Forze conservative* e la sezione 3.3.1 dedicata alle variabili cicliche ed alla funzione di Routh. I capitoli 4 e 5 sono stati ampliati introducendo tutta la parte dedicata al punto materiale vincolato su una superficie e la sezione 5.6. Per quanto riguarda il capitolo 6, i primi quattro paragrafi sono stati riscritti, seguendo però la stessa impostazione degli *Appunti*. Il capitolo 7 è quello che è stato più modificato rispetto al

capitolo 6 degli *Appunti*. Nello specifico è stata aggiunta tutta la sezione dedicata alla *Geometria delle masse* (argomento che veniva trattato in dettaglio nelle esercitazioni del corso). Nel capitolo dedicato ai *Principi Variazionali* è stata aggiunta la sezione 8.3 mentre il paragrafo *Il principio di minima azione* è stato riscritto. Il capitolo 9 è rimasto essenzialmente invariato.

I curatori desiderano ringraziare la Prof.ssa Annick Magnier per aver messo a disposizione il materiale del Prof. Ricci e per il costante incoraggiamento senza il quale difficilmente si sarebbe potuti giungere al completamento di quest'opera. Un particolare ringraziamento va al Prof. Giuseppe Anichini, Direttore del Dipartimento di Matematica ed Informatica "U. Dini", che ha accolto il progetto con favore ed al Prof. Graziano Gentili per i numerosi suggerimenti. Si desidera infine ringraziare Il Dott. Fulvio Guatelli, Direttore editoriale della Florence University Press, per il prezioso aiuto durante la fase editoriale.

Capitolo 1

Introduzione allo spazio affine, alle coordinate curvilinee ed ai sistemi di vettori applicati

In questo capitolo daremo dei brevi richiami di algebra lineare per passare alla definizione di spazio affine, o meglio di spazio affine euclideo. Si postula infatti che lo spazio che percepiamo, dove appunto si trovano i corpi oggetto della meccanica, sia uno spazio affine tridimensionale.

1.1 Spazi vettoriali su \mathbb{R}

Sia V un insieme i cui elementi saranno chiamati **vettori**, o **vettori liberi**, dotato di due operazioni:

- la somma, denotata con $+$, che associa a due vettori $\mathbf{u}, \mathbf{v} \in V$, un terzo vettore $\mathbf{u} + \mathbf{v} \in V$;
- il prodotto di un vettore per uno scalare (un numero reale), che associa al vettore $\mathbf{v} \in V$ e allo scalare $\lambda \in \mathbb{R}$ un altro vettore denotato $\lambda\mathbf{v}$.

V si dice uno **spazio vettoriale reale** o **spazio vettoriale sui reali**, se le suddette operazioni verificano le seguenti proprietà:

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \forall \mathbf{u}, \mathbf{v} \in V$.
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
- In V esiste un vettore, denotato con 0 , detto **vettore nullo**, tale che $\mathbf{v} + 0 = \mathbf{v}, \forall \mathbf{v} \in V$.
- Dato un qualsiasi vettore $\mathbf{v} \in V$, esiste un unico vettore, denotato con $-\mathbf{v}$, e detto **opposto** di \mathbf{v} , tale che $\mathbf{v} + (-\mathbf{v}) = 0$.
- $\forall \mathbf{v} \in V$, si ha $1\mathbf{v} = \mathbf{v}$, e $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}, \forall \alpha, \beta \in \mathbb{R}$.

- $\forall \mathbf{u}, \mathbf{v} \in V$, e $\forall \alpha, \beta \in \mathbb{R}$, valgono le seguenti eguaglianze $(\alpha + \beta) \mathbf{v} = \alpha \mathbf{v} + \beta \mathbf{v}$, e $\alpha(\mathbf{u} + \mathbf{v}) = \alpha \mathbf{u} + \alpha \mathbf{v}$.

Uno spazio vettoriale reale di fondamentale importanza è \mathbb{R}^n , cioè l'insieme delle n -uple in colonna di numeri reali, dette anche **vettori colonna**,

$$\mathbb{R}^n = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \text{ tali che } x_i \in \mathbb{R}, i = 1, 2, \dots, n \right\},$$

dotato dell'operazione di somma e prodotto per scalari così definite

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \mathbf{x} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Il vettore nullo $\mathbf{0}$ è la colonna costituita da tutti zero. Talvolta con \mathbb{R}^n si intende l'insieme delle n -uple in riga di numeri reali (x_1, x_2, \dots, x_n) , dette **vettori riga**, dove la somma ed il prodotto per gli scalari sono definite componente per componente. Quando sarà importante distinguere i due modi di scrittura, specificheremo fra vettore colonna di \mathbb{R}^n , o vettore riga di \mathbb{R}^n .

Diamo adesso la definizione di **sottospazio vettoriale**. Un sottoinsieme $W \subset V$ spazio vettoriale reale, si dice un sottospazio di V se:

- $\mathbf{0} \in W$.
- $\forall \mathbf{u}, \mathbf{w} \in W$, allora $\mathbf{u} + \mathbf{w} \in W$.
- Se $\mathbf{u} \in W$, allora $\lambda \mathbf{u} \in W, \forall \lambda \in \mathbb{R}$.

Vediamo adesso la definizione di **combinazione lineare**. Dati m vettori $\mathbf{v}_1, \dots, \mathbf{v}_m$ di uno spazio vettoriale V e m scalari $\alpha_1, \dots, \alpha_m$, si dice combinazione lineare degli m vettori con gli m scalari, il vettore $\mathbf{w} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_m \mathbf{v}_m$. I numeri reali $\alpha_1, \dots, \alpha_m$ si dicono **coefficienti della combinazione lineare**.

Diremo che gli m vettori $\mathbf{v}_1, \dots, \mathbf{v}_m$ di V sono **linearmente indipendenti** se la combinazione lineare $\alpha_1 \mathbf{v}_1 + \dots + \alpha_m \mathbf{v}_m$, dà il vettore nullo soltanto quando tutti i coefficienti della combinazione sono nulli. Se invece accade che la combinazione lineare $\alpha_1 \mathbf{v}_1 + \dots + \alpha_m \mathbf{v}_m$ si annulla anche se non tutti i coefficienti sono nulli allora gli m vettori $\mathbf{v}_1, \dots, \mathbf{v}_m$ si dicono **linearmente dipendenti**.

Si dice che gli n vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ di V **generano lo spazio vettoriale** V se $\forall \mathbf{u} \in V$ esistono n scalari β_1, \dots, β_n per cui

$$\mathbf{u} = \beta_1 \mathbf{v}_1 + \dots + \beta_n \mathbf{v}_n,$$

cioè ogni vettore di V si può esprimere come combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$. I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ vengono detti **generatori** di V . Lo spazio V viene poi detto **finitamente generato** se ammette un insieme finito di generatori. E' evidente che se

$\{v_1, \dots, v_n\}$ sono un insieme di generatori, aggiungendo a tale insieme altri vettori otteniamo sempre un insieme di generatori: infatti non è detto che i generatori siano tra loro linearmente indipendenti.

Un esempio di spazio finemente generato è \mathbb{R}^n , in cui un insieme di generatori è

$$\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}. \quad (1.1)$$

Se $\{v_1, \dots, v_n\}$ sono un insieme di generatori di V e sono linearmente indipendenti allora si dice **base** di V . Una proprietà fondamentale è data dal seguente

Teorema 1.1.1 *I vettori $\{v_1, \dots, v_n\}$ sono una base di V se e solo se ogni vettore di V si può esprimere in modo univoco come combinazione lineare di v_1, \dots, v_n .*

Dim. Se l'insieme $\{v_1, \dots, v_n\}$ costituisce una base allora, dalla definizione di base sappiamo che, $\forall u \in V$, possiamo scrivere

$$u = \beta_1 v_1 + \dots + \beta_n v_n.$$

Vediamo che tale rappresentazione è unica. Infatti, se non lo fosse avremo anche

$$u = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

con, in generale, $\alpha_i \neq \beta_i$, per qualche i . Ma allora avremo anche

$$(\beta_1 - \alpha_1) v_1 + \dots + (\beta_n - \alpha_n) v_n = 0,$$

che comporta, data la lineare indipendenza dei vettori v_1, \dots, v_n , $\beta_i = \alpha_i, \forall i = 1, 2, \dots, n$, che quindi contraddice l'ipotesi $\alpha_i \neq \beta_i$, per qualche i .

Supponiamo adesso che dato un qualunque vettore $u \in V$, questo si possa scrivere *in modo unico* come combinazione lineare di v_1, \dots, v_n . Dalla definizione di generatori deduciamo che $\{v_1, \dots, v_n\}$ costituisce un insieme di generatori di V . Per mostrare che $\{v_1, \dots, v_n\}$ è un base bisogna provare che v_1, \dots, v_n sono linearmente indipendenti, ovvero che

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0 \implies \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

Ora, siccome $0 = 0v_1 + \dots + 0v_n$, dall'unicità della rappresentazione ricaviamo immediatamente

$$\alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

□

Una base rappresenta l'insieme minimo per generare V . Infatti, tornando all'insieme finito di generatori di V , è facile rendersi conto che in tale insieme possono stare anche vettori che non sono necessari in quanto combinazione lineari di altri vettori di V . Il seguente teorema (la cui dimostrazione viene omessa) ci garantisce che, se V è finemente generato, ovvero è generato da un numero finito di generatori, allora da essi si riesce ad escludere i vettori "superflui" fino a giungere ad una base.

Teorema 1.1.2 *Dato V spazio vettoriale reale finemente generato, allora V ammette almeno una base.*

□

E' quindi evidente che ogni insieme di generatori contiene almeno una base. Ora, uno spazio vettoriale ammette infinite basi, tuttavia ogni base è caratterizzata dallo stesso numero di vettori. Infatti, supponendo che V sia finemente generato, è possibile dimostrare il seguente risultato

Teorema 1.1.3 *Ogni base di V possiede lo stesso numero di vettori.*

Dim. La dimostrazione non è difficile ma abbastanza lunga. La riassumeremo quindi nei suoi punti essenziali. Si comincia col far vedere che se $A = \{v_1, \dots, v_m\}$ è un insieme di vettori linearmente indipendenti e $U = \{u_1, \dots, u_p\}$ un insieme di generatori¹ di V , allora $p \geq m$. Tale dimostrazione si fa per assurdo supponendo $p < m$. Se così fosse allora si identifica un sottoinsieme $A_p \subset A$, costituito da p vettori $A_p = \{v_1, \dots, v_p\}$, e si esprimono in funzione dei vettori di U . Se si introduce quindi la matrice \mathbb{A} , $n \times n$, dei coefficienti delle combinazioni lineari, risulta chiaro, sfruttando il teorema di Rouché-Capelli, che $\det \mathbb{A} \neq 0$, e quindi \mathbb{A} è invertibile. Allora possiamo esprimere i vettori di U come combinazione lineare dei vettori di A_p , dove \mathbb{A}^{-1} sarà la matrice dei coefficienti delle combinazioni lineari. Ora, per ipotesi, i vettori di U generano tutto V , e quindi anche v_{p+1} potrà essere espresso come combinazione lineare di u_1, \dots, u_p , i quali, a loro volta, sono combinazioni lineari v_1, \dots, v_p . Quindi v_{p+1} è combinazione lineare di v_1, \dots, v_p , il che è assurdo dato che i vettori di A sono per ipotesi linearmente indipendenti. Di conseguenza deve essere $p \geq m$.

Adesso supponiamo di avere due basi B_1 e B_2 , di V , tali che B_1 ha m vettori mentre B_2 ne ha p , e che $m > p$, per esempio (il caso $m < p$ è del tutto analogo). Allora, leggiamo B_1 come l'insieme di vettori linearmente indipendenti e B_2 come l'insieme di generatori. Da quanto visto prima deve essere $p \geq m$, e quindi assumere $m > p$ porta ad un assurdo. Si conclude quindi che B_1 e B_2 hanno necessariamente lo stesso numero di vettori.

□

Diamo quindi la definizione di **dimensione** di uno spazio V finemente generato: *la dimensione di V è il numero di vettori di una qualunque base di V (tale numero è sempre lo stesso). Se la dimensione di V è n , scriveremo brevemente $\dim V = n$.*

L'insieme (1.1) è una base di \mathbb{R}^n , ed è usualmente detta **base canonica**.

Vediamo adesso il concetto di **componenti**, o **coordinate, di un vettore**. Sia V uno spazio vettoriale di dimensione n . Sia $\{v_1, \dots, v_n\}$, una base di V . Dato un qualunque vettore $u \in V$, potremo scrivere

$$u = u_1 v_1 + \dots + u_n v_n,$$

dove i coefficienti della combinazione lineare sono univocamente determinati. Tali coefficienti sono detti **componenti**, o **coordinate, di u rispetto alla base $\{v_1, \dots, v_n\}$.**

¹Ricordiamo che stiamo lavorando su spazi finemente generati.

Le componenti di \mathbf{u} vengono quindi espresse tramite la n -upla

$$\mathbf{u} \rightarrow \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},$$

scrivendo anche

$$\mathbf{u} = u_1\mathbf{v}_1 + u_2\mathbf{v}_2 + \dots + u_n\mathbf{v}_n. \tag{1.2}$$

Quindi, una volta specificata la base di V , ad ogni vettore \mathbf{u} possiamo associare una ed una sola² n -upla: risulta quindi definita una **corrispondenza biunivoca** fra lo spazio V e lo spazio \mathbb{R}^n . In tal senso, nel seguito di quest'opera, intenderemo, salvo che non diversamente specificato, con la stessa lettera in grassetto \mathbf{u} sia il vettore dello spazio vettoriale V sia la n -upla di \mathbb{R}^n che lo rappresenta (rispetto ad un data base). *Ovviamente le componenti di un vettore dipendono dalla base scelta, cambiando base allo stesso vettore \mathbf{u} viene associato un diverso vettore colonna.* In altri termini, il vettore \mathbf{u} *prescinde dalla base* mentre la n -upla *dipende dalla base scelta*. In effetti, la scrittura (1.2), specificando esplicitamente sia le componenti che la base, mette ben in luce questo stretto legame fra la n -upla di \mathbb{R}^n e la base prescelta.

Le considerazioni appena illustrate mostrano che vale il seguente

Teorema 1.1.4 *Tutti gli spazi vettoriali ad n dimensioni (n fissato) su \mathbb{R} , sono isomorfi fra loro e quindi sono isomorfi ad \mathbb{R}^n .*

□

In altri termini, il teorema precedente afferma che se V e U sono due spazi vettoriali su \mathbb{R} , aventi la stessa dimensione n , si può stabilire una corrispondenza biunivoca fra i vettori di V e quelli di U , cioè $\mathbf{v} \longleftrightarrow \mathbf{u}$, se $\mathbf{v} \in V$ e $\mathbf{u} \in U$. Inoltre, se \mathbf{v}, \mathbf{v}' , sono vettori V e \mathbf{u}, \mathbf{u}' , è la corrispondente coppia di vettori di U , cioè $\mathbf{v} \longleftrightarrow \mathbf{u}$, e $\mathbf{v}' \longleftrightarrow \mathbf{u}'$, allora $\mathbf{v} + \mathbf{v}' \longleftrightarrow \mathbf{u} + \mathbf{u}'$, e $\alpha\mathbf{v} \longleftrightarrow \alpha\mathbf{u}$, per ogni $\alpha \in \mathbb{R}$. Si può quindi affermare che, astrattamente, U e V sono lo stesso spazio.

1.2 Spazi vettoriali euclidei

Allo scopo di definire le nozioni di lunghezza di un vettore e di angolo tra vettori si introduce una nuova operazione tra vettori: il prodotto scalare. Se V è uno spazio vettoriale reale di dimensione finita, definiamo **prodotto scalare** in V la seguente operazione

$$\cdot : V \times V \longrightarrow \mathbb{R}$$

²Infatti se al vettore \mathbf{u} si potesse associare due n -uple, cioè

$$\mathbf{u} = \begin{cases} \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n, \\ \beta_1\mathbf{v}_1 + \dots + \beta_n\mathbf{v}_n, \end{cases}$$

facendo la differenza otterremmo

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + \dots + (\alpha_n - \beta_n)\mathbf{v}_n = 0,$$

che può essere soddisfatta soltanto se $(\alpha_i - \beta_i) = 0, \forall i = 1, \dots, n$, dal momento che gli n vettori \mathbf{v}_i sono linearmente indipendenti.

che gode delle seguenti proprietà:

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}, \forall \mathbf{u}, \mathbf{v} \in V$.
2. $\mathbf{v} \cdot \mathbf{v} > 0$, se $\mathbf{v} \neq 0$, ed è zero solo nel caso in cui $\mathbf{v} = 0$.
3. \cdot è lineare sia nel primo che nel secondo argomento, ovvero

$$\begin{aligned} (\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) \cdot \mathbf{u} &= \alpha_1 \mathbf{v}_1 \cdot \mathbf{u} + \alpha_2 \mathbf{v}_2 \cdot \mathbf{u}, \\ \mathbf{v} \cdot (\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2) &= \alpha_1 \mathbf{v} \cdot \mathbf{u}_1 + \alpha_2 \mathbf{v} \cdot \mathbf{u}_2. \end{aligned}$$

Lo spazio vettoriale V dotato del prodotto scalare viene detto **spazio vettoriale euclideo**, che usualmente si denota (V, \cdot) . Talvolta, sfruttando la linearità sul primo e secondo argomento e la simmetria, si dice anche che il **prodotto scalare è una forma bilineare** (proprietà 3) **simmetrica** (proprietà 1), **definita positiva** (proprietà 2).

Nel caso in cui $V = \mathbb{R}^n$, l'applicazione $\cdot : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, così definita

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n, \quad (1.3)$$

soddisfa le proprietà del prodotto scalare. Dunque (\mathbb{R}^n, \cdot) è uno spazio vettoriale euclideo. Osserviamo inoltre che se leggiamo l' n -upla \mathbf{x} come una matrice $n \times 1$, e denotiamo con $\mathbf{y}^T = (y_1, y_2, \dots, y_n)$, ovvero l' n -upla riga che corrisponde ad una matrice $1 \times n$, il prodotto scalare può anche essere visto come prodotto righe per colonne fra matrici

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y}^T \mathbf{x} = \mathbf{x}^T \mathbf{y}.$$

Dato (V, \cdot) spazio vettoriale euclideo, l'applicazione $\| \cdot \| : V \rightarrow \mathbb{R}$, che opera così:

$$\| \mathbf{v} \| = \sqrt{\mathbf{v} \cdot \mathbf{v}},$$

viene detta **modulo**, o **norma**, o **lunghezza**, del vettore \mathbf{v} . Talvolta *la norma del vettore \mathbf{v} potrà anche essere denotata con $|\mathbf{v}|$ anziché con $\|\mathbf{v}\|$* .

In \mathbb{R}^n , la norma (o modulo o lunghezza), associata al prodotto scalare standard dato dalla (1.3), di una n -upla \mathbf{x} è

$$\| \mathbf{x} \| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Dalla proprietà del prodotto scalare si deducono facilmente le seguenti proprietà:

- $\| \mathbf{v} \| > 0$, se $\mathbf{v} \neq 0$, ed è zero solo nel caso in cui $\mathbf{v} = 0$.
- $\forall \lambda \in \mathbb{R}, \|\lambda \mathbf{v}\| = |\lambda| \|\mathbf{v}\|$.

Due importanti risultati sono riassunti nel seguente

Teorema 1.2.1 *Se (V, \cdot) è uno spazio vettoriale euclideo (reale), allora $\forall \mathbf{u}, \mathbf{v} \in V$, si ha:*

1. $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$, *disuguaglianza di Cauchy-Schwarz.*
2. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, *disuguaglianza triangolare o di Minkowski.*

Dim. Dimostriamo prima la disuguaglianza di Cauchy-Schwarz. E' chiaro che la disuguaglianza è verificata se almeno uno dei due vettori è nullo. Supponiamo quindi che siano entrambi non nulli. Consideriamo $\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$, con $\alpha, \beta \in \mathbb{R}$,

$$\mathbf{w} \cdot \mathbf{w} = (\alpha\mathbf{u} + \beta\mathbf{v}) \cdot (\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha^2 \|\mathbf{u}\|^2 + \beta^2 \|\mathbf{v}\|^2 + 2\alpha\beta\mathbf{u} \cdot \mathbf{v} \geq 0.$$

Prendendo adesso $\alpha = \|\mathbf{v}\|^2$ e $\beta = -\mathbf{u} \cdot \mathbf{v}$, si ottiene

$$\|\mathbf{v}\|^4 \|\mathbf{u}\|^2 + (\mathbf{u} \cdot \mathbf{v})^2 \|\mathbf{v}\|^2 - 2\|\mathbf{v}\|^2 (\mathbf{u} \cdot \mathbf{v})^2 = \|\mathbf{v}\|^2 \left(\|\mathbf{v}\|^2 \|\mathbf{u}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \right) \geq 0.$$

Siccome $\mathbf{v} \neq 0$, possiamo dividere per $\|\mathbf{v}\|^2$, ed ottenere così la disuguaglianza

$$\|\mathbf{v}\|^2 \|\mathbf{u}\|^2 \geq (\mathbf{u} \cdot \mathbf{v})^2, \implies |\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Per quanto riguarda la disuguaglianza triangolare, sfruttando la disuguaglianza di Cauchy-Schwarz abbiamo

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} \\ &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\| \|\mathbf{v}\| = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2. \end{aligned}$$

Notiamo che se $\mathbf{u} \cdot \mathbf{v} = 0$, allora abbiamo $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, che altro non è che il teorema di Pitagora.

□

A partire dalla disuguaglianza di Cauchy-Schwarz possiamo definire l'angolo compreso fra due vettori. Infatti, se $\mathbf{u} \neq 0$ e $\mathbf{v} \neq 0$, abbiamo

$$-1 \leq \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1.$$

Per cui, possiamo definire l'angolo $\theta \in [0, \pi]$ fra i vettori \mathbf{u} e \mathbf{v} come quell'angolo il cui coseno è dato da

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}. \tag{1.4}$$

Dalla (1.4) scaturisce questa importante relazione fra norme dei vettori \mathbf{u} e \mathbf{v} , prodotto scalare e angolo compreso

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta.$$

Da cui segue la definizioni di ortogonalità fra vettori: diremo che \mathbf{u} e \mathbf{v} sono **ortogonali**, e scriveremo $\mathbf{u} \perp \mathbf{v}$, se

$$\mathbf{u} \cdot \mathbf{v} = 0, \iff \text{l'angolo } \theta \text{ fra } \mathbf{u} \text{ e } \mathbf{v} \text{ è retto.}$$

Diamo adesso la definizione di **base ortonormale**. Se (V, \cdot) è uno spazio vettoriale di dimensione n , una base $\{e_1, \dots, e_n\}$ si dice ortonormale, o **base canonica**, se

$$\|e_i\| = 1, \quad \forall i = 1, 2, \dots, n \quad \text{e} \quad e_i \cdot e_j = 0, \quad \forall i \neq j,$$

ovvero³

$$e_i \cdot e_j = \delta_{ij} = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases} \quad (1.5)$$

I vettori delle basi ortonormali si denotano in genere con la lettera e , e vengono usualmente detti **versori**. Notiamo che la base (1.1) di \mathbb{R}^n è ortonormale se il prodotto scalare è dato dalla (1.3).

Dato (V, \cdot) spazio vettoriale euclideo, si può provare che esiste un metodo che permette di determinare sempre una base ortonormale. Ciò prova, in generale, l'esistenza di una base ortonormale. Tale metodo viene detto procedimento di *ortonormalizzazione di Gram-Schmidt*.

Se si scrive l'espressione del prodotto scalare utilizzando le componenti dei vettori *rispetto ad una base ortonormale* su uno spazio vettoriale Euclideo di dimensione n , si ottiene la stessa espressione del prodotto scalare standard su \mathbb{R}^n , cioè quello definito dalla (1.3). In altri termini abbiamo questo

Teorema 1.2.2 *Se (V, \cdot) è uno spazio vettoriale euclideo, e $\{e_1, \dots, e_n\}$ è una base ortonormale, allora dati comunque due vettori u e w*

$$u = u_1 e_1 + \dots + u_n e_n, \Rightarrow a \ u \ \text{si associa l}'n\text{-upla } \underline{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},$$

$$w = w_1 e_1 + \dots + w_n e_n, \Rightarrow a \ w \ \text{si associa l}'n\text{-upla } \underline{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

si ha

$$u \cdot w = \underbrace{\underline{u} \cdot \underline{w}}_{\substack{\text{p. scalare} \\ \text{in } \mathbb{R}^n \text{ definito} \\ \text{dalla (1.3)}}} = \underline{u}^T \underline{w} = \underline{w}^T \underline{u}. \quad (1.6)$$

Dim. La dimostrazione è facilissima. Infatti, ricordando la (1.5),

$$\begin{aligned} u \cdot w &= \left(\sum_{i=1}^n u_i e_i \right) \cdot \left(\sum_{j=1}^n w_j e_j \right) = \sum_{i=1}^n \sum_{j=1}^n u_i \underbrace{(e_i \cdot e_j)}_{\delta_{ij}} w_j \\ &= \sum_{i=1}^n u_i w_i. \end{aligned}$$

³La δ_{ij} è detta δ di Kronecker. Leopold Kronecker (1823 – 1891) è stato un matematico tedesco.

Vale anche il viceversa. Se $\forall \mathbf{u}, \mathbf{w} \in V$ vale la (1.6) allora la base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è ortonormale. La dimostrazione è lasciata per esercizio.

□

Due conseguenze di questo risultato sono le seguenti:

- **le componenti di un qualunque vettore \mathbf{v} rispetto ad una base ortonormale $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ hanno questa forma particolare**

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{e}_1) \mathbf{e}_1 + (\mathbf{v} \cdot \mathbf{e}_2) \mathbf{e}_2 + \dots + (\mathbf{v} \cdot \mathbf{e}_n) \mathbf{e}_n .$$

In altri termini, se $\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \dots + v_n \mathbf{e}_n$, allora

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \stackrel{\text{Se } B \text{ è}}{\equiv} \text{ortonormale} \begin{pmatrix} \mathbf{v} \cdot \mathbf{e}_1 \\ \mathbf{v} \cdot \mathbf{e}_2 \\ \vdots \\ \mathbf{v} \cdot \mathbf{e}_n \end{pmatrix} . \tag{1.7}$$

Infatti, se moltiplichiamo scalarmente \mathbf{v} per \mathbf{e}_1 , poi per \mathbf{e}_2 , e così via otteniamo

$$\mathbf{v} \cdot \mathbf{e}_1 = (v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \dots + v_n \mathbf{e}_n) \cdot \mathbf{e}_1 \stackrel{(1.5)}{=} v_1 ,$$

$$\mathbf{v} \cdot \mathbf{e}_2 = (v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \dots + v_n \mathbf{e}_n) \cdot \mathbf{e}_2 \stackrel{(1.5)}{=} v_2 ,$$

⋮

$$\mathbf{v} \cdot \mathbf{e}_n = (v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \dots + v_n \mathbf{e}_n) \cdot \mathbf{e}_n \stackrel{(1.5)}{=} v_n .$$

Gli scalari $(\mathbf{v} \cdot \mathbf{e}_i)$, $i = 1, 2, \dots, n$, vengono detti **proiezioni ortogonali di \mathbf{v} sui vettori della base**.

- Alla base ortonormale $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di V corrisponde la base canonica (1.1) di \mathbb{R}^n .

Rimarchiamo ancora che la (1.6) vale solo nel caso in cui B sia una base ortonormale. Ci chiediamo dunque: se (V, \cdot) è uno spazio vettoriale euclideo di dimensione n , ed in esso è data una base $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, non ortonormale, come si traduce il prodotto scalare fra vettori nel prodotto fra n -uple? Se $\mathbf{u} = \left(\sum_{i=1}^n u_i \mathbf{v}_i\right)$, $\mathbf{w} = \left(\sum_{i=1}^n w_i \mathbf{v}_i\right)$, potremo scrivere

$$\begin{aligned} \mathbf{u} \cdot \mathbf{w} &= \left(\sum_{i=1}^n u_i \mathbf{v}_i\right) \cdot \left(\sum_{j=1}^n w_j \mathbf{v}_j\right) = \sum_{i=1}^n \sum_{j=1}^n u_i (\mathbf{v}_i \cdot \mathbf{v}_j) w_j \\ &= \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \cdot \underbrace{\begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 & \cdots & \mathbf{v}_1 \cdot \mathbf{v}_n \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 & \cdots & \mathbf{v}_2 \cdot \mathbf{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_n \cdot \mathbf{v}_1 & \mathbf{v}_n \cdot \mathbf{v}_2 & \cdots & \mathbf{v}_n \cdot \mathbf{v}_n \end{pmatrix}}_{\mathbb{G}} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} , \tag{1.8} \end{aligned}$$

cioè $\mathbf{u} \cdot \mathbf{w} = \underline{\mathbf{u}} \cdot \mathbb{G} \underline{\mathbf{w}} \equiv \underline{\mathbf{u}}^T \mathbb{G} \underline{\mathbf{w}}$, dove $\underline{\mathbf{u}}, \underline{\mathbf{w}} \in \mathbb{R}^n$. La matrice \mathbb{G} viene detta **matrice metrica** o **matrice associata alla forma bilineare**, oppure anche **tensore metrico**. Evidentemente, una volta definito il prodotto scalare, \mathbb{G} dipende dalla base prescelta: se infatti $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ fosse ortonormale avremo $\mathbb{G} \equiv \mathbb{I}$. Notiamo che la norma di \mathbf{u} è

$$\|\mathbf{u}\|^2 = \underline{\mathbf{u}} \cdot \mathbb{G} \underline{\mathbf{u}} \equiv \underline{\mathbf{u}}^T \mathbb{G} \underline{\mathbf{u}}.$$

Le principali proprietà di \mathbb{G} sono riassunte nel seguente

Teorema 1.2.3 *Se (V, \cdot) è uno spazio vettoriale euclideo e $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, cui corrisponde la matrice \mathbb{G} , allora:*

- \mathbb{G} è simmetrica;
- $\det \mathbb{G} \neq 0$, ovvero \mathbb{G} è invertibile.

Dim. La simmetria è una conseguenza banale della simmetria del prodotto scalare. Proviamo che $\det \mathbb{G} \neq 0$. Infatti se fosse $\det \mathbb{G} = 0$, allora il sistema lineare

$$\mathbb{G} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0,$$

avrebbe soluzioni non nulle. Se ne consideriamo una qualsiasi e ad essa associamo il vettore

$$\mathbf{w} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n,$$

questo è sicuramente non nullo, $\mathbf{w} \cdot \mathbf{w} = \|\mathbf{w}\|^2 > 0$. Del resto

$$\|\mathbf{w}\|^2 = (x_1, x_2, \dots, x_n) \underbrace{\mathbb{G} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{=0} = 0,$$

che è assurdo.

□

Notiamo inoltre che l'identificazione (1.7) **non è in generale vera se B non è una base ortonormale**. Supponendo al solito $\mathbf{u} = u_1 \mathbf{v}_1 + u_2 \mathbf{v}_2 + \dots + u_n \mathbf{v}_n$, abbiamo

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v}_1 &= \sum_{i=1}^n u_i (\mathbf{v}_1 \cdot \mathbf{v}_i), \\ &\vdots \\ \mathbf{u} \cdot \mathbf{v}_n &= \sum_{i=1}^n u_i (\mathbf{v}_n \cdot \mathbf{v}_i), \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} \mathbf{u} \cdot \mathbf{v}_1 \\ \vdots \\ \mathbf{u} \cdot \mathbf{v}_n \end{pmatrix} = \mathbb{G} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}. \quad (1.9)$$

In altre parole, il generico vettore \mathbf{u} è rappresentato, rispetto alla *stessa base* B , da due diverse n -uple, cioè

$$\mathbf{u} \longrightarrow \left\{ \begin{array}{l} \left(\begin{array}{c} u_1 \\ \vdots \\ u_n \end{array} \right), \text{ tale che } \mathbf{u} = \sum_{i=1}^n u_i \mathbf{v}_i, \\ \left(\begin{array}{c} \mathbf{u} \cdot \mathbf{v}_1 \\ \vdots \\ \mathbf{u} \cdot \mathbf{v}_n \end{array} \right). \end{array} \right.$$

E' chiaro che se B è ortonormale i due vettori di \mathbb{R}^n coincidono. Per passare dall'uno all'altro si deve utilizzare la matrice \mathbb{G} che, come abbiamo provato, è invertibile. Le due n -uple si distinguono l'una dall'altra tramite diverse denominazioni: l' n -upla

$\left(\begin{array}{c} \mathbf{u} \cdot \mathbf{v}_1 \\ \vdots \\ \mathbf{u} \cdot \mathbf{v}_n \end{array} \right)$ viene detta n -upla delle **componenti**, o **coordinate**, **covarianti** di \mathbf{u} ,

mentre $\left(\begin{array}{c} u_1 \\ \vdots \\ u_n \end{array} \right)$ viene detta n -upla delle **componenti**, o **coordinate**, **controvarianti**

di \mathbf{u} . La relazione che consente di passare dalle componenti controvarianti a quelle covarianti è la (1.9). L'inversa, cioè

$$\left(\begin{array}{c} u_1 \\ \vdots \\ u_n \end{array} \right) = \mathbb{G}^{-1} \left(\begin{array}{c} \mathbf{u} \cdot \mathbf{v}_1 \\ \vdots \\ \mathbf{u} \cdot \mathbf{v}_n \end{array} \right),$$

permette invece di calcolare le componenti controvarianti a partire da quelle covarianti.

Nella figura 1.1 sono riportate graficamente sia le componenti controvarianti che quelle covarianti del medesimo vettore, **nel caso particolare in cui i vettori della base $\mathbf{v}_1, \mathbf{v}_2$ siano normalizzati**. In tal caso, infatti, le componenti covarianti di un vettore coincidono con le proiezioni ortogonali del vettore sugli assi della base. Questo però non è vero se i vettori della base non sono normalizzati. Tuttavia, anche se impropriamente, le componenti covarianti di un vettore sono sovente lette come le "proiezioni" del vettore sugli assi della base.

Nel corso di quest'opera non staremo quasi mai a distinguere fra componenti covarianti o controvarianti di un vettore. Questo perché molto spesso avremo a che fare con basi ortonormali. Ovviamente, quando scriveremo

$$\mathbf{u} = u_1 \mathbf{v}_1 + u_2 \mathbf{v}_2 + \dots + u_n \mathbf{v}_n,$$

le componenti $u_i, i = 1, 2, \dots, n$, sono quelle controvarianti, che, per basi non ortonormali, **non** coincidono le "proiezioni" di \mathbf{u} sui vettori della base.

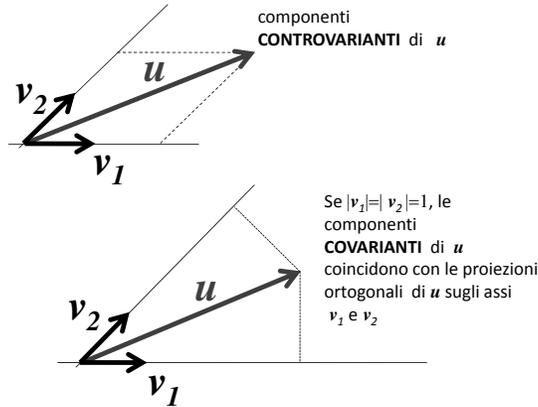


Figura 1.1: L'interpretazione geometrica delle componenti covarianti del vettore u è semplice nel caso in cui i vettori della base siano normalizzati, cioè $|v_1| = |v_2| = 1$. Infatti, in tal caso $u \cdot v_i = |u| \cos \theta_i$, $i = 1, 2$, essendo θ_i l'angolo compreso fra u e v_i . In questo caso quindi le componenti covarianti altro non sono che le proiezioni ortogonali di u sugli assi v_1 e v_2 . Evidentemente se $|v_i| \neq 1$, $i = 1, 2$, allora $u \cdot v_i = |u| |v_i| \cos \theta_i$, e dunque le componenti covarianti $u \cdot v_i$, $i = 1, 2$, non coincidono con le proiezioni ortogonali del vettore u sugli assi v_i , $i = 1, 2$.

1.3 Trasformazioni di basi ortonormali

In uno spazio vettoriale euclideo (V, \cdot) sono date due basi ortonormali $B = \{e_1, \dots, e_n\}$ e $B' = \{e'_1, \dots, e'_n\}$, sì che il generico vettore x è rappresentato da due differenti n -uple, cioè

$$x \longrightarrow \begin{cases} \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} & \text{nella base } B, \\ \underline{x}' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} & \text{nella base } B'. \end{cases}$$

Vogliamo adesso stabilire che relazione c'è fra la n -upla \underline{x} e la n -upla \underline{x}' . Esprimiamo ciascun versore di B nella base B' ,

$$e_i = \sum_{j=1}^N (e_i \cdot e'_j) e'_j = \sum_{j=1}^N a_{ij} e'_j, \quad i = 1, \dots, n, \quad (1.10)$$

dove

$$a_{ij} = e_i \cdot e'_j, \quad i, j = 1, \dots, n, \quad (1.11)$$

ciò introduciamo la matrice

$$\mathbb{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 \cdot \mathbf{e}'_1 & \cdots & \mathbf{e}_1 \cdot \mathbf{e}'_n \\ \mathbf{e}_2 \cdot \mathbf{e}'_1 & \cdots & \mathbf{e}_2 \cdot \mathbf{e}'_n \\ \vdots & & \vdots \\ \mathbf{e}_n \cdot \mathbf{e}'_1 & \cdots & \mathbf{e}_n \cdot \mathbf{e}'_n \end{pmatrix}, \quad (1.12)$$

detta **matrice del cambiamento di base da B' in B** . Esprimiamo ora \mathbf{x} nella base B

$$\mathbf{x} = \sum_{i=1}^N x_i \mathbf{e}_i, \quad \text{con } x_i = \mathbf{x} \cdot \mathbf{e}_i,$$

e nella base B'

$$\mathbf{x} = \sum_{j=1}^N x'_j \mathbf{e}'_j, \quad \text{con } x'_j = \mathbf{x} \cdot \mathbf{e}'_j.$$

Riprendendo la (1.10)

$$x_i = \mathbf{x} \cdot \mathbf{e}_i = \mathbf{x} \cdot \left(\sum_{j=1}^N a_{ij} \mathbf{e}'_j \right) = \sum_{j=1}^N a_{ij} \underbrace{(\mathbf{x} \cdot \mathbf{e}'_j)}_{x'_j} = \sum_{j=1}^N a_{ij} x'_j,$$

otteniamo

$$x_i = \sum_{j=1}^N a_{ij} x'_j, \quad \Rightarrow \quad \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}}_{\mathbb{A}} \underbrace{\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}}_{\mathbf{x}'}$$

La proprietà fondamentale di \mathbb{A} è quella di essere una **matrice ortogonale**, cioè

$$\mathbb{A}^T \mathbb{A} = \mathbb{A} \mathbb{A}^T = \mathbb{I}.$$

Infatti, dal momento che B e B' sono ortonormali si ha

$$\begin{aligned} \delta_{kj} &= \mathbf{e}_k \cdot \mathbf{e}_j = \left(\sum_{s=1}^N a_{ks} \mathbf{e}'_s \right) \cdot \left(\sum_{i=1}^N a_{ji} \mathbf{e}'_i \right) \\ &= \sum_{s=1}^N \sum_{i=1}^N a_{ks} a_{ji} \underbrace{\mathbf{e}_s \cdot \mathbf{e}_i}_{\delta_{si}} = \sum_{i=1}^N a_{ki} a_{ji} \\ &= \sum_{i=1}^N a_{ki} a_{ij}^T, \end{aligned}$$

cioè $\mathbb{A}\mathbb{A}^T = \mathbb{I}$. Del resto

$$\begin{aligned} \delta_{mn} &= e'_m \cdot e'_n = \left[\sum_{s=1}^N \underbrace{(e'_n \cdot e_s)}_{a_{sn}} e_s \right] \cdot \left[\sum_{r=1}^N \underbrace{(e'_m \cdot e_r)}_{a_{rm}} e_r \right] \\ &= \sum_{s=1}^N \sum_{r=1}^N a_{sn} a_{rm} \underbrace{e_s \cdot e_r}_{\delta_{sr}} = \sum_{r=1}^N a_{rn} a_{rm} \\ &= \sum_{r=1}^N a_{nr}^T a_{rm}, \end{aligned}$$

e quindi $\mathbb{A}^T \mathbb{A} = \mathbb{I}$.

1.4 Spazio affine e coordinate curvilinee

Definizione 1.4.1 Uno spazio affine è una terna (A, V, F) , dove: A è un insieme i cui elementi sono detti punti; V è uno spazio vettoriale sul campo reale di dimensione finita (V viene detto spazio vettoriale sottostante lo spazio affine); F è un'applicazione

$$F : A \times A \rightarrow V$$

$$A \times A \ni (P, Q) \xrightarrow{F} F(P, Q) = \mathbf{v} \in V, \quad \begin{array}{l} \text{il vettore } \mathbf{v} \text{ si denota anche} \\ \text{con } (P - Q) \end{array}$$

che gode delle seguenti proprietà:

1. \forall coppia $P \in A$, e $\mathbf{v} \in V$, esiste uno ed un solo $Q \in A$, tale che $\underbrace{(P - Q)}_{F(P, Q)} = \mathbf{v}$.
2. Dati comunque tre punti P, Q, S in A , vale

$$\underbrace{(P - Q)}_{F(P, Q)} + \underbrace{(Q - S)}_{F(Q, S)} = \underbrace{(P - S)}_{F(P, S)}.$$

La dimensione di uno spazio affine A è quella di V . Il generico vettore \mathbf{v} di V viene anche detto **vettore libero**, mentre la coppia $(P, \mathbf{v}) \in A \times V$, viene spesso detta **vettore applicato nel punto P** .

Se in A si fissa un punto O , allora A si identifica con V , nel senso che tutti i punti P di A sono in corrispondenza biunivoca con i vettori $F(P, O) = (P - O) \in V$.

Osserviamo che uno spazio vettoriale può sempre pensarsi come spazio affine. Gli elementi di A , cioè i punti, sono i vettori. Inoltre, se P corrisponde al vettore \mathbf{v} e Q corrisponde a \mathbf{u} , il vettore associato a $P - Q$ è $\mathbf{v} - \mathbf{u}$.

Definizione 1.4.2 Un **riferimento affine**, (detto anche **riferimento cartesiano**) nello spazio affine⁴ A , è costituito da:

⁴Con il termine spazio affine A , si intende sempre la terna (A, V, F) .

- un punto O , detto origine;
- una base $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$, di vettori di V . Per semplicità la base verrà denotata nel seguito anche come $\{\mathbf{w}_i\}$.

Il riferimento cartesiano in A , si denota con $\{O, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$, o, più semplicemente, con $\{O, \mathbf{w}_i\}$. Quindi, dato un riferimento cartesiano su A , il generico punto $P \in A$ si identifica con la n -upla di \mathbb{R}^n

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

che, come sappiamo, rappresenta le **componenti del vettore** $(P - O)$, o anche **coordinate del punto** P , rispetto alla base $\{\mathbf{w}_i\}$ prescelta,

$$P - O = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \dots + x_n\mathbf{w}_n.$$

In sostanza, quando parliamo di spazio affine dotato di riferimento cartesiano, intervengono tre spazi:

- Lo spazio affine A , o spazio dei punti P .
- Lo spazio vettoriale V .
- Lo spazio delle coordinate \mathbb{R}^n .

L'importanza della teoria degli spazi affini nella meccanica è giustificata dal fatto che si assume che lo spazio dove ci troviamo, ovvero lo spazio delle nostre percezioni, è uno spazio affine di dimensione tre.

Definizione 1.4.3 *Uno spazio affine (A, V, F) , si dice euclideo se lo spazio vettoriale V è euclideo, ovvero in esso è definito un prodotto scalare.*

Negli spazi affini euclidei si può definire la **distanza fra punti** $P, Q \in A$, come il modulo del vettore $(P - Q)$, che spesso si denota con $|P - Q|$, anziché con $\|P - Q\|$.

Consideriamo adesso uno spazio affine euclideo, in cui $\{O, \mathbf{e}_i\}$, è un **riferimento cartesiano ortogonale**, cioè la base $\{\mathbf{e}_i\}$, è una base ortonormale. Come detto, ogni punto $P \in A$ è individuato da un n -upla, i cui elementi sono le componenti del vettore $P - O$, che, come sappiamo, coincidono anche con le proiezione ortogonali di $P - O$ sui vettori della base. Se Q è un dominio di \mathbb{R}^n , ovvero $Q \subset \mathbb{R}^n$, si considera questa applicazione $\mathbf{x} : Q \rightarrow \mathfrak{D} \subset \mathbb{R}^n$,

$$Q \ni \mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} \xrightarrow{\mathbf{x}} \mathbf{x}(\mathbf{q}) = \begin{pmatrix} x_1(q_1, q_2, \dots, q_n) \\ x_2(q_1, q_2, \dots, q_n) \\ \vdots \\ x_n(q_1, q_2, \dots, q_n) \end{pmatrix} \in \mathfrak{D} \subset \mathbb{R}^n.$$

Il codominio \mathfrak{D} in generale è un sottoinsieme di \mathbb{R}^n , e ad esso corrisponderà un dominio \mathfrak{D}_A nello spazio affine A . Se l'applicazione $\mathbf{x}(\mathbf{q})$ gode delle seguenti proprietà:

1. $\mathbf{x}(\mathbf{q}) \in C^1(Q)$;
2. $\forall \mathbf{q} \in Q$, l'applicazione è invertibile;

allora si dice che $\mathbf{x}(\mathbf{q})$ definisce un nuovo sistema di coordinate su \mathfrak{D} , dette **coordinate curvilinee**.

La condizione 2 è verificata se la matrice $n \times n$,

$$\mathbb{J} = \begin{pmatrix} \frac{\partial x_1}{\partial q_1} & \frac{\partial x_1}{\partial q_2} & \cdots & \frac{\partial x_1}{\partial q_n} \\ \frac{\partial x_2}{\partial q_1} & \frac{\partial x_2}{\partial q_2} & \cdots & \frac{\partial x_2}{\partial q_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial q_1} & \frac{\partial x_n}{\partial q_2} & \cdots & \frac{\partial x_n}{\partial q_n} \end{pmatrix},$$

ha rango massimo, ovvero $\forall \mathbf{q} \in Q$ si ha $\det \mathbb{J} \neq 0$. In altri termini, le n funzioni

$$\begin{cases} x_1 = x_1(q_1, q_2, \dots, q_n), \\ x_2 = x_2(q_1, q_2, \dots, q_n), \\ \vdots \\ x_n = x_n(q_1, q_2, \dots, q_n), \end{cases}$$

rappresentano un sistema di coordinate curvilinee in \mathfrak{D} , se gli n vettori di V ,

$$\frac{\partial \mathbf{x}}{\partial q_1}, \frac{\partial \mathbf{x}}{\partial q_2}, \dots, \frac{\partial \mathbf{x}}{\partial q_n},$$

le cui componenti rispetto alla base $\{\mathbf{e}_i\}$ sono i vettori di \mathbb{R}^n

$$\begin{pmatrix} \frac{\partial x_1}{\partial q_1} \\ \frac{\partial x_2}{\partial q_1} \\ \vdots \\ \frac{\partial x_n}{\partial q_1} \end{pmatrix}, \begin{pmatrix} \frac{\partial x_1}{\partial q_2} \\ \frac{\partial x_2}{\partial q_2} \\ \vdots \\ \frac{\partial x_n}{\partial q_2} \end{pmatrix}, \dots, \begin{pmatrix} \frac{\partial x_1}{\partial q_n} \\ \frac{\partial x_2}{\partial q_n} \\ \vdots \\ \frac{\partial x_n}{\partial q_n} \end{pmatrix},$$

sono linearmente indipendenti⁵ $\forall \mathbf{q} \in Q$. In tal caso diremo che $\frac{\partial \mathbf{x}}{\partial q_i}$, $i = 1, 2, \dots, n$, formano un **referimento associato alle coordinate curvilinee** \mathbf{q} , e quindi una base dello spazio vettoriale V soggiacente lo spazio affine A . Tale base però cambia da punto a punto e perciò viene detta **base locale** o **base puntuale**. In generale i vettori $\frac{\partial \mathbf{x}}{\partial q_i}$, $i = 1, 2, \dots, n$, si denotano con \mathbf{u}_i , e la base puntuale con $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$, o

⁵Infatti, scrivendo la matrice \mathbf{J} come

$$\mathbb{J} = \begin{bmatrix} \frac{\partial \mathbf{x}}{\partial q_1} & \frac{\partial \mathbf{x}}{\partial q_2} & \cdots & \frac{\partial \mathbf{x}}{\partial q_n} \end{bmatrix},$$

la condizione di rango massimo implica l'indipendenza dei vettori $\frac{\partial \mathbf{x}}{\partial q_i}$.

semplicemente con $\{\mathbf{u}_i\}$. Diremo poi che le coordinate curvilinee \mathbf{q} sono **ortogonali** se $\mathbf{u}_i \cdot \mathbf{u}_j = 0$, se $i \neq j$. In tal caso si usano introdurre i **versori della base locale** come

$$\boldsymbol{\kappa}_i = \frac{\partial \mathbf{x}}{\partial q_i} / \left| \frac{\partial \mathbf{x}}{\partial q_i} \right|, \quad i = 1, 2, \dots, n,$$

tali che $\boldsymbol{\kappa}_i \cdot \boldsymbol{\kappa}_j = \delta_{ij}$.

Una proprietà importante dei vettori \mathbf{u}_i , è quella di essere punto per punto tangenti alle rispettive curve coordinate. Infatti, se fissiamo i valori di $n - 1$ coordinate e lasciamo variare una sola coordinata (per esempio fissiamo $(q_1, q_2, \dots, q_{i-1}, q_{i+1}, \dots, q_n)$ e lasciamo “libera” la q_i), otteniamo una curva parametrizzata dalla coordinata variabile. Il vettore $\frac{\partial \mathbf{x}}{\partial q_i}$ è tangente alla curva parametrizzata dalla coordinata q_i . Da notare che, nel caso specifico della base ortonormale $\{\mathbf{e}_i\}$, il vettore unitario \mathbf{e}_i è tangente all’ i -esimo asse coordinato.

In particolare, dato un generico vettore $\mathbf{v} \in V$, questo si potrà anche esprimere sia in termini della base $\{\mathbf{e}_i\}$, che in termini della base locale $\{\mathbf{u}_i\}$. Quindi, riferendoci a quest’ultima, la n -upla

$$\begin{pmatrix} \mathbf{v} \cdot \mathbf{u}_1 \\ \mathbf{v} \cdot \mathbf{u}_2 \\ \vdots \\ \mathbf{v} \cdot \mathbf{u}_n \end{pmatrix}$$

fornisce le componenti covarianti di \mathbf{v} rispetto alla base puntuale $\{\mathbf{u}_i\}$, cioè

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{v} \cdot \mathbf{u}_n) \mathbf{u}_n.$$

In accordo con la (1.8) si definisce la matrice metrica associata al sistema di coordinate curvilinee come

$$\mathbb{G} = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_n \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_n \\ \vdots & & \vdots \\ \mathbf{u}_n \cdot \mathbf{u}_1 & \cdots & \mathbf{u}_n \cdot \mathbf{u}_n \end{pmatrix},$$

le cui componenti variano da punto a punto. Se quindi utilizziamo la base $\{\mathbf{u}_i\}$ per esprimere i vettori \mathbf{v} e \mathbf{w} tramite le loro componenti controvarianti, cioè $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{u}_i$ e $\mathbf{w} = \sum_{i=1}^n w_i \mathbf{u}_i$, avremo

$$\begin{aligned} \mathbf{v} \cdot \mathbf{w} &= (v_1 \quad \cdots \quad v_n) \mathbb{G} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\ &= (w_1 \quad \cdots \quad w_n) \mathbb{G} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}. \end{aligned}$$

Esempio 1.4.1 Consideriamo il piano affine A con un riferimento cartesiano ortogonale $\{O, \mathbf{e}_x, \mathbf{e}_y\}$. Poniamo $Q = (0, +\infty) \times [0, 2\pi) \subset \mathbb{R}^2$, $q_1 = r$, $q_2 = \theta$, e

consideriamo l'applicazione

$$\mathbf{x}(\mathbf{q}) = \begin{pmatrix} x(r, \theta) = r \cos \theta \\ y(r, \theta) = r \sin \theta \end{pmatrix}$$

che porta il dominio Q in $\mathfrak{D}_A = A - \{O\}$. Determiniamo i vettori $\mathbf{u}_r = \frac{\partial \mathbf{x}}{\partial r}$, $\mathbf{u}_\theta = \frac{\partial \mathbf{x}}{\partial \theta}$, e verifichiamo la loro indipendenza $\forall (r, \theta) \in Q$. Infatti, dando direttamente le loro componenti,

$$\mathbf{u}_r = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad \mathbf{u}_\theta = \begin{pmatrix} -r \sin \theta \\ r \cos \theta \end{pmatrix},$$

abbiamo $\det \mathbb{J} = r \neq 0$, $\forall (r, \theta) \in Q$. La base locale $\{\mathbf{u}_r, \mathbf{u}_\theta\}$ è ortogonale: $\mathbf{u}_r \cdot \mathbf{u}_\theta = 0$. Inoltre \mathbf{u}_θ è tangente alle linee coordinate $r = \text{cost.}$, e θ variabile in $[0, 2\pi)$, mentre \mathbf{u}_r è tangente alle linee coordinate $\theta = \text{cost.}$, r variabile in $(0, \infty)$. \mathbf{u}_r è quindi diretto radialmente.

Esempio 1.4.2 Consideriamo lo spazio tridimensionale con il riferimento cartesiano ortogonale $\{O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, e le coordinate sferiche $q_1 = r$, $q_2 = \varphi$ longitudine e $q_3 = \theta$, co-latitudine (angolo che il vettore $(P - O)$ forma con l'asse z , vedi figura 1.2), definite nel dominio $Q = (0, +\infty) \times [0, 2\pi) \times (0, \pi) \subset \mathbb{R}^3$. Consideriamo poi l'applicazione

$$\mathbf{x}(\mathbf{q}) = \begin{pmatrix} x(r, \varphi, \theta) = r \sin \theta \cos \varphi \\ y(r, \varphi, \theta) = r \sin \theta \sin \varphi \\ z(r, \varphi, \theta) = r \cos \theta \end{pmatrix}.$$

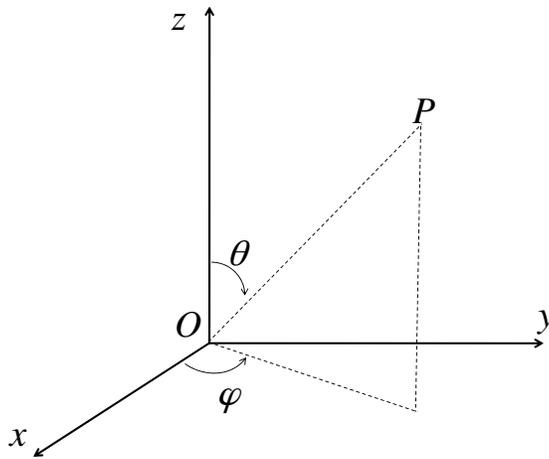
che porta il dominio Q in $\mathfrak{D}_A = A - \{\text{asse } z\}$. Si verifica facilmente che $\det \mathbb{J} = -r^2 \sin \theta \neq 0$, $\forall (r, \varphi, \theta) \in Q$. La base locale (non normalizzata) è

$$\mathbf{u}_r = \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}, \quad \mathbf{u}_\varphi = \begin{pmatrix} -r \sin \theta \sin \varphi \\ r \sin \theta \cos \varphi \\ 0 \end{pmatrix}, \quad \mathbf{u}_\theta = \begin{pmatrix} r \cos \theta \cos \varphi \\ r \cos \theta \sin \varphi \\ -r \sin \theta \end{pmatrix},$$

mentre quella normalizzata è $\boldsymbol{\kappa}_r = \mathbf{u}_r$, $\boldsymbol{\kappa}_\varphi = \frac{\mathbf{u}_\varphi}{r \sin \theta}$, $\boldsymbol{\kappa}_\theta = \frac{\mathbf{u}_\theta}{r}$. È facile verificare che la base locale è ortogonale. Possiamo esprimere il vettore $(P - O)$ nella base locale. Abbiamo

$$\begin{aligned} (P - O) &= r \sin \theta \cos \varphi \mathbf{e}_x + r \sin \theta \sin \varphi \mathbf{e}_y + r \cos \theta \mathbf{e}_z \\ &= r \underbrace{(\sin \theta \cos \varphi \mathbf{e}_x + \sin \theta \sin \varphi \mathbf{e}_y + \cos \theta \mathbf{e}_z)}_{\mathbf{u}_r}, \end{aligned}$$

Figura 1.2: Coordinate sferiche.



cioè

$$(P - O) \rightarrow \mathbf{x} = \begin{cases} x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z, & \text{coordinate cartesiane,} \\ r\boldsymbol{\kappa}_r, & \text{coordinate sferiche.} \end{cases} \quad (1.13)$$

In particolare, $\|P - O\| = r$.

1.5 Prodotto vettoriale e sistemi di vettori applicati

1.5.1 Il prodotto vettoriale

Si comincia con la definizione di prodotto vettoriale fra terne.

Definizione 1.5.1 Se $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, si definisce **prodotto vettoriale**, o **prodotto esterno**, delle terne \mathbf{a} e \mathbf{b} il vettore

$$\mathbf{a} \wedge \mathbf{b} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}. \quad (1.14)$$

Introducendo la base canonica (1.1) $\mathbf{e}_1, \mathbf{e}_2$ ed \mathbf{e}_3 , si ha formalmente

$$\mathbf{a} \wedge \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}. \quad (1.15)$$

E' facile dimostrare che il prodotto vettoriale così come definito dalla (1.14) gode delle seguenti proprietà:

- $\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$;
- $\forall \alpha, \beta \in \mathbb{R}, (\alpha \mathbf{a} + \beta \mathbf{b}) \wedge \mathbf{c} = \alpha (\mathbf{a} \wedge \mathbf{c}) + \beta (\mathbf{b} \wedge \mathbf{c})$;
- $\mathbf{a} \wedge \mathbf{b}$ è ortogonale sia ad \mathbf{a} che a \mathbf{b} . Infatti $(\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{a} = (\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{b} = 0$;
- $\mathbf{e}_1 \wedge \mathbf{e}_2 = \mathbf{e}_3, \mathbf{e}_3 \wedge \mathbf{e}_1 = \mathbf{e}_2, \mathbf{e}_2 \wedge \mathbf{e}_3 = \mathbf{e}_1$;
- dati \mathbf{a}, \mathbf{b} e \mathbf{c} qualunque

$$[\mathbf{a} \wedge \mathbf{b}] \cdot \mathbf{c} = [\mathbf{c} \wedge \mathbf{a}] \cdot \mathbf{b} = [\mathbf{b} \wedge \mathbf{c}] \cdot \mathbf{a} ; \quad (1.16)$$

- il prodotto vettoriale soddisfa la cosiddetta identità di Jacobi

$$\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) + \mathbf{b} \wedge (\mathbf{c} \wedge \mathbf{a}) + \mathbf{c} \wedge (\mathbf{a} \wedge \mathbf{b}) = 0. \quad (1.17)$$

Consideriamo adesso lo spazio affine euclideo A di dimensione 3. Sia $\{O, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$, un riferimento cartesiano ortonormale centrato in O . Diremo che il **riferimento è positivamente orientato** o che la **terna $\{O, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ è levogira**, o **destrorsa**, se $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ sono orientati come pollice, indice e medio della mano destra. In altri termini, la terna è levogira se la rotazione di $\pi/2$ che porta il primo versore sul secondo è antioraria se vista dal semispazio verso cui punta il terzo versore. In figura 1.3 sono riportate una terna ortonormale levogira ed una non levogira.

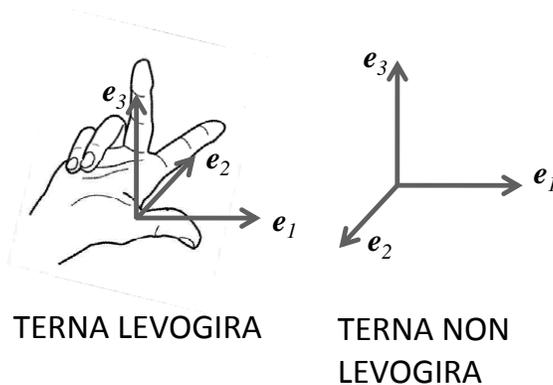


Figura 1.3: Terna ortonormale levogira e terna ortonormale non levogira.

Due riferimenti cartesiani ortonormali $\{O, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ e $\{Q, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si diranno **concordemente orientati** se la matrice del cambiamento di base \mathbb{A} ha determinante uguale ad 1, $\det \mathbb{A} = 1$. I riferimenti $\{O, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ e $\{Q, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ non sono concordemente orientati se $\det \mathbb{A} = -1$.

Nota 1.5.1 Con l'introduzione del prodotto vettoriale nello spazio affine di dimensione 3 abbiamo introdotto il concetto di destra e sinistra. Questi sono concetti puramente

relativi, almeno nel nostro contesto. Quello che ha un significato assoluto è la differenza fra le due orientazioni. Nel caso della retta, per esempio, è possibile specificare due versi, ma definirne uno come positivo (perché diretto da “sinistra verso destra”) e l’altro come negativo si può fare solo se ci riferiamo ad un uomo che osserva la retta. Prescindendo dall’osservatore concreto quello che si può stabilire è soltanto la differenza dei due versi: dire che uno è positivo e l’altro è negativo è una pura convenzione. Nello spazio quindi esistono due tipi di riferimento cartesiano ortogonale: quelli di un tipo sono collegati fra loro da matrici con determinante uguale ad 1, mentre quelli di tipo diverso sono collegati da trasformazioni con determinante uguale a -1 . In astratto non si può quindi dire quali siano gli uni e quali gli altri a meno di non ricorrere ad un osservatore concreto che confronta la base cartesiana con la sua mano destra.

Consideriamo lo spazio affine euclideo di dimensione 3. In esso vale la seguente interpretazione geometrica del prodotto vettoriale: dati tre punti A , B e C , il prodotto vettoriale $(A - C) \wedge (B - C)$, ha la direzione ortogonale al piano individuato dai vettori $(A - C)$ e $(B - C)$. Il verso del vettore $(A - C) \wedge (B - C)$ è tale che la rotazione (di angolo inferiore a π) che porta $(A - C)$ su $(B - C)$ sia antioraria se vista dal semispazio che contiene il vettore $(A - C) \wedge (B - C)$. Infine, per quel che riguarda il modulo, vale

$$|(A - C) \wedge (B - C)| = |(A - C)||B - C| \sin \varphi, \quad (1.18)$$

dove $\varphi \in [0, \pi]$, è l’angolo compreso fra i vettori $(A - C)$ e $(B - C)$. Per dimostrare tali affermazioni, consideriamo

$$\frac{(A - C)}{|(A - C)|} \wedge \frac{(B - C)}{|(B - C)|},$$

cioè il prodotto vettoriale fra versori. Senza perdere di generalità possiamo considerare la terna ortonormale levogira $\{A, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ centrata in A , dove

$$\begin{aligned} \mathbf{w}_1 &= \frac{(A - C)}{|(A - C)|}, \\ \mathbf{w}_2 &= \text{versore del piano individuato dai vettori} \\ &\quad (A - C) \text{ e } (B - C), \text{ ortogonale a } \mathbf{w}_1, \end{aligned}$$

e dove \mathbf{w}_3 è dato da $\mathbf{w}_3 = \mathbf{w}_1 \wedge \mathbf{w}_2$. Le componenti di $\frac{(B - C)}{|(B - C)|}$ rispetto a tale base sono

$$\frac{(B - C)}{|(B - C)|} = \cos \varphi \mathbf{w}_1 + \sin \varphi \mathbf{w}_2, \quad \varphi \in [0, 2\pi),$$

per cui

$$\frac{(A - C)}{|(A - C)|} \wedge \frac{(B - C)}{|(B - C)|} = \mathbf{w}_1 \wedge (\cos \varphi \mathbf{w}_1 + \sin \varphi \mathbf{w}_2) = \sin \varphi \mathbf{w}_3,$$

ovvero $(A - C) \wedge (B - C) = |(A - C)||B - C| \sin \varphi \mathbf{w}_3$. Se adesso ne prendiamo il modulo otteniamo esattamente la (1.18), dove $\varphi \in [0, \pi]$ indica l’angolo fra i due vettori.

1.5.2 Vettori applicati e momento rispetto ad un polo

Sia, al solito, A spazio affine di dimensione 3 nel quale lavoriamo e sia V il soggiacente spazio vettoriale. Abbiamo già incontrato nella sezione 1.4 la definizione di vettore applicato: è la coppia $(P, \mathbf{v}) \in A \times V$.

Definizione 1.5.2 Si dice **momento di un vettore applicato** (P, \mathbf{v}) rispetto ad un punto O , il vettore

$$\mathbf{m}(O) = (P - O) \wedge \mathbf{v}.$$

Il punto O si dice **polo** o **centro di riduzione**.

Ricordando la (1.18) e osservando la figura 1.4, si ha

$$|\mathbf{m}(O)| = |\mathbf{v}| \underbrace{|(P - O)| \sin \varphi}_b = b |\mathbf{v}|,$$

dove b si dice braccio.

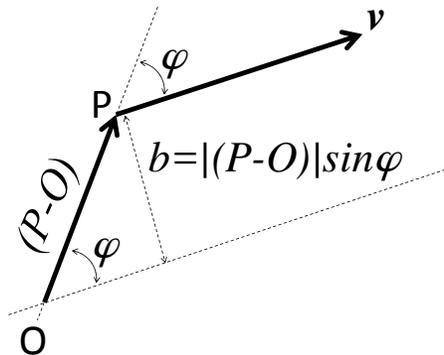


Figura 1.4: Momento del vettore applicato (P, \mathbf{v}) rispetto al polo O .

Il momento ha le seguenti proprietà:

- $\mathbf{m}(O) = \mathbf{m}(O')$, se $(O - O')$ è parallelo a \mathbf{v} . Infatti

$$\begin{aligned} \mathbf{m}(O') &= (P - O') \wedge \mathbf{v} = \mathbf{m}(O) = [(P - O) + (O - O')] \wedge \mathbf{v} \\ &= (P - O) \wedge \mathbf{v} + \underbrace{(O - O') \wedge \mathbf{v}}_{=0 \text{ poiché } (O - O') \parallel \mathbf{v}} = \mathbf{m}(O). \end{aligned}$$

Se si definisce **retta d'applicazione** del vettore applicato (P, \mathbf{v}) , la retta passante per P e parallela a \mathbf{v} , il risultato appena illustrato ci dice che il *momento di un vettore applicato rispetto ad un polo O , non cambia se si sposta \mathbf{v} lungo la propria retta di applicazione.*

- In generale, se $(O - O')$ non è parallelo a \mathbf{v}

$$\mathbf{m}(O') = \mathbf{m}(O) + (O - O') \wedge \mathbf{v}. \quad (1.19)$$

La (1.19) è nota come formula di trasposizione dei momenti.

Un sistema di N vettori applicati è un insieme

$$\mathcal{S} = \{(P_i, \mathbf{v}_i) \in A \times V : i = 1, \dots, N\}.$$

Definizione 1.5.3 Si dice *risultante* del sistema \mathcal{S} il vettore

$$\mathbf{R} = \sum_{i=1}^N \mathbf{v}_i,$$

e momento risultante rispetto al centro di riduzione O

$$\mathbf{m}(O) = \sum_{i=1}^N \mathbf{m}_i(O) = \sum_{i=1}^N (P_i - O) \wedge \mathbf{v}_i.$$

Se cambiamo polo e consideriamo O' , applicando la (1.19) si ottiene

$$\begin{aligned} \mathbf{m}(O') &= \sum_{i=1}^N \mathbf{m}_i(O') = \sum_{i=1}^N [\mathbf{m}_i(O) + (O - O') \wedge \mathbf{v}_i] \\ &= \mathbf{m}(O) + (O - O') \wedge \underbrace{\sum_{i=1}^N \mathbf{v}_i}_{\mathbf{R}} = \mathbf{m}(O) + (O - O') \wedge \mathbf{R}. \end{aligned} \quad (1.20)$$

Quindi $\mathbf{m}(O') = \mathbf{m}(O)$, se $\mathbf{R} = 0$, oppure se $(O - O')$ è parallela ad \mathbf{R} . Inoltre, proprio dalla (1.20) discende una proprietà notevole

$$\mathbf{m}(O') \cdot \mathbf{R} = \mathbf{m}(O) \cdot \mathbf{R} + \underbrace{[(O - O') \wedge \mathbf{R}] \cdot \mathbf{R}}_{=0} = \mathbf{m}(O) \cdot \mathbf{R}.$$

Se quindi si introduce la quantità

$$\mathfrak{S} = \mathbf{m}(O) \cdot \frac{\mathbf{R}}{|\mathbf{R}|},$$

detta **invariante scalare**, questa è indipendente dal polo O . In particolare, possiamo sempre scomporre $\mathbf{m}(O)$ lungo il versore $\frac{\mathbf{R}}{|\mathbf{R}|}$

$$\mathbf{p} = \left(\mathbf{m}(O) \cdot \frac{\mathbf{R}}{|\mathbf{R}|} \right) \frac{\mathbf{R}}{|\mathbf{R}|} = \mathfrak{S} \frac{\mathbf{R}}{|\mathbf{R}|},$$

e nella rimanente parte $\mathbf{n}(O) = \mathbf{m}(O) - \mathbf{p} = \mathbf{m}(O) - \mathfrak{S} \frac{\mathbf{R}}{|\mathbf{R}|}$, che risulta ortogonale a \mathbf{p} . Infatti

$$\begin{aligned} \mathbf{n}(O) \cdot \mathbf{p} &= \mathfrak{S} \mathbf{n}(O) \cdot \frac{\mathbf{R}}{|\mathbf{R}|} = \mathfrak{S} \left(\mathbf{m}(O) - \mathfrak{S} \frac{\mathbf{R}}{|\mathbf{R}|} \right) \cdot \frac{\mathbf{R}}{|\mathbf{R}|} \\ &= \mathfrak{S} (\mathfrak{S} - \mathfrak{S}) = 0. \end{aligned}$$

Il vettore \mathbf{p} , detto **invariante vettoriale**, non dipende dal polo ma soltanto dal sistema di vettori applicati \mathcal{S} (cioè dai vettori \mathbf{v}_i e dalla loro disposizione geometrica nello spazio). Inoltre,

$$\mathbf{m}(O) = \mathbf{p} + \mathbf{n}(O), \quad \Rightarrow \quad |\mathbf{m}(O)| = \sqrt{|\mathbf{n}(O)|^2 + \mathfrak{S}^2}. \quad (1.21)$$

Quindi, partendo da un generico punto O , si vuole trovare, se esiste, il luogo geometrico dei punti Q per i quali il momento risultante abbia modulo minimo, cioè $|\mathbf{n}(Q)| = 0$. Applicando la (1.20) ed imponendo $\mathbf{m}(Q) = \mathbf{p}$, abbiamo

$$\mathbf{m}(Q) = \mathbf{m}(O) + (O - Q) \wedge \mathbf{R} = \mathbf{p},$$

ovvero, supponendo O centro del sistema di riferimento ed indicando con $\mathbf{x} = (Q - O)$ il vettore delle coordinate di Q ,

$$\mathbf{x} \wedge \mathbf{R} + \mathbf{p} = \mathbf{m}(O), \quad (1.22)$$

che è l'equazione di una retta parallela ad \mathbf{R} . Tale retta è detta **asse centrale**. Quindi, per tutti i punti Q che soddisfano la (1.22), cioè che giacciono sull'asse centrale, si ha $\mathbf{m}(Q) = \mathbf{p}$ (v. figura 1.5).

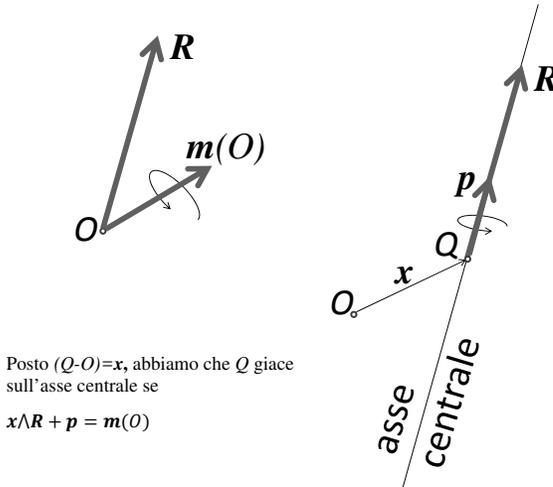


Figura 1.5: Asse centrale.

Un particolare sistema di vettori applicata è la **coppia**.

Definizione 1.5.4 Una **coppia** è un sistema di due vettori applicati

$$\mathcal{S} = \{(P, \mathbf{v}), (Q, -\mathbf{v})\}.$$

Poiché la risultante di una coppia è nulla, il momento $\mathbf{m}(O)$ della coppia *non* dipende dal polo O . In particolare, riferendoci alla figura 1.6

$$|\mathbf{m}(O)| = d |\mathbf{v}|.$$

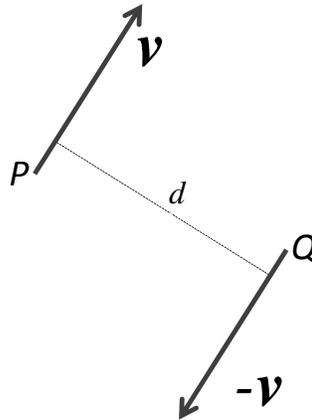


Figura 1.6: Schema di una coppia.

1.5.3 Sistemi equivalenti di vettori applicati

Definizione 1.5.5 Consideriamo due sistemi di vettori applicati

$$S = \{(P_i, \mathbf{v}_i), i = 1, \dots, N\},$$

e

$$S' = \{(A_j, \mathbf{u}_j), j = 1, \dots, M\},$$

di risultanti \mathbf{R} e \mathbf{R}' , e momento risultante $\mathbf{m}(O)$ e $\mathbf{m}'(O)$, con O fissato. Diremo che S e S' sono **equivalenti** se

$$\mathbf{R} = \mathbf{R}', \quad \mathbf{m}(O) = \mathbf{m}'(O). \quad (1.23)$$

Osserviamo subito che se la (1.23) è vera per un generico punto O , allora è vera per tutti i punti dello spazio affine. Infatti, applicando la (1.20)

$$\begin{aligned} \mathbf{m}(Q) &= \mathbf{m}(O) + (O - Q) \wedge \mathbf{R}, \\ \mathbf{m}'(Q) &= \mathbf{m}'(O) + (O - Q) \wedge \mathbf{R}', \end{aligned}$$

che, in virtù della (1.23), risultano esser ancora uguali.

In particolare, due sistemi di vettori applicati S e S' sono equivalenti se possiamo trasformarli uno nell'altro mediante le seguenti operazioni:

1. rimpiazzare due o più vettori applicati in un punto con la loro risultante;
2. scomporre un vettore nelle somma di più vettori;
3. cancellare due vettori uguali ed opposti che sono applicati nello stesso punto;
4. applicare allo stesso punto due vettori uguali e contrari;

5. spostare un vettore lungo la propria retta d'applicazione.

Utilizzando tali operazione è facile mostrare che, dato un vettore applicato (P, v) ed un qualunque centro di riduzione O , il sistema (P, v) è equivalente, nel senso della definizione 1.5.5, ad un vettore v applicato in O e ad una coppia di momento $m(O) = (P - O) \wedge v$. La dimostrazione è schematicamente riportata nella figura 1.7.

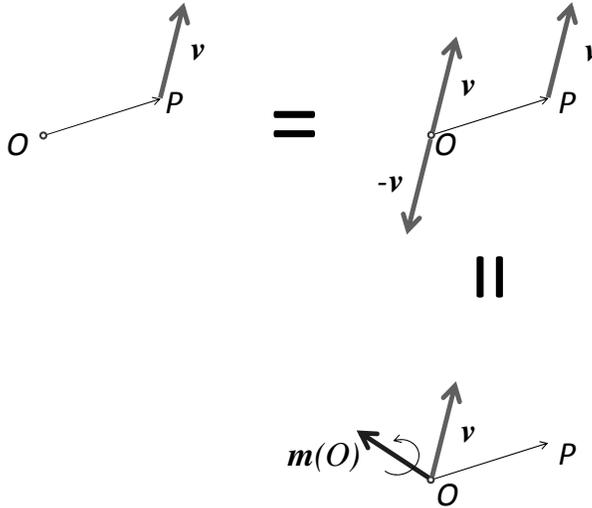


Figura 1.7: Un vettore v applicato in P è equivalente allo stesso vettore applicato in O e ad una coppia.

L'utilità del concetto di equivalenza di due sistemi di vettori è evidente: l'idea infatti è quella di sostituire un sistema complesso di vettori applicati con uno molto più semplice, costituito da un solo vettore ed una sola coppia. In sostanza, il risultato che ci accingiamo a dimostrare afferma che, fissato un arbitrario polo O e dato un sistema di vettori applicati (complesso quanto si vuole) questo è equivalente ad unico vettore (la risultante) applicato proprio in O ed ad una opportuna coppia, che indicheremo con $m(O)$. Se quindi si vuole calcolare il momento del sistema rispetto ad un altro centro di riduzione, per esempio rispetto al punto Q , si può sostituire il sistema con quello semplice: cioè con la risultante R applicata in O e la coppia $m(O)$, per cui

$$m(Q) = m(O) + (Q - O) \wedge R.$$

Teorema 1.5.1 *Un qualunque sistema di vettori applicati*

$$S = \{(P_i, v_i), i = 1, \dots, N\},$$

è equivalente ad un altro sistema S' costituito da una coppia e da un solo vettore applicato.

Dimostriamo il teorema considerando, per semplicità, un sistema di soli tre vettori applicati (la generalizzazione ad $N > 3$ vettori è banale)

$$\mathcal{S} = \{(P_1, \mathbf{v}_1), (P_2, \mathbf{v}_2), (P_3, \mathbf{v}_3)\}.$$

La dimostrazione è sostanzialmente grafica ed è riportata nella figura 1.8.

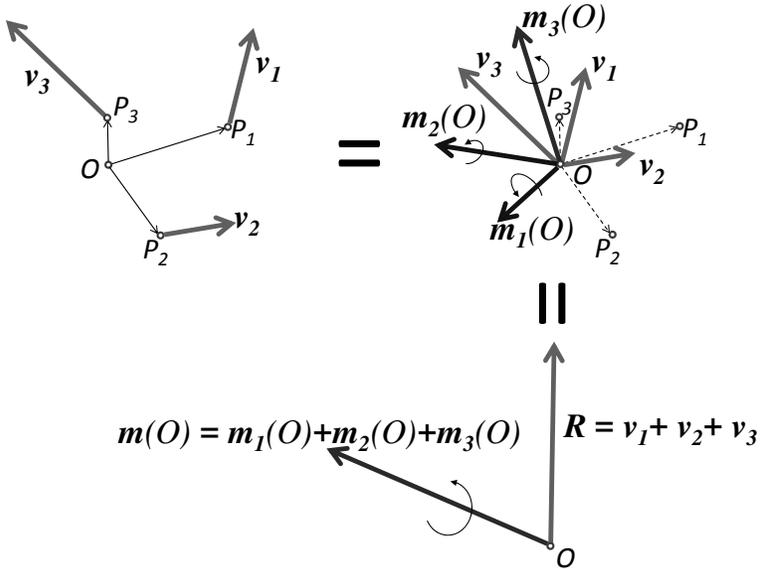


Figura 1.8: Ogni sistema di vettori applicati può essere ridotto ad una coppia e ad un vettore. Notare che, dapprima si riduce di ogni vettore applicato ad un vettore applicato in O più una coppia (vedi figura 1.7), dopodiché si considerano le risultanti di tutti i vettori applicati in O e di tutte le coppie.

In particolare, se O appartiene all'asse centrale, il sistema equivalente è costituito dalla risultante \mathbf{R} (applicata in O), mentre la coppia $\mathbf{m}(O)$ coincide con l'invariante vettoriale \mathbf{p} .

1.5.4 Sistemi di vettori paralleli

I vettori applicati del sistema $\mathcal{S} = \{(P_i, \mathbf{v}_i), i = 1, \dots, N\}$, si dicono **paralleli**, se esiste un vettore \mathbf{e} , per cui

$$\mathbf{v}_i = \beta_i \mathbf{e}, \text{ con } \beta_i \in \mathbb{R}, i = 1, \dots, N.$$

In questo caso, ponendo $\beta = \sum_{i=1}^N \beta_i$, la risultante è $\mathbf{R} = \beta \mathbf{e}$, ed il momento rispetto ad un qualunque polo O è

$$\mathbf{m}(O) = \left[\sum_{i=1}^N \beta_i (P_i - O) \right] \wedge \mathbf{e}.$$

E' banale provare che l'invariante scalare \mathfrak{S} si annulla. Di conseguenza, ricordando la scomposizione (1.21), $\mathbf{m}(O) = \mathbf{n}(O)$, e quindi $\mathbf{m}(O) = 0$, se O appartiene all'asse centrale. Quindi, applicando il teorema 1.5.1, abbiamo che il sistema è equivalente ad un unico vettore (la risultante) applicato in un qualunque punto dell'asse centrale. Se quindi Q è un polo che non giace sull'asse centrale

$$\mathbf{m}(Q) = (Q - O) \wedge \mathbf{R},$$

dove O è un qualunque punto dell'asse centrale. Quindi, se vogliamo semplificare un sistema di vettori applicati paralleli è necessario individuarne l'asse centrale. A tal proposito, vale il seguente

Teorema 1.5.2 *L'asse centrale di un sistema di vettori paralleli è la retta parallela a e , passante per il punto C definito da*

$$C - O = \frac{1}{\beta} \sum_{i=1}^N \beta_i (P_i - O), \quad (1.24)$$

detto *centro del sistema*.

Dim. Per provare che C appartiene all'asse centrale basta dimostrare che $\mathbf{m}(C) = 0$. Dunque

$$\mathbf{m}(C) = \sum_{i=1}^N (P_i - C) \wedge \beta_i \mathbf{e} = \left[\sum_{i=1}^N \beta_i (P_i - C) \right] \wedge \mathbf{e} = 0.$$

Infatti

$$\begin{aligned} \sum_{i=1}^N \beta_i (P_i - C) &= \sum_{i=1}^N \beta_i ((P_i - O) - (C - O)) \\ &= \beta (C - O) - \underbrace{\left(\sum_{i=1}^N \beta_i \right)}_{\beta} (C - O) = 0. \end{aligned}$$

Se poi prendiamo un altro punto Q tale che⁶ $(Q - C) = \lambda e$, con $\lambda \in \mathbb{R}$, applicando la (1.20) si ha

$$\mathbf{m}(Q) = \underbrace{\mathbf{m}(C)}_{=0} + \underbrace{(C - Q)}_{-\lambda e} \wedge \underbrace{\mathbf{R}}_{\beta e} = 0.$$

□

⁶Cioè Q giace sulla retta parallela ad e passante per C .

Capitolo 2

Equazioni differenziali

2.1 Introduzione

Il problema del moto, della sua descrizione e delle sue “cause”, è il punto chiave per la comprensione di gran parte delle applicazioni della matematica al mondo fisico.

La storia è lunga e va molto indietro nel tempo. In effetti questo problema aveva angustiato i pensatori greci classici (si pensi al famoso paradosso del Achille e la tartaruga, “risolto” da Zenone con la *negazione* della possibilità stessa del moto). Già allora era apparsa la stretta connessione del moto con la *continuità*, intesa come infinita (almeno potenzialmente) divisibilità dello spazio e del tempo.

La questione era destinata a trascinarsi per lungo tempo: per una chiara definizione di continuità (con la definizione dei numeri reali) si dovette attendere la metà del diciannovesimo secolo.

Ciò nonostante gli strumenti concettuali per la descrizione del moto si fecero strada molto prima, anche se in modo vago (come tutti i concetti degni di questo nome!). Nell’opera di Galileo la velocità appare come un concetto “primitivo”, una proprietà dei corpi in movimento, che non viene definita in funzione di altro (si indica però dei modi per calcolarla, per esempio facendo urtare un corpo contro un altro corpo fermo).

La chiave della “risoluzione” si trova nel lavoro di Newton che dà origine al calcolo differenziale. La velocità (istantanea) viene identificata con la *derivata* del moto, inteso come la funzione che associa la posizione nello spazio al tempo in cui il “mobile” la occupa¹.

¹Per quanto questo modo di pensare possa apparire molto plausibile, in effetti maschera un bel po’ di problemi concettuali. Per parlare di velocità si deve identificare ad ogni istante il corpo in moto con un punto geometrico; ma come facciamo a sapere veramente *quando* il corpo sta transitando in una data posizione nello spazio? Chiaramente dobbiamo “illuminare” il corpo e misurare l’intervallo di tempo che lo stesso impiega per transitare fra due punti dello spazio. La fisica moderna, attraverso la teoria della meccanica quantistica e la teoria della relatività, ci ha fatto capire che questa procedura concettuale non è corretta quando si ha a che fare con corpi “realmente” piccoli e veloci. Infatti se “illuminiamo” un corpo piccolo, quale per esempio un elettrone, perturbiamo significativamente, ed in modo casuale, il suo moto e quindi non siamo più in grado di misurarne la velocità.

Inoltre, anche restando nell’ambito puramente classico, la definizione di velocità come derivata della posizione assume che si abbia chiaro cosa si intende per “spazio” e soprattutto per “tempo”: nell’ambito della meccanica classica lo spazio è rappresentato matematicamente da uno spazio euclideo tridimensionale, il tempo da un continuo unidimensionale che, come afferma Newton, “scorre uniformemente”: “Tempus

L'accelerazione, che è la variazione della velocità rispetto al tempo, appare allora come la derivata della funzione che ad ogni istante associa la velocità istantanea.

Conoscendo ad ogni istante l'accelerazione è possibile ricostruire la velocità e quindi la posizione con l'operazione inversa della derivazione, l'integrazione. Ma questo sembra generare un regresso all'infinito: per conoscere l'accelerazione forse dovrei conoscere la sua variazione (la "derivata terza" del moto) e così via.

Il "miracolo" è che con l'accelerazione si può chiudere il regresso. Una lettura "purementemente matematica" della seconda legge² della dinamica newtoniana è che è *sempre possibile trovare una funzione della posizione del corpo, della sua velocità e del tempo, che determina la sua accelerazione.*

Questo, insieme con le leggi che regolano la struttura delle forze (il principio di azione-reazione, le specifiche leggi per le varie forze, come la legge del quadrato inverso per la gravitazione), trasforma il problema del moto in un problema matematico ben definito: la soluzione di un' **equazione differenziale**, o più in generale, di un **sistema di equazioni differenziali**.

2.2 Considerazioni generali

Cominciamo con un caso ben noto: il modello più comune per descrivere una molla è quello in cui il corpo subisce una forza di richiamo proporzionale a quanto si è allontanato da un punto fissato, il centro di attrazione. Questo si traduce nel modello matematico

$$m\ddot{x} = -Kx \quad (2.1)$$

dove m denota la massa del corpo, K è una costante positiva detta costante elastica, la funzione del tempo $x(t)$ indica la posizione del corpo e \ddot{x} la sua derivata seconda.

La (2.1) è un esempio di equazione differenziale del secondo ordine, cioè che contiene la funzione incognita e le sue derivate fino al secondo ordine.

Il problema fondamentale della Meccanica si può enunciare in questi termini: *date le forze, determinare il moto*, ovvero espressa la forza in funzione della posizione del corpo, della sua velocità e del tempo, determinare la posizione in funzione del tempo, ovvero la funzione $x(t)$, in modo che sia soddisfatta l'equazione

$$\ddot{x} = F(x, \dot{x}, t).$$

Diamo ora alcune definizioni fondamentali

absolutum, verum et mathematicum in se et natura sua sine relatione ad externum quodvis, aequabiliter fluit". Questo è il "modello" di spazio-tempo che fa da "sfondo" alla meccanica newtoniana.

Oggi le nostre idee sono forse un po' più confuse. Per esempio, secondo Roger Penrose: *The temporal ordering that we 'appear' to perceive is, I am claiming, something that we impose upon our perceptions in order to make sense of them in relation to the uniform forward time-progression of an external physical reality*. Si vedano gli articoli su tempo nel sito web di storia della matematica della St.Andrews University http://www-gap.dcs.st-and.ac.uk/history/HistTopics/Time_1.html

http://www-gap.dcs.st-and.ac.uk/history/HistTopics/Time_2.html

²*Mutationem motus proportionalem esse vi motrici impressae et fieri secundum lineam rectam qua vis illa imprimitur* ovvero "la variazione del moto (accelerazione) è proporzionale alla forza applicata ed è nella direzione di tale forza". Qui la forza è pensata come una quantità (vettoriale) nota in funzione dello stato cinematico del corpo (posizione e velocità).

Definizione 2.2.1 Sia data una funzione $f : \mathbb{R}^{n+2} \rightarrow \mathbb{R}$. Un'equazione differenziale (ordinaria) di ordine n è una relazione, che coinvolge una funzione incognita $x(t) : \mathbb{R} \rightarrow \mathbb{R}$, assieme con le sue derivate fino all'ordine n ,

$$f(x^{(n)}(t), x^{(n-1)}(t), \dots, x'(t), x(t), t) = 0$$

dove $x^{(k)}$ indica la derivata k -esima di x , che deve essere soddisfatta per ogni³ valore della variabile indipendente t . L'equazione è detta in **forma normale** se ha la forma

$$x^{(n)} = f(x^{(n-1)}, \dots, x', x, t)$$

L'equazione si dice **autonoma** se la funzione f non dipende da t .

Accanto alle equazioni scalari possiamo anche considerare le **equazioni differenziali vettoriali** (dette anche **sistemi di equazioni differenziali**),

$$\mathbf{F}(x^{(n)}, x^{(n-1)}, \dots, x', x, t) = 0, \quad (2.2)$$

dove l'incognita $x(t)$ è ora una funzione vettoriale a valori in \mathbb{R}^d e la funzione $\mathbf{F} : \mathbb{R}^{d \times (n+1)+1} \rightarrow \mathbb{R}^d$. L'equazione vettoriale si dirà in forma normale se (2.2) può essere scritta come

$$x^{(n)} = \mathbf{F}(x^{(n-1)}, \dots, x', x, t). \quad (2.3)$$

In effetti basta limitarsi a equazioni differenziali vettoriali del primo ordine. Infatti qualsiasi equazione di ordine superiore può essere ridotta a un sistema del primo ordine introducendo un numero opportuno di variabili ausiliarie (cioè nuove funzioni incognite). Per fissare le idee vediamo come si passa da un'equazione scalare del secondo ordine a un sistema di due equazioni del primo ordine.

Sia quindi data l'equazione

$$\ddot{x} = f(\dot{x}, x, t). \quad (2.4)$$

Poniamo

$$y = \dot{x} \quad (2.5)$$

avremo di conseguenza che $\ddot{x} = \dot{y}$ e possiamo scrivere la (2.4) e la (2.5) come

$$\begin{cases} \dot{x} = y, \\ \dot{y} = f(y, x, t), \end{cases}$$

ovvero come

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \quad \text{dove} \quad \mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad \mathbf{F}(\mathbf{x}(t), t) = \begin{pmatrix} y \\ f(x, y, t) \end{pmatrix}.$$

La generalizzazione di questa riduzione al primo ordine per sistemi di equazioni differenziali di ordine superiore è ovvia.

³A voler essere precisi, la definizione che abbiamo dato dovrebbe essere "localizzata": la funzione f può essere definita solo in un sottoinsieme aperto di \mathbb{R}^{n+2} , così come la eventuale soluzione $x(t)$ può essere definita solo in un certo intervallo (t_1, t_2) . Quest'ultima limitazione è molto importante in quanto, anche per funzioni f "semplici", p.e. $f(x', x) = x' - x^2$, le soluzioni, che pure esistono, *non* sono definite per tutti i valori della variabile indipendente t ma solo in intervalli limitati.

2.3 Il problema di Cauchy

È quasi immediato rendersi conto che un'equazione differenziale non ha soluzione unica. Basta prendere la più semplice possibile

$$\dot{x} = 0 \quad (2.6)$$

e verificare che tutte le funzioni *costanti* sono soluzioni di (2.6).

L'equazione (2.6) ha però una sola soluzione che soddisfa *anche* la condizione

$$x(t_0) = x_0 \quad (2.7)$$

dove t_0 è un tempo fissato e x_0 un valore fissato (la soluzione è, ovviamente, $x(t) = x_0$ per ogni t).

La condizione (2.7) si dice **condizione iniziale**, o di Cauchy, e determina, sotto opportune condizioni, in modo univoco la soluzione di un'equazione differenziale (del primo ordine). Il problema di determinare la soluzione di un'equazione differenziale che soddisfi a una data condizione iniziale è detto **problema ai dati iniziali** o **problema di Cauchy**.

Enunciamo il teorema fondamentale sull'esistenza e unicità delle soluzioni del problema di Cauchy. Esso richiede che l'equazione sia in forma normale (2.3). Enunciamo il teorema per il semplice caso di una sola equazione differenziale, ovvero

$$\begin{cases} \dot{x} = f(x, t), \\ x(t_0) = x_0. \end{cases} \quad (2.8)$$

Teorema 2.3.1 *Sia data una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, continua rispetto alle variabili x e t e lipschitziana⁴ rispetto alla variabile x in un intorno di $(x_0, t_0) \in \mathbb{R}^2$.*

Esistenza: *Esiste un $\delta > 0$ e una funzione $x : (t_0 - \delta, t_0 + \delta) \rightarrow \mathbb{R}$, di classe C^1 in $(t_0 - \delta, t_0 + \delta)$ che soddisfa il problema di Cauchy (2.8).*

Unicità: *Se $\bar{\delta} > 0$ e $y : (t_0 - \bar{\delta}, t_0 + \bar{\delta}) \rightarrow \mathbb{R}$, un'altra funzione che soddisfa anch'essa (2.8), allora $x(t) \equiv y(t)$ in $(t_0 - \delta_m, t_0 + \delta_m)$, con $\delta_m = \min\{\delta, \bar{\delta}\}$.*

Nota 2.3.1 *Una prima osservazione è che il teorema ha carattere **locale** ovvero non garantisce l'esistenza della soluzione per tutti i tempi $t \in \mathbb{R}$. Questo perché per una funzione f generica, anche molto regolare, la soluzione può “esplodere” in tempo finito. Come esempio si prenda l'equazione*

$$\dot{x} = x^2$$

che, se imponiamo la condizione iniziale $x(0) = 1$, ha per soluzione $x(t) = 1/(1-t)$. Questa soluzione è definita solo nell'intervallo $t < 1$.

Concludiamo osservando che la “lipschitzianità” richiesta alla f non è un “requisito tecnico”, ma è legata alla struttura del problema di Cauchy. Se infatti consideriamo il seguente problema

$$\begin{cases} \dot{x} = \frac{2\sqrt[3]{x}}{3}, \\ x(t_0) = 0, \end{cases}$$

⁴Una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$, si dice *lipschitziana* (o Lipschitz continua) se esiste una costante L tale che, per ogni coppia di punti x e $y \in \mathbb{R}$, f soddisfa la disuguaglianza $|f(x) - f(y)| \leq L|x - y|$.

è banale rendersi conto che esso ha due soluzioni: $x(t) \equiv 0$, e $x(t) = t^{3/2}$. E infatti la funzione $\sqrt[3]{x}$ non è lipschitziana in $x = 0$.

Nota 2.3.2 Visto che la soluzione che ci si aspetta è comunque locale, è facile capire perché le ipotesi sulla regolarità della f sono “localizzate”. Si richiede infatti che la funzione f sia continua e lipschitziana **non per tutti gli** $(x, t) \in \mathbb{R}^2$, **ma solo in un intorno del punto** (x_0, t_0) .

2.3.1 Equazioni autonome

Come abbiamo detto un’equazione differenziale, o un sistema di equazioni differenziali, in cui il tempo non compare esplicitamente si dice **autonoma**. Queste equazioni hanno una fondamentale proprietà detta **invarianza temporale**: data una soluzione $x(t)$ e un valore del tempo T qualsiasi, allora $x(t - T)$ è ancora soluzione dell’equazione. Supponiamo infatti di avere il seguente problema di Cauchy

$$\begin{cases} \frac{dx}{dt} = f(x), \\ x(0) = a, \end{cases} \quad (2.9)$$

la cui soluzione è $\hat{x}(t)$. Trasliamo temporalmente di T il problema (2.9), considerando

$$\begin{cases} \frac{dy}{dt} = f(y), \\ y(T) = a. \end{cases} \quad (2.10)$$

E’ facile provare che $y(t) = \hat{x}(t - T)$ è soluzione di (2.10). Infatti

$$\frac{dy(t)}{dt} = \frac{d\hat{x}(t - T)}{dt} = \frac{d\hat{x}(t - T)}{d(t - T)} \underbrace{\frac{d(t - T)}{dt}}_{=1} \stackrel{(2.9)_1}{=} f(\hat{x}(t - T)) = f(y(t)),$$

e

$$y(T) = \hat{x}(t - T)|_{t=T} = \hat{x}(0) = a.$$

Ancora più importante è capire come si interpreta questa proprietà. Supponiamo che, a partire da un certo istante, venga effettuato un certo esperimento su una quantità la cui evoluzione nel tempo è descritta da un’equazione differenziale autonoma. Sia $x(t)$ la soluzione dell’equazione. Supponiamo poi di effettuare lo stesso esperimento (con le medesime condizioni iniziali) T giorni dopo (si ipotizza che il tempo sia misurato in giorni). Se nulla è cambiato nelle condizioni in cui si svolge l’esperimento⁵, fisicamente ci aspettiamo che le due evoluzioni siano uguali, ovvero che coincidano se traslate temporalmente. E infatti, posto $y(t)$ la soluzione ottenuta a partire da T , abbiamo

$$y(t) = x(t - T).$$

In pratica, nelle equazioni autonome, il legame tra la funzione x e le sue derivate **non dipende** dalla variabile indipendente (il tempo) e di conseguenza l’equazione differenziale non cambia in forma se *traslata temporalmente*.

⁵Qui supponiamo di poter trascurare gli “errori sperimentali”!

Esempio 2.3.1 Consideriamo i seguenti problemi di Cauchy

$$P_A : \begin{cases} \dot{x} = x, \\ x(0) = 1, \end{cases} \quad P_B : \begin{cases} \dot{x} = tx, \\ x(0) = 1. \end{cases}$$

Evidentemente P_A coinvolge un'equazione autonoma mentre l'equazione di P_B non è autonoma.

La soluzione di P_A è la seguente

$$\frac{\dot{x}}{x} = 1, \implies \underbrace{\int_0^t \frac{\dot{x}(t')}{x(t')} dt'}_{\ln \frac{x(t)}{x(0)} = \ln x(t)} = \underbrace{\int_0^t dt'}_t, \implies x(t) = e^t.$$

Se adesso consideriamo

$$\begin{cases} \dot{y} = y, \\ y(T) = 1, \end{cases} \quad (2.11)$$

è banale verificare che la soluzione è $y(t) = e^{(t-T)}$. In altri termini, la soluzione di (2.11) è la soluzione di P_A traslata temporalmente di T : cioè $y(t) = x(t-T)$. Dunque il problema di Cauchy P_A non cambia per traslazioni temporali.

Consideriamo adesso il problema P_B la cui soluzione è

$$\frac{\dot{x}}{x} = t, \implies \underbrace{\int_0^t \frac{\dot{x}(t')}{x(t')} dt'}_{\ln \frac{x(t)}{x(0)} = \ln x(t)} = \underbrace{\int_0^t t' dt'}_{\frac{t^2}{2}}, \implies x(t) = \exp \left\{ \frac{t^2}{2} \right\},$$

da cui $x(t-T) = \exp \left\{ \frac{(t-T)^2}{2} \right\}$. Risolviamo il problema P_B a partire da T

$$\begin{cases} \dot{y} = ty, \\ y(T) = 1. \end{cases} \implies y(t) = \exp \left\{ \frac{1}{2} (t^2 - T^2) \right\}.$$

In questo caso $x(t-T)$ non è più soluzione: traslando di T la variabile temporale, otteniamo due diversi problemi di Cauchy.

2.3.2 Equazioni reversibili

Un'altra classe di equazioni differenziali rilevanti nella Meccanica è quella delle equazioni dette **reversibili** ovvero quelle per cui se $x(t)$ è una soluzione allora lo è anche $x(-t)$. Questo accade in Meccanica tutte le volte che abbiamo sistemi isolati o anche in sistemi non isolati ma dove le forze esterne non dipendano dal tempo e gli effetti delle forze *dissipative* (dovute, per esempio, agli attriti) siano trascurabili. In questi casi le equazioni avranno la forma $\ddot{x} = f(x)$ di cui si verifica immediatamente la reversibilità (in effetti: cambiando il segno del tempo la derivata prima cambia di segno, ma la derivata seconda ritorna del segno di partenza).

2.3.3 Equazioni integrabili

Nonostante il teorema di esistenza e unicità ci garantisca che il problema di Cauchy è univocamente risolvibile (almeno localmente), tuttavia non è possibile, in generale, dare un'espressione analitica della soluzione tramite una combinazione finita di "funzioni elementari"⁶ e neppure trovare un algoritmo che permetta di esprimere la soluzione tramite un numero finito di operazioni di integrazione e di inversione di funzioni. Qualora questo sia possibile diremo che l'equazione differenziale è "integrabile". Lo studio di una classe di equazioni del *primo ordine* (cioè che coinvolgono le sole derivate prime) integrabile sarà sviluppato nella sezione 2.4. Nei paragrafi 2.5, 2.6 analizzeremo due classi di equazioni del *secondo ordine* (che coinvolgono cioè soltanto derivate prime e seconde) integrabili.

2.4 Equazioni del primo ordine in forma normale

In questo paragrafo analizzeremo la classe più importante delle equazioni del primo ordine in forma normale: le cosiddette **equazioni a variabili separabili**. Per capire cosa intendiamo per equazione a variabili separabili consideriamo la seguente equazione autonoma del primo ordine (in forma normale)

$$\dot{y} = g(y) h(t). \quad (2.12)$$

In questo caso l'equazione è integrabile per con il *metodo di separazione delle variabili*. Riscriviamo infatti (2.12) come

$$\underbrace{\frac{\dot{y}}{g(y)}}_{\substack{\text{espressione} \\ \text{che dipende} \\ \text{da } y \text{ e } \dot{y}}} = \underbrace{h(t)}_{\substack{\text{espressione} \\ \text{che dipende} \\ \text{da } t}}. \quad (2.13)$$

Introduciamo la funzione

$$G(y) = \int \frac{dy}{g(y)}, \quad (2.14)$$

cioè una primitiva della funzione $\frac{1}{g(y)}$, e

$$H(t) = \int h(t) dt, \quad (2.15)$$

che è una qualsiasi primitiva della funzione $h(t)$. Basta osservare ora che il primo membro della (2.13) è la derivata rispetto al tempo della funzione $G(y(t))$

$$\frac{dG(y(t))}{dt} = \frac{dG}{dy} \frac{dy}{dt} = \frac{1}{g(y)} \frac{dy}{dt},$$

⁶Per "funzione elementare" si intende generalmente una funzione che appartenga al seguente "catalogo": polinomi, funzioni razionali, funzioni trigonometriche ed esponenziali, loro inverse e tutte funzioni ottenute combinando un numero finito di queste funzioni.

mentre il secondo membro della (2.13) è la derivata della funzione $H(t)$ rispetto al tempo. La (2.13) è dunque un'uguaglianza tra le derivate di due funzioni e di conseguenza le due funzioni possono al più differire per una costante

$$G(y(t)) = H(t) + \text{costante}, \quad (2.16)$$

da cui, indicando con G^{-1} l'inversa della G , otteniamo

$$y(t) = G^{-1}(H(t) + \text{costante}). \quad (2.17)$$

Vediamo adesso il caso “più semplice” delle equazioni a variabili separabili:

$$\dot{y} = g(y), \quad \Leftrightarrow \quad \frac{\dot{y}}{g(y)} = 1, \quad (2.18)$$

che corrisponde ad $h(t) = 1$, $\implies H(t) = t$. La (2.16) dà quindi luogo a

$$G(y(t)) = t + \text{costante}, \quad \implies \quad y(t) = G^{-1}(t + \text{costante}). \quad (2.19)$$

La comparsa di una costante di integrazione non deve meravigliare: sappiamo che la (2.12) ammette un'unica soluzione soltanto qualora si specifichi il dato iniziale $y(t_0) = y_0$, cosa che non abbiamo fatto quando abbiamo ricavato la (2.17). Vediamo quindi come si procede quando è noto il dato iniziale. Per semplificare il più possibile consideriamo l'equazione (2.18) con il dato iniziale $y(t_0) = y_0$. Abbiamo quindi il seguente problema di Cauchy

$$\begin{cases} \dot{y} = g(y), \\ y(t_0) = y_0. \end{cases} \quad (2.20)$$

Adesso integriamo fra t_0 e t la (2.18), tenendo conto dal dato $y(t_0) = y_0$,

$$\begin{aligned} t - t_0 &= \int_{t_0}^t ds = \int_{t_0}^t \frac{\dot{y}(s)}{g(y(s))} ds \stackrel{\eta=y(s)}{=} \int_{y(t_0)}^{y(t)} \frac{d\eta}{g(\eta)} \\ &\stackrel{(2.14)}{=} G(y(t)) - G(y(t_0)), \end{aligned} \quad (2.21)$$

ottenendo

$$G(y(t)) = G(y(t_0)) + t - t_0.$$

Possiamo dunque scrivere la soluzione come⁷

$$y(t) = G^{-1}(t - t_0 + G(y_0)). \quad (2.22)$$

Abbiamo quindi espresso la soluzione dell'equazione differenziale tramite il calcolo di un integrale e l'inversione di una funzione: questo è ciò che si intende per “integrazione” di un'equazione differenziale.

Il calcolo che ha portato alla risoluzione di (2.20), si generalizza al caso in cui l'equazione sia la (2.12), cioè

$$\begin{cases} \dot{y} = g(y)h(t), \\ y(t_0) = y_0. \end{cases} \quad (2.23)$$

In questo caso invece della (2.21) abbiamo $G(y(t)) - G(y(t_0)) = H(t) - H(t_0)$, cioè

$$y(t) = G^{-1}(H(t) - H(t_0) + G(y_0)).$$

⁷Notiamo dalla (2.22) è banale verificare $y(t)|_{t=t_0} = y_0$.

Esempio 2.4.1 Vediamo la soluzione generale della seguente equazione a variabili separabili

$$\dot{y} = \lambda y.$$

In questo caso, ricordando la (2.14) e la (2.15)

$$G(y) = \ln y,$$

$$H(t) = \lambda t.$$

Di conseguenza, dalla (2.16) e (2.17)

$$\ln y = \lambda t + \text{cost.}, \implies y = e^{\lambda t + \text{cost.}} = \underbrace{e^{\text{cost.}}}_C e^{\lambda t},$$

per cui la soluzione generale è $y = C e^{\lambda t}$.

Esempio 2.4.2 Consideriamo l'equazione del tipo

$$\dot{y} = \lambda y + f(t). \quad (2.24)$$

Generalizzando quanto visto nell'esempio 2.4.1 si cerca la soluzione del tipo $y = C(t) e^{\lambda t}$. Sostituendo nella (2.24)

$$\dot{C} e^{\lambda t} + C \lambda e^{\lambda t} = C \lambda e^{\lambda t} + f(t), \implies \dot{C} = e^{-\lambda t} f(t).$$

Si ha quindi

$$C(t) = \int_0^t e^{-\lambda \tau} f(\tau) d\tau + K,$$

dove K è una generica costante di integrazione. Abbiamo quindi

$$y(t) = \int_0^t e^{\lambda(t-\tau)} f(\tau) d\tau + K e^{\lambda t}.$$

Analizziamo anche il caso particolare $f(t) = A e^{\lambda t}$, che troveremo nel seguito. La soluzione è

$$y(t) = A \int_0^t e^{\lambda(t-\tau)} e^{\lambda \tau} d\tau + K e^{\lambda t} = A e^{\lambda t} \int_0^t d\tau + K e^{\lambda t} = e^{\lambda t} (At + K).$$

2.5 Equazioni del secondo ordine in forma normale del tipo $\ddot{q} = f(\dot{q})$

Una generica equazione del second'ordine, anche autonoma, non può in generale essere "integrata". Tuttavia ci sono importanti eccezioni. La prima è quella delle equazioni della forma

$$\ddot{q} = f(\dot{q}), \quad (2.25)$$

che chiaramente sono equazioni del primo ordine a variabili separabili nell'incognita $y(t)$, una volta posto $y = \dot{q}$. Possiamo quindi esprimere $y(t)$ come abbiamo fatto nella (2.18), e quindi integrare ancora rispetto al tempo, ottenendo così la $q(t)$.

Esempio 2.5.1 Un esempio significativo di equazione tipo (2.25) è l'equazione che descrive il moto di un grave soggetto alla forza peso e all'attrito viscoso dell'aria (la resistenza viscosa è, in prima approssimazione, proporzionale alla velocità stessa del grave)

$$\ddot{h} = -g - \frac{\eta}{m}\dot{h}, \quad (2.26)$$

dove h rappresenta la quota del grave, g è l'accelerazione di gravità, m la massa del grave e η , costante, è il coefficiente di attrito viscoso. Ponendo $y = \dot{h}$, e $\mu = \eta/m$, la (2.26) si trasforma nella seguente equazione del primo ordine a variabili separabili

$$\frac{\dot{y}}{g + \mu y} = -1, \implies \frac{1}{\mu} \frac{d}{dt} [\ln(g + \mu y)] = -\frac{dt}{dt},$$

da cui otteniamo

$$\ln(g + \mu y) = -\mu t + \tilde{C}, \implies y = \frac{1}{\mu} \left(-g + e^{-\mu t + \tilde{C}} \right),$$

che scriveremo come

$$y(t) = -\frac{g}{\mu}t + \frac{C_1}{\mu}e^{-\mu t},$$

dove $C_1 = e^{\tilde{C}}$ è la prima costante d'integrazione. Integrando ancora

$$h(t) = -\frac{g}{\mu}t - \frac{C_1}{\mu^2}e^{-\mu t} + C_2,$$

dove C_2 è la seconda costante d'integrazione. In particolare, se consideriamo il problema di Cauchy

$$\begin{cases} \ddot{h} = -g - \mu\dot{h}, \\ \dot{h}(0) = 0, \\ h(0) = h_0, \end{cases}$$

che corrisponde al caso del grave lasciato cadere da fermo da un'altezza h_0 , troviamo la seguente soluzione

$$h(t) = -\frac{gt}{\mu} + \frac{g}{\mu^2}(1 - e^{-\mu t}) + h_0.$$

Notiamo che, se μt è "piccolo", ovvero stiamo considerando i primi momenti della caduta, considerando l'approssimazione

$$e^{-\mu t} \approx 1 - \mu t + \frac{1}{2}\mu^2 t^2,$$

otteniamo $h(t) \approx h_0 - \frac{1}{2}gt^2$, che è la formula del moto uniformemente accelerato.

Al contrario, se μt è "grande", si ottiene

$$h(t) \approx h_0 + \frac{g}{\mu^2} - \frac{gt}{\mu},$$

corrispondente ad un moto di caduta con velocità costante proporzionale alla massa ed inversamente proporzionale al coefficiente di attrito.

2.6 Equazioni del secondo ordine del tipo $\ddot{q} = f(q)$: caso conservativo

Un'altra classe di equazioni del secondo ordine integrabili è quella costituita dalle equazioni della forma

$$\ddot{q} = f(q). \quad (2.27)$$

Tale classe di equazioni è fondamentale per la Meccanica poiché di questa forma sono tutte le equazioni di moto di un sistema conservativo con un solo grado di libertà. La (2.27) è riconducibile a un'equazione del primo ordine tramite la seguente osservazione: moltiplichiamo ambo i membri della (2.27) per la funzione \dot{q}

$$\ddot{q}\dot{q} = f(q)\dot{q}, \quad (2.28)$$

e indichiamo con

$$V(q) = - \int f(q) dq, \quad (2.29)$$

una qualsiasi primitiva, cambiata di segno, della funzione $f(q)$. In particolare, la (2.27) potrà essere scritta come

$$\ddot{q} = - \frac{dV(q)}{dq} = f(q). \quad (2.30)$$

La funzione $V(q)$ è usualmente detta **energia potenziale**, anche se, a questo livello, non c'è corrispondenza fra l'equazione (2.30) ed un reale sistema fisico.

Nel primo membro della (2.28) riconosciamo la derivata rispetto al tempo della funzione $\dot{q}^2/2$. Il secondo membro è anch'esso una derivata rispetto al tempo, infatti

$$- \frac{dV(q(t))}{dt} = - \frac{dV(q)}{dq} \frac{dq}{dt} = - \frac{d}{dq} \left(- \int f(q) dq \right) \frac{dq}{dt} = f(q)\dot{q}.$$

Possiamo quindi riscrivere la (2.28) come

$$\frac{d}{dt} \left(\frac{\dot{q}^2}{2} + V(q) \right) = 0,$$

e dunque ottenere

$$\frac{\dot{q}^2}{2} = E - V(q), \quad (2.31)$$

dove abbiamo indicato con E il *valore, arbitrario, della costante di integrazione*⁸, che viene determinato in base alle condizioni iniziali. In genere la costante E viene detta **energia meccanica**.

L'equazione (2.31) può essere risolta rispetto a \dot{q} (purché $V(q) \leq E$). Supponiamo quindi di avere assegnato, per l'equazione differenziale originaria (2.27), condizioni iniziali con $q(t_0) = q_0$ e $\dot{q}(t_0) = v_0 > 0$ (aver fissato positivo il segno di v_0 non

⁸Il fatto che si siano usate le lettere V ed E dovrebbe richiamare alla mente la conservazione dell'energia totale del sistema $E = \frac{\dot{q}^2}{2} + V(q)$. Infatti, talvolta il termine $\frac{\dot{q}^2}{2}$ è detto energia cinetica.

pregiudica la generalità di quanto segue). Possiamo quindi calcolare esplicitamente la costante di integrazione E

$$E = \frac{\dot{q}^2(t_0)}{2} + V(q_0) = \frac{v_0^2}{2} + V(q_0),$$

e quindi risolvere algebricamente la (2.31) ottenendo⁹

$$\dot{q} = \sqrt{2(E - V(q))}, \implies \frac{\dot{q}}{\sqrt{2(E - V(q))}} = 1, \quad (2.32)$$

che è un'equazione del tipo (2.12) e quindi può essere risolta per separazione delle variabili

$$t - t_0 = \int_{q_0}^{q(t)} \frac{d\eta}{\sqrt{2(E - V(\eta))}}. \quad (2.33)$$

Per determinare esplicitamente la soluzione $q(t)$, si deve risolvere l'integrale in (2.33) ed esplicitare la funzione di $q(t)$ ottenuta dall'integrazione. Questo procedimento può scontrarsi (molto spesso) con l'impossibilità di esprimere l'integrale a secondo membro della (2.33) in termini di funzioni elementari. Questo accade, per esempio, per il moto del pendolo semplice¹⁰. Infatti in questo caso abbiamo l'equazione di moto è

$$\ddot{\theta} + \frac{g}{l} \sin \theta = 0, \quad (2.34)$$

la cui soluzione, espressa nella forma (2.33), è

$$t - t_0 = \int_{\theta_0}^{\theta(t)} \frac{d\eta}{\sqrt{2 \left(E - 2 \frac{g}{l} \sin^2 \left(\frac{\eta}{2} \right) \right)}}. \quad (2.35)$$

Questo integrale però non può essere esplicitato tramite una combinazione finita di funzioni elementari¹¹. Torneremo sul moto del pendolo nell'esempio 2.6.3.

Anche nel caso che il processo di integrazione e inversione possa essere portato a termine, resta il fatto che non sempre un'espressione esplicita della soluzione è particolarmente facile da “decifrare”, ovvero resta il problema di capire come effettivamente si comporti la soluzione descritta da tale espressione analitica.

Esempio 2.6.1 Consideriamo un punto materiale P di massa m vincolato a scorrere senza attrito lungo una retta. P è collegato tramite una molla di costante elastica k (avete massa e lunghezza a riposo trascurabili) ad un punto fisso O della retta. Se $q(t)$ denota l'ascissa di P lungo la retta (O è l'origine), l'equazione di moto è

$$m\ddot{q} = -kq, \implies \ddot{q} = -\omega^2 q, \text{ con } \omega^2 = \frac{k}{m}, \text{ pulsazione.} \quad (2.36)$$

⁹Avendo assunto $v_0 > 0$, cioè velocità iniziale positiva, si deve considerare la radice positiva.

¹⁰La posizione del pendolo è univocamente determinata dall'angolo (orientato) θ che il pendolo forma con la verticale.

¹¹Le funzioni che provengono da integrali del tipo di quello che compare nella (2.35) sono dette funzioni ellittiche.

Tale sistema è detto **oscillatore armonico**, e la (2.36) è l'equazione caratteristica. Ricordando la (2.29) introduciamo la funzione $V(q)$

$$V(q) = - \int (-\omega^2 q) dq = \frac{\omega^2}{2} q^2,$$

da cui

$$\frac{\dot{q}^2}{2} + \frac{\omega^2}{2} q^2 = E, \implies \dot{q}^2 = 2 \left(E - \frac{\omega^2}{2} q^2 \right). \quad (2.37)$$

Si osserva dunque che, a meno di considerare la soluzione banale $q(t) \equiv 0$, l'energia E è sicuramente positiva. Quest'ultima si determina sulla base delle condizioni iniziali: $E = 1/2 [\dot{q}^2(t_0) + \omega^2 q^2(t_0)]$, se t_0 è l'istante iniziale. Procedendo come nella (2.33) si ottiene

$$t - t_0 = \int_{q_0}^{q(t)} \frac{d\eta}{\sqrt{2E - \omega^2 \eta^2}} = \frac{1}{\omega} \int_{q_0}^q \frac{d\eta}{\sqrt{\left(\frac{\sqrt{2E}}{\omega}\right)^2 - \eta^2}}.$$

Siccome $\int \frac{d\eta}{\sqrt{a^2 - \eta^2}} = \arcsin\left(\frac{\eta}{a}\right)$, otteniamo

$$\arcsin\left(\frac{q(t)}{\sqrt{2E}/\omega}\right) - \arcsin\left(\frac{q_0}{\sqrt{2E}/\omega}\right) = \omega(t - t_0).$$

Possiamo quindi concludere che la forma generale della soluzione dell'equazione dell'oscillatore armonico è

$$q(t) = A \sin(\omega t + \phi), \quad (2.38)$$

dove A e ϕ dipendono dalle condizioni iniziali mentre ω , definita nella (2.36), dipende dalle caratteristiche fisiche del sistema. In particolare, la soluzione è periodica di periodo $T = 2\pi/\omega$.

Osserviamo infine che la soluzione (2.38) si può anche ottenere con altri metodi (si rimanda all'esempio 2.10.4).

2.6.1 Analisi qualitativa nel caso conservativo

Molte informazioni sulla soluzione possono essere ottenute tramite una "analisi qualitativa", che prescinde dal calcolo esplicito della soluzione stessa.

Cominciamo dal **caso conservativo**, ovvero da un'equazione nella forma (2.27). In questo caso abbiamo detto che le soluzioni dell'equazione sono tali che, al variare di t la quantità

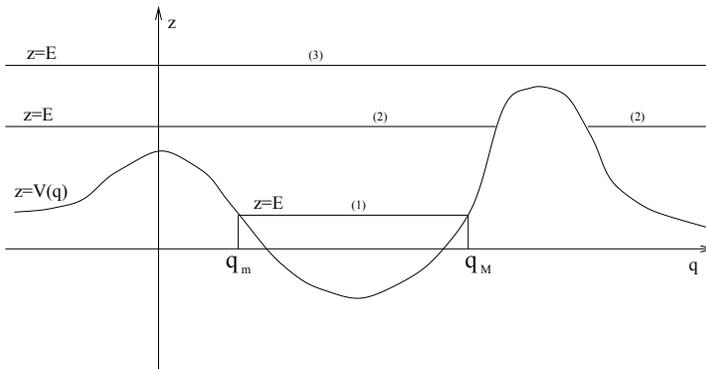
$$\frac{\dot{q}^2}{2} + V(q), \quad (2.39)$$

resta costante, o, come si dice, è un **integrale primo** dell'equazione differenziale¹². Abbiamo quindi

$$\dot{q}^2 = 2[E - V(q)], \quad (2.40)$$

¹²Data un'equazione differenziale $\dot{q} = f(\dot{q}, q, t)$, una funzione $g(\dot{q}, q, t)$ delle variabili \dot{q} , q e t tale che $\frac{dg(\dot{q}(t), q(t), t)}{dt} = 0$, per ogni funzione $q(t)$ soluzione dell'equazione differenziale, si dice un **integrale primo**.

Figura 2.1: Grafico della funzione energia potenziale e livelli dell'energia.



dove E è **un numero che viene fissato dalle condizioni iniziali**. Poiché il primo membro della (2.40) è non negativo, il moto si dovrà svolgere in un intervallo dell'asse delle q dove è soddisfatta la disuguaglianza

$$V(q) \leq E. \quad (2.41)$$

In effetti l'insieme delle soluzioni di (2.41) può essere formato da più intervalli disgiunti¹³: per continuità il moto potrà svolgersi solo su una componente connessa di questo insieme. Quindi, dati $V(q)$ ed E , possiamo avere questi casi:

1. il moto si svolge in un intervallo limitato $[q_m, q_M]$, e può essere:
 - a. periodico;
 - b. aperiodico, ovvero di "periodo" infinito;
2. punti di equilibrio, ovvero $q(t) = \text{costante}$ è soluzione;
3. il moto si svolge in un intervallo illimitato della forma $[q_M, +\infty)$ (oppure del tipo $(-\infty, q_m]$) o su tutta la retta reale;
4. il moto corrisponde ad una *separatrice*.

Caso 1.a. Moto limitato e periodico

Nel primo caso abbiamo un moto limitato tra due valori, uno minimo q_m e uno massimo q_M , dove q_m e q_M sono le intersezioni del grafico $z = V(q)$ con la retta $z = E$, come mostrato in figura 2.1, ovvero le soluzioni di $V(q) = E$. Vogliamo mostrare che in questo caso il moto è periodico e il periodo è determinato da

$$T = 2 \int_{q_m}^{q_M} \frac{d\eta}{\sqrt{2(E - V(\eta))}}, \quad (2.42)$$

¹³Per essere precisi gli intervalli disgiunti sono gli intervalli aperti delle soluzioni della disuguaglianza stretta.

se le **due radici** q_m e q_M **dell'equazione** $V(q) = E$ **sono semplici**. Con il termine radice semplice intendiamo che la funzione

$$F(q) = 2(E - V(q)) \geq 0, \quad (2.43)$$

può essere scritta come:

- $F(q) = (q_M - q) \Phi(q)$, con $\Phi(q) \geq \alpha > 0$, per q in un intorno di q_M ;
- $F(q) = (q - q_m) \Psi(q)$, con $\Psi(q) \geq \beta > 0$, per q in un intorno di q_m .

Supponiamo che a un certo istante (che possiamo prendere, senza perdere in generalità, come $t = 0$) il sistema si trovi nel punto $q_0 \in (q_m, q_M)$. Essendo fissato il valore dell'energia E , il modulo della velocità $|v_0| = |\dot{q}(0)|$ è determinato dalla (2.32). Ancora senza perdere di generalità possiamo assumere che $v_0 > 0$. Il sistema si muoverà quindi verso il punto q_M ; quando sarà sufficientemente vicino a q_M la sua velocità comincerà a decrescere verso zero. Infatti, quando $q \rightarrow q_M$, si ha che $(E - V(q)) \rightarrow 0$ e quindi, dalla (2.40) si deduce che $\dot{q} \rightarrow 0$. Mostriamo che il sistema raggiunge il punto q_M in un tempo finito. Il tempo necessario per “raggiungere” q_M si ricava dalla¹⁴ (2.33)

$$t_M = \int_{q_0}^{q_M} \frac{d\eta}{\sqrt{2(E - V(\eta))}} \stackrel{(2.43)}{=} \int_{q_0}^{q_M} \frac{d\eta}{\sqrt{F(\eta)}}. \quad (2.44)$$

Ora la funzione integranda diverge in q_M , e quindi l'integrale potrebbe divergere, cosa che implicherebbe $t_M \rightarrow \infty$. Ma la convergenza dell'integrale è garantita dal fatto che q_M è una radice semplice. Possiamo quindi scrivere¹⁵

$$t_M = \int_{q_0}^{q_M} \frac{d\eta}{\sqrt{F(\eta)}} = \int_{q_0}^{q_M} \frac{d\eta}{\sqrt{(q_M - \eta) \Phi(\eta)}} \leq \frac{1}{\sqrt{\alpha}} \underbrace{\int_{q_0}^{q_M} \frac{d\eta}{\sqrt{(q_M - \eta)}}}_{\text{integrale convergente}} < \infty.$$

Il sistema quindi impiega un tempo finito per raggiungere q_M . Per $t = t_M$, abbiamo $q(t_M) = q_M$, e $\dot{q}(t_M) = 0$. **Il sistema giunge con velocità nulla in q_M , dove il moto si inverte.** Infatti, ricordando la (2.30) e la (2.43) abbiamo

$$\begin{aligned} \ddot{q}(t_M) &= - \left. \frac{dV(q)}{dq} \right|_{q_M} = \frac{1}{2} \left. \frac{dF(q)}{dq} \right|_{q_M} = \frac{1}{2} \left. \frac{d}{dq} ((q_M - q) \Phi(q)) \right|_{q_M} \\ &= \frac{1}{2} [(q_M - q) \Phi'(q) - \Phi(q)]_{q_M} = -\frac{1}{2} \Phi(q_M) \leq -\frac{\alpha}{2} < 0, \end{aligned}$$

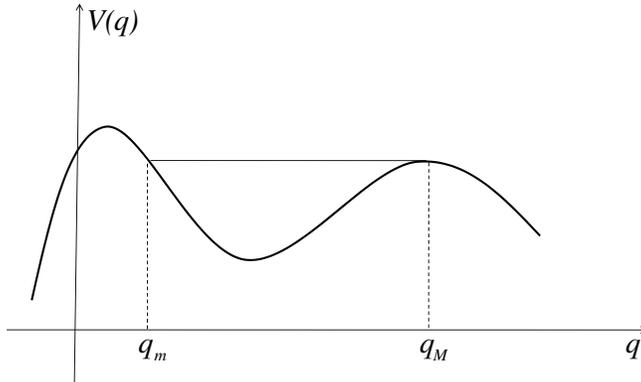
che appunto significa inversione del moto. Quindi il sistema “riparte” dalla posizione $q = q_M$, muovendosi questa volta verso q_m . La procedura si ripete per il percorso che va da q_0 verso q_m , e quindi otteniamo la (2.42) come formula per il tempo T impiegato

¹⁴Si ricordi che abbiamo assunto $t_0 = 0$.

¹⁵Osserviamo che l'integrale $\int_{q_0}^{q_M} \frac{d\eta}{\sqrt{q_M - \eta}}$, è convergente, infatti

$$\int_{q_0}^{q_M} \frac{d\eta}{\sqrt{q_M - \eta}} = \int_0^{q_M - q_0} \frac{ds}{\sqrt{s}} = 2\sqrt{q_M - q_0}.$$

Figura 2.2: Moto aperiodico.



dalla soluzione per ritornare al punto di partenza. Poiché l'equazione è autonoma, le soluzioni sono invarianti per traslazioni temporali¹⁶, ovvero la soluzione che parte da q_0 con velocità v_0 al tempo $t = T$, è ottenuta per traslazione da quella che, al tempo $t = 0$, partiva da q_0 con la stessa velocità v_0 : in altri termini $q(t) = q(t - T)$. Ne segue che il moto è periodico.

Il periodo T , ad eccezione di particolari forme di $V(q)$ (che verranno studiate nella sezione 2.9), dipende da E . Ma poiché quest'ultimo è definito dalla posizione e velocità iniziali, avremo che T , in generale, dipende dalle condizioni iniziali.

Caso 1.b. Moto aperiodico

Consideriamo la figura 2.2, in cui la condizione iniziale q_0 appartiene al solito intervallo limitato a sinistra da q_m e a destra da q_M . La soluzione $q(t)$ non è periodica. Questo caso è caratterizzato dal fatto che un valore di energia $E = V(q_M)$ corrispondente a un massimo relativo isolato q_M . Di conseguenza q_M è **almeno radice doppia dell'equazione** $V(q) = E$. Quindi, ricordando la (2.43), per q in un intorno di q_M scriveremo

$$F(q) = (q_M - q)^2 \Xi(q),$$

con $\Xi(q) \geq 0$.

Vediamo qual'è il comportamento della soluzione quando q si avvicina al valore q_M . Per fissare le idee supponiamo di aver scelto come condizioni iniziali, al tempo $t = 0$, $q_0 < q_M$ e $v_0 > 0$. Possiamo scrivere quindi la soluzione corrispondente usando la formula implicita (2.33). In particolare, sfruttando la (2.44), possiamo determinare

¹⁶Se $q(t)$ è soluzione di $\ddot{q} = f(q)$, è facile verificare che anche $q(t - T)$ è soluzione.

il tempo t_M impiegato per “raggiungere” q_M . Si ottiene¹⁷

$$t_M = \int_{q_0}^{q_M} \frac{d\eta}{\sqrt{(q_M - \eta)^2 \Xi(\eta)}} \geq \frac{1}{\sqrt{\max \Xi}} \underbrace{\int_{q_0}^{q_M} \frac{d\eta}{q_M - \eta}}_{\text{integrale divergente}} = +\infty.$$

Ne segue che la soluzione $q(t)$, pur essendo monotona crescente, non raggiunge mai la posizione q_M . Alla stessa conclusione saremmo arrivati se avessimo considerato $v_0 < 0$. In questo caso la soluzione avrebbe impiegato un tempo finito per raggiungere q_m (supponendo quest’ultima una radice semplice) e ritornare nella posizione q_0 con velocità positiva. Dopo di che si ripete la procedura appena descritta, deducendo che la soluzione impiega un tempo infinito per giungere sino a q_M . Il moto però non è monotono siccome avviene un’inversione in corrispondenza di q_m .

Il moto che abbiamo appena descritto verrà ripreso anche nel Caso 4, dove tratteremo il caso generale del moto lungo le *separatrici*.

Caso 2. I punti di equilibrio

I valori estremi di E che corrispondono ai valori di minimo o massimo relativo della funzione $V(q)$ generano delle *soluzioni di equilibrio* ovvero delle soluzioni che sono funzioni costanti del tempo. Riferendoci alla figura 2.3 e ricordando la (2.30) abbiamo che $q(t) = q_e = \text{costante}$, è soluzione dell’equazione di moto se $\left. \frac{dV}{dq} \right|_{q_e} = 0$. Infatti in questo caso al sistema competono velocità e accelerazione nulle, e quindi il sistema resta fermo nel punto di estremo q_e .

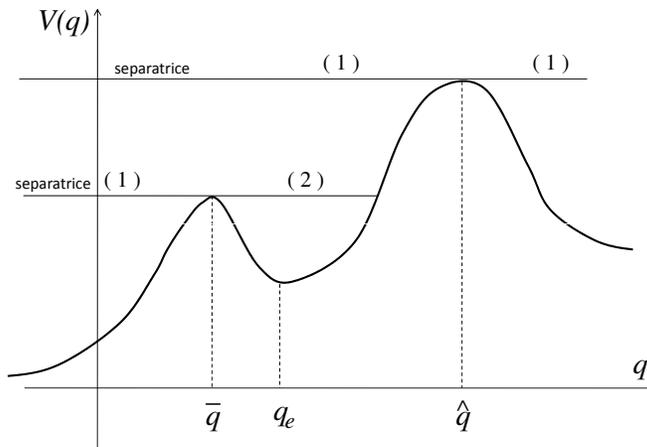
Se poi il punto di minimo q_e , è un minimo isolato (ovvero esiste un intorno $q_1 < q_e < q_2$ in cui $V(q) > V(q_e)$ per $q \neq q_e$) allora la posizione di equilibrio è *stabile*: ogni moto che parte “sufficientemente vicino” a q_e e con velocità “sufficientemente piccola” resta “vicino” a q_e per ogni tempo t (approfondiremo questo concetto nella sezione 2.8). Questo fatto è di immediata verifica osservando che, se il moto si trova a un certo istante in un punto q_0 vicino a q_e , ed ha velocità v_0 piccola in modulo, al valore di energia $E_1 = \frac{1}{2}v_0^2 + V(q_0)$ corrisponde un intervallo $E_1 \geq V(q)$ che è un intorno “piccolo” di q_e . Possiamo quindi concludere che il moto periodico descritto nel Caso 1.a avviene soltanto in un intorno di un punto di minimo di $V(q_e)$, ovvero di un punto di stabilità.

I massimi relativi della funzione $V(q)$ corrispondono ancora a posizioni di equilibrio: infatti ha ancora $\left. \frac{dV}{dq} \right|_{q_e} = 0$. In questo caso però, per quanto si considerino condizioni iniziali (q_0, v_0) “vicine” a $(q_e, 0)$, le corrispondenti soluzioni che non restano “vicine” a $(q_e, 0)$. In tal caso parleremo di equilibrio instabile.

¹⁷L’integrale $\int_{q_0}^{q_M} \frac{d\eta}{q_M - \eta}$ è divergente. Infatti

$$\int_{q_0}^{q_M} \frac{d\eta}{q_M - \eta} = \int_0^{q_M - q_0} \frac{ds}{s} = \lim_{s \rightarrow 0} \left(\ln \frac{q_M - q_0}{s} \right) = +\infty.$$

Figura 2.3: Punti di equilibrio e separatrici.



Caso 3. Moto illimitato

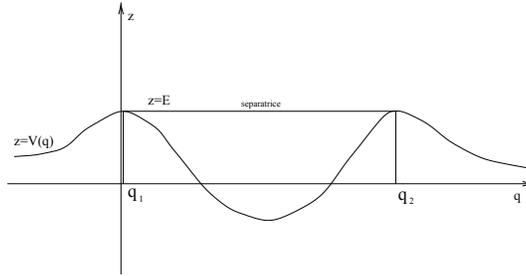
Considerando la figura 2.1, questo caso corrisponde alle rette orizzontali (2) e (3). L'insieme delle q che soddisfano la (2.41) o è tutto l'asse reale (retta (3) della figura 2.1) o comunque è un intervallo illimitato a destra o a sinistra (retta (2) della figura 2.1). Nel primo caso \dot{q} non si annulla mai e quindi a seconda della velocità iniziale il moto si svolge nel verso delle q positive oppure nel verso di quelle negative: la soluzione $q(t)$ è **strettamente monotona**.

Nel secondo caso esistono valori di q (le intersezioni fra $V(q)$ e la retta E) in cui \dot{q} si annulla. Tali intersezioni corrispondono comunque a radici semplici dell'equazione $E = V(q)$, e pertanto, se la v_0 (velocità iniziale) ha il segno "giusto", possono venir raggiunte in un tempo finito. Per esempio, riferendoci alla retta (2) di figura 2.1, se $q(0) = q_m$, e $v_0 > 0$, la soluzione raggiunge in un tempo finito l'intersezione di sinistra fra la retta (2) e $V(q)$, dopo di che il moto si inverte e si svolgerà nel verso delle q negative. Se viceversa avessimo avuto $v_0 < 0$, allora $q(t)$ decresce monotonamente senza mai invertire il moto.

Per riassumere, in questo caso abbiamo che, comunque si scelgano le condizioni iniziali, $\lim_{t \rightarrow \infty} q(t) = \pm\infty$. In particolare, se $V(q)$ è tale che $V(q) \xrightarrow{|q| \rightarrow \infty} \text{costante}$, allora la (2.40) comporta che anche $\dot{q} \rightarrow \text{costante}$, quando $t \rightarrow \infty$.

Caso 4. Le separatrici

Volendo caratterizzare questo caso, possiamo dire che, a differenza del precedente, il moto è (o può essere) limitato (cioè $|q|$ non tende all'infinito per $t \rightarrow \infty$) ma non periodico. Il primo tipo di separatrice è quella descritta nel Caso 1.b, moto aperiodico, e trova rappresentazione nel segmento denotato con (2) di figura 2.3. In questo caso una

Figura 2.4: Punti di massimo relativo di $V(q)$ con lo stesso valore dell'energia.

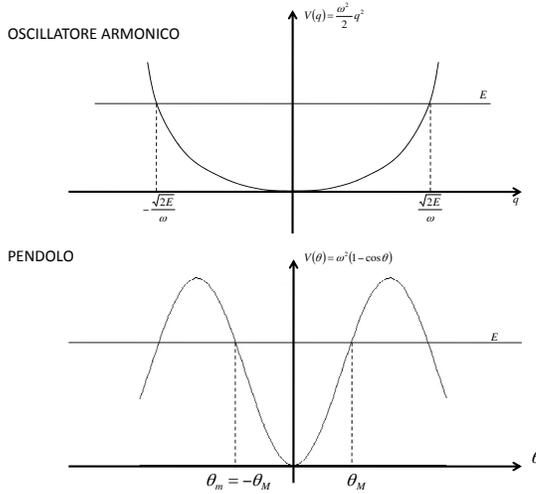
delle due radici di $V(q) = E$ (o anche entrambe) corrisponde ad un massimo isolato di V (riferendoci alla figura 2.3, la radice corrispondente al massimo è \bar{q}). La soluzione $q(t)$, ancorché limitata, non è periodica in quanto impiega un tempo infinito per raggiungere l'estremo corrispondente al massimo isolato. Il moto può essere monotono o meno (ovvero presentare un'inversione) a seconda della condizione iniziale sulla velocità.

Il secondo tipo di separatrice è quello che si ha in corrispondenza delle semirette indicate con (1) in figura 2.3. In tal caso si considerano gli intervalli illimitati $(-\infty, \bar{q})$, $(-\infty, \hat{q})$ oppure $(\hat{q}, +\infty)$. Quindi se $q_0 \in (-\infty, \bar{q})$ oppure $q_0 \in (-\infty, \hat{q})$, e la velocità iniziale v_0 è **positiva**, la soluzione $q(t)$ è monotona crescente ma impiega un tempo infinito per raggiungere \bar{q} , oppure \hat{q} . In ogni caso il moto è limitato: $q_0 \leq q(t) < \bar{q}$, con $\lim_{t \rightarrow \infty} q(t) = \bar{q}$; oppure $q_0 \leq q(t) < \hat{q}$, con $\lim_{t \rightarrow \infty} q(t) = \hat{q}$. In altri termini, abbiamo a che fare con un moto del tutto analogo a quello, privo d'inversioni, descritto nel Caso **1.b**. Viceversa v_0 è **negativa**, $q(t)$ è monotona decrescente e $\lim_{t \rightarrow \infty} q(t) = -\infty$ (se ovviamente $V(q) > V(\bar{q})$, per ogni $q < \bar{q}$).

Se $q_0 \in (\hat{q}, \infty)$, e $v_0 < 0$, allora $q(t)$ è monotona decrescente $\lim_{t \rightarrow \infty} q(t) = \hat{q}$. Se invece $v_0 > 0$, allora $q(t)$ si allontana indefinitamente da \hat{q} (ammesso che $V(q) > V(\hat{q}), \forall q > \hat{q}$).

Soluzioni limitate e monotone con $\lim_{t \rightarrow \infty} q(t) = \hat{q}_1$, oppure con $\lim_{t \rightarrow \infty} q(t) = \hat{q}_2$, si hanno quando \hat{q}_1 e \hat{q}_2 sono due punti di massimo relativo per i quali $V(\hat{q}_1) = V(\hat{q}_2)$, vedi figura 2.4. In questo particolare caso abbiamo un moto aperiodico privo d'inversioni indipendentemente dalla velocità iniziale.

Esempio 2.6.2 Calcoliamo il periodo dell'oscillatore armonico. Ricordando l'esempio 2.6.1, abbiamo $V(q) = \frac{\omega^2}{2}q^2$, per cui, per $E > 0$, l'intervallo $[q_m, q_M]$ corrispondente a $V(q) \leq E$, è $\left[-\frac{\sqrt{2E}}{\omega}, \frac{\sqrt{2E}}{\omega} \right]$ (si veda la figura 2.5). Dalla (2.42)

Figura 2.5: $V(q)$ nel caso dell'oscillatore armonico e nel caso del pendolo.

otteniamo¹⁸

$$\begin{aligned}
 T &= 2 \int_{-\frac{\sqrt{2E}}{\omega}}^{\frac{\sqrt{2E}}{\omega}} \frac{dq}{\sqrt{2\left(E - \frac{\omega^2}{2}q^2\right)}} = 4 \int_0^{\frac{\sqrt{2E}}{\omega}} \frac{dq}{\sqrt{2E - \omega^2 q^2}} \\
 &= \frac{4}{\omega} \int_0^{\frac{\sqrt{2E}}{\omega}} \frac{dq}{\sqrt{\left(\frac{\sqrt{2E}}{\omega}\right)^2 - q^2}} = \frac{4}{\omega} \left[\arcsin\left(\frac{q}{\frac{\sqrt{2E}}{\omega}}\right) \right]_0^{\frac{\sqrt{2E}}{\omega}} = \frac{2\pi}{\omega}.
 \end{aligned}$$

Notiamo che T non dipende da E , e quindi è indipendente dalle condizioni iniziali¹⁹. Osserviamo che $q = 0$ è un minimo assoluto, isolato, per l'energia potenziale, e quindi è un equilibrio stabile. Notiamo inoltre che $\omega = V''(0)$.

Esempio 2.6.3 Calcoliamo il periodo del pendolo semplice. Ricordando la (2.34), calcoliamo $V(\theta)$ trovando una primitiva qualsiasi di $-\omega^2 \sin \theta$, dove $\omega^2 = g/l$. Abbiamo

$$V(\theta) = - \int_0^\theta (-\omega^2 \sin \theta') d\theta' = \omega^2 (1 - \cos \theta) = 2\omega^2 \sin^2\left(\frac{\theta}{2}\right).$$

¹⁸ $\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin\left(\frac{x}{a}\right)$.

¹⁹Questo fatto è ovvio se si guarda la forma della soluzione generale (2.38); è importante notare però che il risultato è stato ottenuto senza risolvere l'equazione.

e di conseguenza, se $0 \leq E < 2\omega^2$, il periodo è

$$T = 2 \int_{-\theta_M}^{\theta_M} \frac{d\theta}{\sqrt{2 \left(E - 2\omega^2 \sin^2 \left(\frac{\theta}{2} \right) \right)}}, \quad (2.45)$$

dove $\theta_M \in [0, \pi]$, è la soluzione di $2\omega^2 \sin^2 \left(\frac{\theta}{2} \right) = E$. La costante E , che corrisponde all'energia cinetica più potenziale della massa unitaria, si calcola in base ai dati iniziali

$$E = \frac{\dot{\theta}^2(0)}{2} + V(0) = \frac{\dot{\theta}^2(0)}{2} + 2\omega^2 \sin^2 \left(\frac{\theta(0)}{2} \right).$$

Quindi supponendo $\dot{\theta}(0) = 0$, il pendolo parte da fermo, e $\theta(0) = \theta_0$, abbiamo $E = 2\omega^2 \sin^2 \left(\frac{\theta_0}{2} \right)$. Quindi l'intervallo, centrato nell'origine, corrispondente a $V(\theta) \leq E$, è $[-\theta_0, \theta_0]$. Possiamo quindi riscrivere la (2.45) come

$$\begin{aligned} T &= 2 \int_{-\theta_0}^{\theta_0} \frac{d\theta}{\sqrt{2 \left(2\omega^2 \sin^2 \left(\frac{\theta_0}{2} \right) - 2\omega^2 \sin^2 \left(\frac{\theta}{2} \right) \right)}} \\ &= \frac{2}{\omega} \int_0^{\theta_0} \frac{d\theta}{\sin \frac{\theta_0}{2} \sqrt{1 - \frac{\sin^2 \frac{\theta}{2}}{\sin^2 \frac{\theta_0}{2}}}}. \end{aligned}$$

Ponendo $\beta = \sin \frac{\theta_0}{2} < 1$, ed operando il cambio di variabile $y = \frac{\sin \frac{\theta}{2}}{\sin \frac{\theta_0}{2}} = \frac{1}{\beta} \sin \left(\frac{\theta}{2} \right)$, che implica

$$dy = \frac{1}{2\beta} \cos \left(\frac{\theta}{2} \right) d\theta = \frac{1}{2\beta} \sqrt{1 - \sin^2 \left(\frac{\theta}{2} \right)} d\theta = \frac{1}{2\beta} \sqrt{1 - \beta^2 y^2} d\theta,$$

otteniamo la seguente espressione per il periodo

$$T = \frac{4}{\omega} \int_0^1 \frac{dy}{\sqrt{1 - y^2} \sqrt{1 - \beta^2 y^2}}.$$

Siccome $\beta y = \sin \left(\frac{\theta}{2} \right) < 1$, possiamo sviluppare in serie di Taylor $1/\sqrt{1 - \beta^2 y^2}$,

$$\frac{1}{\sqrt{1 - \beta^2 y^2}} = 1 + \frac{\beta^2 y^2}{2} + \frac{3}{8} \beta^4 y^4 + \dots,$$

ottenendo così

$$\begin{aligned}
 T &= \frac{4}{\omega} \int_0^1 \frac{1}{\sqrt{1-y^2}} \left(1 + \frac{\beta^2 y^2}{2} + \frac{3}{8} \beta^4 y^4 + \dots \right) dy \\
 &= \frac{4}{\omega} \underbrace{\int_0^1 \frac{dy}{\sqrt{1-y^2}}}_{\pi/2} + \frac{2}{\omega} \underbrace{\int_0^1 \frac{\beta^2 y^2}{\sqrt{1-y^2}} dy}_{\beta^2 \pi/4} + \dots \\
 &= \frac{2\pi}{\omega} \left[1 + \frac{\beta^2}{4} + \dots \right].
 \end{aligned}$$

Ricordando che $\omega = \sqrt{g/l}$, e che $\beta = \sin(\theta_0/2)$, abbiamo

$$T = 2\pi \sqrt{\frac{l}{g}} \left[1 + \frac{1}{4} \sin^2 \frac{\theta_0}{2} + \dots \right],$$

che mette in luce la ben nota non isocronia del pendolo: il periodo del pendolo, a differenza di quello dell'oscillatore armonico, dipende dal punto di partenza. E' evidente che se $\theta_0 \ll 1$, ovvero se il pendolo viene lasciato andare da una posizione prossima alla verticale, allora possiamo trascurare $\sin^2 \frac{\theta_0}{2}$, e gli ordini superiori giungendo

all'espressione "classica" $T \approx 2\pi \sqrt{\frac{l}{g}}$.

Osservando la figura 2.5, possiamo dedurre le seguenti proprietà qualitative del moto: (i) se $0 \leq E < 2\omega^2$, abbiamo un moto periodico e limitato (di cui abbiamo calcolato il periodo); (ii) se $E = 2\omega^2$ abbiamo la separatrice, un moto limitato ma aperiodico; (iii) se $E > 2\omega^2$, il moto è illimitato. In tal caso l'energia iniziale E è "sufficiente" per far ruotare il pendolo intorno al proprio asse. Il caso (ii) corrisponde, per esempio, al caso limite in cui il pendolo parte praticamente dalla posizione verticale superiore ed impiega un tempo infinito per ritornarvi.

Esempio 2.6.4 Analizziamo qualitativamente le soluzioni nel caso in cui

$$V(q) = V_o (e^{-2q} - 2e^{-q}), \quad V_o > 0. \quad (2.46)$$

Il grafico di $V(q)$ è riportato nella figura 2.6.

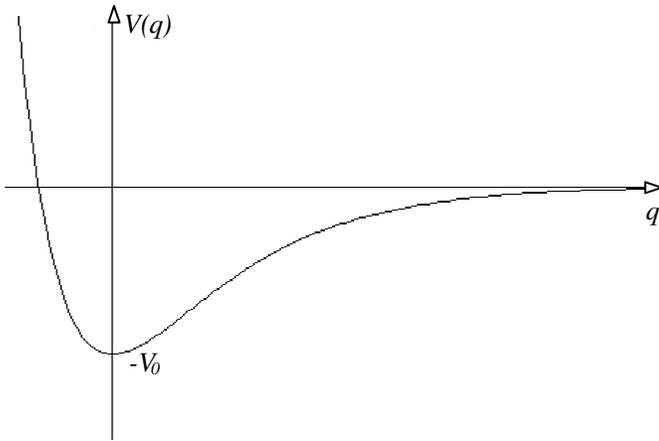
Notiamo subito che $E < -V_o$, non è ammissibile. Per $E = -V_o$ abbiamo un punto di equilibrio stabile, corrispondente a $q = 0$, mentre per $-V_o < E < 0$, abbiamo soluzioni periodiche limitate. $E = 0$, dà luogo alla separatrice, che corrisponde ad un moto illimitato ma anche aperiodico, mentre per $E > 0$ abbiamo moti illimitati.

Calcoliamo adesso il periodo delle oscillazioni supponendo $E = -\beta V_o$, con $\beta \in (0, 1)$. Dobbiamo determinare le intersezioni di $V(q)$ con la retta $-\beta V_o$,

$$-\beta V_o = V_o (e^{-2q} - 2e^{-q}).$$

Posto $\xi = e^q$, abbiamo questa equazione da risolvere

$$\xi^2 - \frac{2}{\beta} \xi + \frac{1}{\beta} = 0 \quad (2.47)$$

Figura 2.6: Grafico di $V(q)$, dato dalla 2.46.

le cui soluzioni sono

$$\begin{cases} \xi_+ = \frac{1 + \sqrt{1 - \beta}}{\beta}, \\ \xi_- = \frac{1 - \sqrt{1 - \beta}}{\beta}, \end{cases} \quad \Rightarrow \quad \begin{cases} q_+ = \ln \left(\frac{1 + \sqrt{1 - \beta}}{\beta} \right), \\ q_- = \ln \left(\frac{1 - \sqrt{1 - \beta}}{\beta} \right). \end{cases}$$

Ricordando che $|E| = \beta V_0$, il periodo è dunque

$$T = 2 \int_{q_-}^{q_+} \frac{dq}{\sqrt{2[-\beta V_0 - V_0(e^{-2q} - 2e^{-q})]}} = \frac{2}{\sqrt{2|E|}} \int_{q_-}^{q_+} \frac{dq}{\sqrt{-1 - \frac{e^{-2q}}{\beta} + \frac{2e^{-q}}{\beta}}}.$$

Ponendo adesso $\xi = e^q$, e tenendo presente la (2.47), l'integrale si trasforma in

$$T = \frac{2}{\sqrt{2|E|}} \int_{\xi_-}^{\xi_+} \frac{d\xi}{\sqrt{-\left(\xi^2 - \frac{2\xi}{\beta} + \frac{1}{\beta}\right)}} = \frac{2}{\sqrt{2|E|}} \underbrace{\int_{\xi_-}^{\xi_+} \frac{d\xi}{\sqrt{(\xi_+ - \xi)(\xi - \xi_-)}}}_{\pi \text{ (integrale notevole)}},$$

$$\text{cioè } T = \frac{2\pi}{\sqrt{2|E|}}.$$

2.6.2 Periodo delle oscillazioni in vicinanza di punti di equilibrio stabile

Supponiamo che $V(q)$ abbia un minimo isolato in $q = q_e$, che per quanto accennato nel caso 2 della sezione 2.6.1, è un punto di equilibrio stabile (la definizione rigorosa di punto di equilibrio stabile sarà illustrata nel paragrafo 2.8).

Il periodo di oscillazione delle soluzioni attorno ad un minimo q_e di $V(q)$, dipende, in generale, dal valore di E (e quindi dalle condizioni iniziali). Tuttavia, se E , benché maggiore di $V(q_e)$, è “vicino” a $V(q_e)$, è possibile trovare un’espressione approssimata del periodo (che sarà tanto più esatta quanto più $E - V(q_e)$ è “piccolo”) indipendente da E . Considerando un intorno di q_e , si approssima $V(q)$ con il suo sviluppo di Taylor (limitato al secondo ordine)

$$V(q) \approx V(q_e) + \underbrace{V'(q_e)}_{=0} (q - q_e) + \frac{1}{2} V''(q_e) (q - q_e)^2, \quad (2.48)$$

dove $V''(q_e) > 0$ (si ricordi che q_e è un minimo isolato). Gli estremi dell’intervallo dove si svolge il moto sono dati da

$$V(q_e) + \frac{1}{2} V''(q_e) (q - q_e)^2 = E, \implies \begin{cases} q_m = q_e - \sqrt{\frac{2(E - V(q_e))}{V''(q_e)}}, \\ q_M = q_e + \sqrt{\frac{2(E - V(q_e))}{V''(q_e)}}. \end{cases}$$

Quindi, introducendo $\xi = q - q_e$, si ha la seguente espressione approssimata di T

$$\begin{aligned} T &\approx 2 \int_{-\sqrt{\frac{2(E - V(q_e))}{V''(q_e)}}}^{\sqrt{\frac{2(E - V(q_e))}{V''(q_e)}}} \frac{d\xi}{\sqrt{2 \left[E - V(q_e) - \frac{1}{2} V''(q_e) \xi^2 \right]}} \\ &= \frac{2}{\sqrt{V''(q_e)}} \int_{-\sqrt{\frac{2(E - V(q_e))}{V''(q_e)}}}^{\sqrt{\frac{2(E - V(q_e))}{V''(q_e)}}} \frac{d\xi}{\sqrt{\frac{2(E - V(q_e))}{V''(q_e)} - \xi^2}}. \end{aligned}$$

Da cui, ricordando che $\int_{-a}^a \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin\left(\frac{x}{a}\right) \Big|_{-a}^a = \pi$, ricaviamo

$$T \approx \frac{2\pi}{\sqrt{V''(q_e)}}. \quad (2.49)$$

Quando $E \rightarrow V(q_e)$ (livello di energia della posizione di equilibrio) il periodo (approssimato) delle piccole oscillazioni non dipende da E , ma soltanto da $V''(q_e)$. In pratica, ricordando l’esempio 2.6.2, la (2.49) altro non è che il periodo di un oscillatore armonico in cui $\omega = \sqrt{V''(q_e)}$.

2.6.3 Generalizzazione del caso conservativo

In questa sezione vogliamo considerare una forma più generale dell'equazione (2.27), analizzando le equazioni del tipo

$$a(q)\ddot{q} + \frac{a'(q)}{2}\dot{q}^2 = f(q), \quad (2.50)$$

dove $a(q)$ è una **funzione strettamente positiva**, e $a'(q) = \frac{da(q)}{dq}$. Introduciamo, al solito, l'energia potenziale $V(q)$ tramite la (2.29), la (2.50) potrà essere riscritta come

$$a(q)\ddot{q} + \frac{a'(q)}{2}\dot{q}^2 = -V'(q), \quad (2.51)$$

essendo $V'(q) = \frac{dV(q)}{dq}$. Procedendo come nella sezione 2.6, se moltiplichiamo la (2.51) per \dot{q} , è immediato verificare che

$$\frac{d}{dt} \left(\frac{a(q)}{2}\dot{q}^2 + V(q) \right) = 0, \quad \implies \quad \frac{a(q)}{2}\dot{q}^2 + V(q) = E, \quad (2.52)$$

dove, al solito, l'energia E è specificata dalle condizioni iniziali $q(0)$, e $\dot{q}(0)$. Il termine $\frac{a(q)}{2}\dot{q}^2$ è detto **energia cinetica** e viene usualmente indicato con T . Tuttavia, limitatamente a questa sezione, poniamo $\mathcal{E} = \frac{a(q)}{2}\dot{q}^2$, al fine di evitare confusione di simboli.

Anche in questo caso il moto potrà svolgersi solo in una componente connessa di $V(q) \leq E$, e la (2.33) diventa

$$t - t_0 = \int_{q_0}^{q(t)} \sqrt{\frac{a(\eta)}{2(E - V(\eta))}} d\eta.$$

L'analisi effettuata per il caso $a = 1$, sulla periodicità o aperiodicità del moto può essere ripetuta in maniera analoga.

Come vedremo nella sezione 2.8, i punti di equilibrio stabile dell'equazione (2.51) risultano essere sempre i minimi isolati di $V(q)$. Vediamo allora come si generalizza la formula (2.49) nel caso in cui l'energia cinetica sia $\frac{a(q)}{2}\dot{q}^2$. Se introduciamo $\xi = q - q_e$, abbiamo $\dot{\xi} = \dot{q}$, e sviluppando l'energia cinetica in un intorno di q_e , cioè di $\xi = 0$, abbiamo

$$\mathcal{E} \approx \frac{1}{2} \left[a(q_e) + a'(q_e)\xi + \frac{1}{2}a''(q_e)\xi^2 \right] \dot{\xi}^2.$$

Ora se vogliamo un'approssimazione limitata ai termini quadratici in ξ , e $\dot{\xi}$, come si dice un'approssimazione al secondo ordine, otteniamo

$$\mathcal{E} \approx \frac{1}{2}a(q_e)\dot{\xi}^2.$$

Di conseguenza, ricordando la (2.48), la riscrittura approssimata della (2.52) è

$$\frac{1}{2}a(q_e)\dot{\xi}^2 + \frac{1}{2}V''(q_e)\xi^2 = E - V(q_e),$$

che è esattamente la (2.37) dell'oscillatore armonico, se poniamo

$$\omega^2 = \frac{V''(q_e)}{a(q_e)},$$

ed $\frac{E - V(q_e)}{a(q_e)}$ è l'energia meccanica. Pertanto, siccome il periodo dell'oscillatore armonico è $2\pi/\omega$, abbiamo che il periodo (approssimato) delle piccole oscillazioni attorno a q_e è

$$T \approx 2\pi \sqrt{\frac{a(q_e)}{V''(q_e)}}.$$

2.7 Il piano delle fasi

Abbiamo già visto che ogni equazione differenziale scalare del secondo ordine, $\ddot{q} = f(\dot{q}, q, t)$ può essere sempre trasformata in un sistema di due equazioni in due incognite, semplicemente definendo $v = \dot{q}$ e quindi ponendo

$$\begin{cases} \dot{q} = v, \\ \dot{v} = f(v, q, t). \end{cases} \quad (2.53)$$

La soluzione di (2.53) è ora una coppia di funzioni $(q(t), v(t))$. Possiamo rappresentare la soluzione tramite un punto che si muove in un piano cartesiano di ascissa q e ordinata v . Questa rappresentazione è particolarmente utile nel caso dei sistemi autonomi, il piano (q, v) così definito è detto **piano delle fasi**.

$$\begin{cases} \dot{q} = v, \\ \dot{v} = f(v, q). \end{cases} \quad (2.54)$$

Definiamo **orbita** del sistema il luogo dei punti del piano $\{(q, v) \mid q = q(t), v = v(t), -\infty < t < +\infty\}$, dove $(q(t), v(t))$ è una soluzione di (2.54). Poiché, in virtù del teorema di esistenza e unicità, nel caso di un sistema autonomo due soluzioni $(q_1(t), v_1(t))$ e $(q_2(t), v_2(t))$ che corrispondono alle stesse condizioni iniziali assunte a due istanti diversi, t_1 e t_2 (cioè $(q_1(t_1) = q_0, v_1(t_1) = v_0)$ e $(q_2(t_2) = q_0, v_2(t_2) = v_0)$) sono l'una la traslata temporale dell'altra, i.e. $q_1(t) = q_2(t - (t_2 - t_1))$, $v_1(t) = v_2(t - (t_2 - t_1))$, ne risulta che le orbite di queste due soluzioni coincidono, ovvero un punto (q, v) appartiene all'orbita della prima soluzione se e solo se appartiene all'orbita della seconda. Di conseguenza per ogni punto del piano delle fasi passa una e una sola orbita.

Nella figura 2.7 abbiamo riportato un esempio delle orbite nel piano delle fasi dell'equazione (2.27), ovvero del sistema

$$\begin{cases} \dot{q} = v, \\ \dot{v} = f(q) = -\frac{dV(q)}{dq}. \end{cases}$$

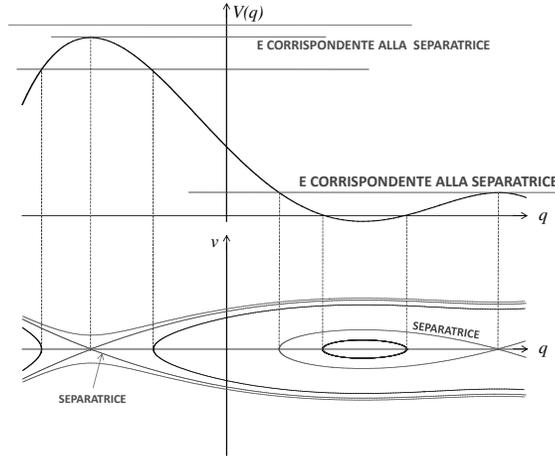


Figura 2.7: Orbite sul piano delle fasi.

Le orbite corrispondono al $V(q)$ rappresentato nella parte alta della figura. In questo caso la rappresentazione delle orbite sul piano delle fasi è agevole. Infatti, sfruttando la (2.40) abbiamo

$$v = \pm \sqrt{2(E - V(q))}.$$

Le curve che rappresentano le orbite (come mostrato nella parte inferiore della figura 2.7) sono simmetriche rispetto all'asse q e si ottengono tracciando schematicamente il grafico di

$$v = \sqrt{2(E - V(q))}, \text{ e poi di } v = -\sqrt{2(E - V(q))}.$$

Quest'ultimo poi si ottiene semplicemente ribaltando il primo grafico rispetto all'asse q . Evidentemente il dominio dove disegnare il grafico $v = \sqrt{2(E - V(q))}$, sarà individuato, una volta fissato E , da punti q che soddisfano $V(q) \leq E$. Inoltre, l'andamento del grafico è facilmente rappresentabile in quanto v è, in prima approssimazione, "proporzionale" alla differenza fra $V(q)$ ed E : maggiore è tale differenza maggiore sarà v .

La rappresentazione delle orbite nel caso di equazioni del tipo (2.51) non è, in generale, fattibile senza l'ausilio di un programma di calcolo. Infatti la (2.52) comporta

$$v = \pm \sqrt{\frac{2(E - V(q))}{a(q)}}.$$

Ora, mentre è facile determinare il dominio della funzione (è esattamente lo stesso procedimento di prima, $V(q) \leq E$), non è assolutamente facile tracciare il grafico di v in funzione di q . In questo caso, infatti, non c'è più "proporzionalità" fra v e la differenza fra E e $V(q)$, a causa del fattore $1/a(q)$.

2.8 Punti di equilibrio, stabilità

Un punto del piano delle fasi (q_e, v_e) si dice **punto di equilibrio** per il sistema autonomo (2.54) se la coppia di funzioni $q(t) = q_e \forall t$, e $v(t) = v_e \forall t$, è soluzione del sistema (2.54). In virtù della forma del sistema (2.54), si ha immediatamente $v_e = 0$, mentre q_e deve essere una soluzione dell'equazione (non più differenziale) $f(q_e, 0) = 0$. Talvolta q_e viene anche detta **configurazione di equilibrio**.

Come abbiamo visto nelle precedenti sezioni, nel caso dell'equazione $\ddot{q} = f(q)$ una configurazione di equilibrio q_e è un punto di massimo e minimo relativo della funzione $V(q)$, siccome $f(q) = -V'(q)$. Abbiamo già osservato che i minimi (isolati) della V sono punti di equilibrio "stabile". Vogliamo ora precisare questo concetto introducendo la definizione di **stabilità secondo Lyapunov**²⁰.

La definizione si applica in generale ai punti di equilibrio di un sistema di n equazioni differenziali in n incognite, o equazione differenziale in \mathbb{R}^n , e quindi la daremo direttamente in questo caso generale.

Consideriamo adesso un sistema di equazioni differenziali autonomo

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}), \quad (2.55)$$

dove $\mathbf{x} : t \rightarrow \mathbf{x}(t) \in \mathbb{R}^n$ è la funzione vettoriale incognita e $\mathbf{F} : \mathbf{x} \rightarrow \mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$ è una funzione da \mathbb{R}^n in \mathbb{R}^n . In seguito indicheremo con $\mathbf{x}(t; \mathbf{x}_0, t_0)$ la soluzione al tempo t che al tempo t_0 occupava il punto \mathbf{x}_0 (in altri termini la soluzione del sistema (2.55) corrispondente ai dati iniziali $\mathbf{x}(t_0) = \mathbf{x}_0$).

Definizione 2.8.1 Il punto $\mathbf{x}_e \in \mathbb{R}^n$ si dice **punto, o configurazione, di equilibrio** per il sistema autonomo (2.55), se la funzione $\mathbf{x}(t) \equiv \mathbf{x}_e$, è soluzione del problema di Cauchy

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}), \\ \mathbf{x}(t_0) \equiv \mathbf{x}_e. \end{cases}$$

E' immediato constatare che i punti di equilibrio \mathbf{x}_e sono tutti e soli quelle n -uple che soddisfano il sistema di equazioni

$$\mathbf{F}(\mathbf{x}_e) = 0, \quad \iff \begin{cases} F_1(x_1, \dots, x_n) = 0, \\ F_2(x_1, \dots, x_n) = 0, \\ \vdots \\ F_n(x_1, \dots, x_n) = 0. \end{cases} \quad (2.56)$$

Definizione 2.8.2 Un punto di equilibrio \mathbf{x}_e si dice **stabile** se: per ogni t_0 e per ogni $r > 0$, esiste un $r_0 > 0$ tale che sia verificata la disuguaglianza

$$\|\mathbf{x}(t; \mathbf{x}_0, t_0) - \mathbf{x}_e\| < r,$$

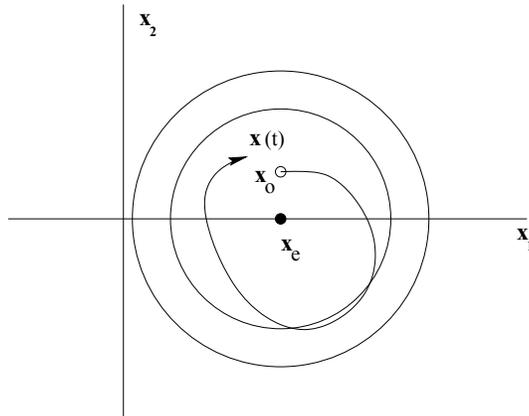
per ogni $t > t_0$ e per tutti i dati iniziali \mathbf{x}_0 che soddisfano²¹

$$\|\mathbf{x}_0 - \mathbf{x}_e\| < r_0.$$

²⁰Aleksandr Mikhailovich Lyapunov, matematico russo, 1857-1918

²¹Si noti la somiglianza tra questa definizione e la definizione "ε-δ" di limite.

Figura 2.8: Stabilità alla Lyapunov



Il concetto di stabilità del punto x_e è schematicamente rappresentato dalla figura 2.8. Il cerchio di raggio maggiore corrisponde alla circonferenza di raggio r . Fissato r , è possibile determinare r_0 tale che, una qualunque orbita che “parte” da un punto all’interno della circonferenza r_0 , rimane all’interno della circonferenza di raggio r .

2.8.1 Il criterio di Lyapunov

Una condizione sufficiente per stabilire se una posizione di equilibrio è stabile è data dal seguente teorema, sempre dovuto a Lyapunov.

Teorema 2.8.1 *Se una posizione di equilibrio x_e per il sistema (2.55), cioè x_e è soluzione della (2.56), ammette una funzione di Lyapunov, allora è stabile.*

Ovviamente per dar senso al teorema bisogna dire cosa è una **funzione di Lyapunov**.

Definizione 2.8.3 *Una funzione Λ definita in un intorno U_{x_e} del punto x_e , si dice **funzione di Lyapunov** per il punto di equilibrio x_e se:*

1. $\Lambda : U_{x_e} \rightarrow \mathbb{R}$ è una funzione di classe C^1 ;
2. $\Lambda(x_e) = 0$ e $\Lambda(x) > 0$ per ogni $x \neq x_e$;
3. $\nabla\Lambda(x) \cdot F(x) \leq 0$ per ogni $x \in U_{x_e}$.

Questa definizione²² necessita di qualche commento. A parte la richiesta “tecnica” di regolarità, la condizione 2 ci dice che la funzione Λ ha un minimo isolato²³ nel punto di equilibrio, mentre la terza condizione implica che la funzione $\Lambda(x(t))$, ottenuta

²²Con il simbolo ∇ si indica l’operatore gradiente, quindi $\nabla\Lambda(x) = (\frac{\partial\Lambda}{\partial x_1}, \dots, \frac{\partial\Lambda}{\partial x_n})$.

²³Un punto x_e è un minimo isolato per una funzione f se esiste un intorno di U di x_e tale che $f(y) > f(x_e)$ per ogni $y \in U$, $y \neq x_e$.

componendo $\Lambda(\mathbf{x})$ con una soluzione $\mathbf{x}(t)$ di (2.55), è una funzione non crescente di t . Infatti

$$\frac{d}{dt}\Lambda(\mathbf{x}(t)) = \nabla\Lambda(\mathbf{x}(t)) \cdot \frac{d\mathbf{x}}{dt} = \nabla\Lambda(\mathbf{x}(t)) \cdot \mathbf{F}(\mathbf{x}(t)) \leq 0. \quad (2.57)$$

Dim. Scegliamo²⁴ un numero $R > 0$ e sia $B_{\mathbf{x}_e}(R)$ la sfera, aperta, di centro \mathbf{x}_e e di raggio R . Sia inoltre $\lambda = \min_{\partial B_{\mathbf{x}_e}(R)} \Lambda(\mathbf{x})$ il valore minimo assunto dalla funzione Λ sul bordo questa sfera²⁵; per le nostre ipotesi si ha $\lambda > 0$.

Siccome Λ è di classe C^1 , esiste finito $\max_{\mathbf{x} \in \overline{B_{\mathbf{x}_e}(R)}} \|\nabla\Lambda(\mathbf{x})\|$, ed è sicuramente positivo²⁶. Poniamo pertanto

$$\mathfrak{G} = \max_{\mathbf{x} \in \overline{B_{\mathbf{x}_e}(R)}} \|\nabla\Lambda(\mathbf{x})\| > 0.$$

Dunque, $\forall \mathbf{x} \in \overline{B_{\mathbf{x}_e}(R)}$, abbiamo

$$\Lambda(\mathbf{x}) = \Lambda(\mathbf{x}) - \underbrace{\Lambda(\mathbf{x}_e)}_{=0} \leq \mathfrak{G} \|\mathbf{x} - \mathbf{x}_e\|. \quad (2.58)$$

Consideriamo adesso $r < \min\left\{\frac{\lambda}{2\mathfrak{G}}, \frac{R}{2}\right\}$, e l'intorno sferico di \mathbf{x}_e di raggio r , cioè $B_{\mathbf{x}_e}(r)$. E' chiaro che²⁷ $B_{\mathbf{x}_e}(r) \subset B_{\mathbf{x}_e}(R)$. Inoltre, in virtù della (2.58), $\forall \mathbf{x} \in B_{\mathbf{x}_e}(r)$, cioè $\forall \mathbf{x}$ tale che $\|\mathbf{x} - \mathbf{x}_e\| < r$, abbiamo

$$\Lambda(\mathbf{x}) \leq \mathfrak{G} \|\mathbf{x} - \mathbf{x}_e\| < \mathfrak{G} r < \mathfrak{G} \min\left\{\frac{\lambda}{2\mathfrak{G}}, \frac{R}{2}\right\} = \min\left\{\frac{\lambda}{2}, \frac{\mathfrak{G}R}{2}\right\} \leq \frac{\lambda}{2}.$$

In altri termini, abbiamo che $\Lambda(\mathbf{x}) < \lambda/2$, $\forall \mathbf{x} \in B_{\mathbf{x}_e}(r)$. Ora, la coppia R e r soddisfa le condizioni della definizione di equilibrio stabile secondo Lyapunov. Infatti se prendiamo una condizione iniziale \mathbf{x}_0 contenuta in $B_{\mathbf{x}_e}(r)$, si ha che $\Lambda(\mathbf{x}(t; \mathbf{x}_0, t_0)) \leq \Lambda(\mathbf{x}_0) < \lambda/2$, e quindi, $\forall t > t_0$, il punto $\mathbf{x}(t; \mathbf{x}_0, t_0)$ deve appartenere a $B_{\mathbf{x}_e}(r)$, e di conseguenza a $B_{\mathbf{x}_e}(R)$. Se infatti la soluzione $\mathbf{x}(t; \mathbf{x}_0, t_0)$ uscisse da $B_{\mathbf{x}_e}(R)$, allora ad un certo istante \hat{t} la soluzione toccherebbe il bordo di $B_{\mathbf{x}_e}(R)$, e quindi²⁸ $\Lambda(\mathbf{x}(\hat{t}; \mathbf{x}_0, t_0)) \geq \lambda$. Ma ciò contraddirebbe l'ipotesi di non crescita della funzione Λ vista come funzione del tempo dal momento che si avrebbe

$$\Lambda(\mathbf{x}_0) = \Lambda(t_0) < \Lambda(\hat{t}) = \Lambda(\mathbf{x}(\hat{t}; \mathbf{x}_0, t_0)).$$

□

2.8.2 Asintotica stabilità

Se \mathbf{x}_e , oltre a essere un punto di equilibrio stabile, soddisfa la seguente condizione: *esiste un intorno $U_{\mathbf{x}_e}$ di \mathbf{x}_e tale che per ogni t_0 e per ogni $\mathbf{x}_0 \in U_{\mathbf{x}_e}$ si ha*

$$\lim_{t \rightarrow +\infty} \mathbf{x}(t; \mathbf{x}_0, t_0) = \mathbf{x}_e$$

²⁴ R andrà scelto in modo che $B_{\mathbf{x}_e}(R) \subset U_{\mathbf{x}_e}$

²⁵Notiamo che il minimo esiste perché la funzione Λ è continua e perché $\partial B_{\mathbf{x}_e}(R)$ è un compatto.

²⁶Se infatti $\max_{\mathbf{x} \in \overline{B_{\mathbf{x}_e}(R)}} \|\nabla\Lambda(\mathbf{x})\| = 0$, allora $\nabla\Lambda(\mathbf{x}) \equiv 0$, e cioè $\Lambda(\mathbf{x}) \equiv 0$, $\forall \mathbf{x} \in B_{\mathbf{x}_e}(R)$, in palese contraddizione con l'ipotesi 2.

²⁷In particolare $B_{\mathbf{x}_e}(r)$ è separato dal bordo di $B_{\mathbf{x}_e}(R)$.

²⁸Si ricordi che $\lambda = \min_{\partial B_{\mathbf{x}_e}(R)} \Lambda(\mathbf{x})$.

la posizione si dice di equilibrio *asintoticamente stabile*.

Anche in questo caso abbiamo un criterio sufficiente per l'asintotica stabilità; se nella terza condizione per la funzione di Lyapunov sostituiamo il segno \leq con la disuguaglianza stretta $<$, che deve valere per tutti gli \mathbf{x} con l'eccezione di \mathbf{x}_e (dove ovviamente abbiamo il segno $=$ poiché sia \mathbf{F} che $\nabla\Lambda$ si annullano), allora la posizione di equilibrio è asintoticamente stabile.

Vediamo la dimostrazione di questa versione del criterio di Lyapunov.

Per prima cosa osserviamo che \mathbf{x}_e è una posizione di equilibrio stabile: dalla stabilità segue che possiamo scegliere due costanti positive R e r_0 in modo che tutte le soluzioni con dato iniziale in $B_{\mathbf{x}_e}(r_0)$ siano contenute in $B_{\mathbf{x}_e}(R)$ per tutti tempi successivi al tempo iniziale. Sia $\mathbf{x}(t)$ una tale soluzione. Poiché Λ è una funzione di Lyapunov, la funzione del tempo $\Lambda(\mathbf{x}(t))$ è una funzione decrescente e quindi esiste $\lim_{t \rightarrow \infty} \Lambda(\mathbf{x}(t)) = \lambda$. Poiché $\Lambda \geq 0$ in $B_{\mathbf{x}_e}(R)$, avremo $\lambda \geq 0$.

Se $\lambda = 0$, allora necessariamente $\mathbf{x}(t) \rightarrow \mathbf{x}_e$, poiché è l'unico punto dove $\Lambda = 0$. Supponiamo invece che sia $\lambda > 0$. Se $B_{\mathbf{x}_e}(\delta)$ denota la sfera di centro \mathbf{x}_e e raggio δ , scegliamo δ in modo che $\Lambda(\mathbf{x}) < \lambda/2$ per ogni²⁹ $\mathbf{x} \in \overline{B_{\mathbf{x}_e}(\delta)}$. Ora per t sufficientemente grande, $\mathbf{x}(t)$ non può appartenere alla sfera $B_{\mathbf{x}_e}(\delta)$; ovvero, per t sufficientemente grande, $\mathbf{x}(t) \in \overline{B_{\mathbf{x}_e}(R)} \setminus B_{\mathbf{x}_e}(\delta)$. Ma la derivata di $\Lambda(\mathbf{x}(t))$ è *strettamente negativa* per la condizione $\nabla\Lambda(\mathbf{x}) \cdot \mathbf{F}(\mathbf{x}) < 0$. Infatti, poiché $\nabla\Lambda(\mathbf{x}) \cdot \mathbf{F}(\mathbf{x})$ è una funzione continua³⁰ definita sul compatto $\overline{B_{\mathbf{x}_e}(R)} \setminus B_{\mathbf{x}_e}(\delta)$, per tutti t sufficientemente grandi

$$\frac{d}{dt}\Lambda(\mathbf{x}(t)) = \nabla\Lambda(\mathbf{x}(t)) \cdot \mathbf{F}(\mathbf{x}(t)) \leq \max_{\mathbf{x} \in \overline{B_{\mathbf{x}_e}(R)} \setminus B_{\mathbf{x}_e}(\delta)} \nabla\Lambda(\mathbf{x}) \cdot \mathbf{F}(\mathbf{x}) = C < 0.$$

Ma questo implica $\lim_{t \rightarrow \infty} \Lambda(\mathbf{x}(t)) = -\infty$, il che è assurdo poiché $\Lambda(\mathbf{x}(t)) \geq 0$.

Un esempio di configurazione di equilibrio asintoticamente stabile verrà illustrato nella sezione 2.11. Un criterio sufficiente per l'asintotica stabilità è dato dal seguente *teorema di Lyapunov sulla stabilità linearizzata*:

Teorema 2.8.2 *Sia \mathbf{x}_e un punto di equilibrio per il sistema (2.55). Sia \mathbb{A} la matrice delle derivate parziali delle componenti di \mathbf{F} calcolate in \mathbf{x}_e , ovvero*

$$\mathbb{A} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \dots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \dots & \frac{\partial F_n}{\partial x_n} \end{pmatrix}.$$

Se tutti gli autovalori di \mathbb{A} hanno parte reale negativa, la posizione di equilibrio è asintoticamente stabile. Più precisamente esiste un numero positivo δ tale che per ogni $\mathbf{x}_0 \in B_{\mathbf{x}_e}(\delta)$ sia ha

$$\|\mathbf{x}(t; t_0, \mathbf{x}_0) - \mathbf{x}_e\| < a\|\mathbf{x}_0 - \mathbf{x}_e\|e^{-\alpha t},$$

con a e α costanti indipendenti da \mathbf{x}_0 .

²⁹Notiamo che la sfera $B_{\mathbf{x}_e}(\delta)$ esiste per la continuità di Λ .

³⁰ Λ è di classe C^1 , e, ovviamente, F è continua

La dimostrazione di questo teorema è piuttosto complessa (ma non specialmente “difficile”) e viene omessa: si può trovare in molti testi sulle equazioni differenziali ordinarie³¹.

2.8.3 Punti di equilibrio per sistemi conservativi: il criterio di Dirichlet

Torniamo alla equazione $\ddot{q} = f(q)$, o meglio al sistema equivalente

$$\begin{cases} \dot{q} = v, \\ \dot{v} = f(q). \end{cases} \quad (2.59)$$

In accordo con quanto detto nel paragrafo precedente, un punto di equilibrio è dato, nel piano (v, q) , da $(0, q_e)$, dove q_e è tale che $f(q_e) = 0$, ovvero da $V'(q_e) = 0$, dove $V(q)$ è l'energia potenziale definita dalla (2.29). Consideriamo la seguente funzione di v e q

$$\Lambda(q, v) = \frac{1}{2}v^2 + V(q) - V(q_e). \quad (2.60)$$

A questo punto possiamo enunciare il seguente

Criterio di Dirichlet: *Una configurazione di equilibrio q_e in cui si realizza un minimo isolato dell'energia potenziale è stabile secondo Lyapunov.*

La dimostrazione è elementare. Infatti è sufficiente far vedere che la funzione $\Lambda(q, v)$ è una funzione di Lyapunov per il punto di equilibrio $(0, q_e)$ (i dettagli della verifica sono lasciati per esercizio).

Quello che abbiamo visto per un'equazione scalare può essere generalizzato al caso vettoriale. Supponiamo di avere un sistema di equazioni del tipo

$$\begin{cases} \ddot{q}_1 = f_1(q_1, \dots, q_n), \\ \ddot{q}_2 = f_2(q_1, \dots, q_n), \\ \vdots \\ \ddot{q}_n = f_n(q_1, \dots, q_n), \end{cases}$$

che, scritto in forma compatta, acquista la forma

$$\ddot{\mathbf{q}} = \mathbf{f}(\mathbf{q}), \quad (2.61)$$

dove $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, e dove $\mathbf{q}(t) : \mathbb{R} \rightarrow \mathbb{R}^n$ è la funzione vettoriale incognita. Diremo che il sistema è **conservativo** se la funzione \mathbf{f} è il gradiente di una funzione scalare $-V(\mathbf{q})$, detta sempre **energia potenziale**. Ovvero

$$\mathbf{f}(\mathbf{q}) = -\nabla V(\mathbf{q}), \quad \text{cioè} \quad \begin{pmatrix} f_1(q_1, \dots, q_n) \\ f_2(q_1, \dots, q_n) \\ \vdots \\ f_n(q_1, \dots, q_n) \end{pmatrix} = \begin{pmatrix} -\frac{\partial V(q_1, \dots, q_n)}{\partial q_1} \\ -\frac{\partial V(q_1, \dots, q_n)}{\partial q_2} \\ \vdots \\ -\frac{\partial V(q_1, \dots, q_n)}{\partial q_n} \end{pmatrix}. \quad (2.62)$$

³¹Si veda p.e. L.S. Pontrjagin “Ordinary differential equations”, Reading (Mass.), Addison-Wesley, 1962.

Anche in questo caso è facile dimostrare che la funzione scalare

$$H(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} |\dot{\mathbf{q}}|^2 + V(\mathbf{q}),$$

è un integrale primo, ovvero che la sua derivata lungo la soluzioni è nulla

$$\frac{dH}{dt} = \dot{\mathbf{q}} \cdot \ddot{\mathbf{q}} + \nabla V \cdot \dot{\mathbf{q}} = \dot{\mathbf{q}} \cdot \underbrace{(\ddot{\mathbf{q}} + \nabla V)}_{\ddot{\mathbf{q}} - \mathbf{f}(\mathbf{q})=0} = 0.$$

Come conseguenza possiamo enunciare il principio di Dirichlet anche per i sistemi conservativi multidimensionali: *se il punto \mathbf{q}_e è un minimo isolato per la funzione $V(\mathbf{q})$ allora $(0, \mathbf{q}_e)$ è un punto di equilibrio stabile.* In questo caso la funzione di Lyapunov è

$$\Lambda(\mathbf{q}, \mathbf{v}) = H(\mathbf{q}, \mathbf{v}) - V(\mathbf{q}_e) = \frac{1}{2} |\mathbf{v}|^2 + V(\mathbf{q}) - V(\mathbf{q}_e).$$

Riprenderemo questo argomento della sezione 5.9 dove studieremo la stabilità dei sistemi lagrangiani.

2.9 I potenziali isocroni

Abbiamo visto nell'esempio 2.6.2 che le oscillazioni di un sistema soggetto a una energia potenziale quadratica

$$V(q) = -\frac{\omega^2}{2} q^2, \quad (2.63)$$

(oscillatore armonico) sono isocrone.

Vogliamo ora dimostrare che (2.63) è, sostanzialmente, l'unica energia potenziale che possiede questa proprietà. In altri termini dimostriamo il seguente:

Teorema 2.9.1 *Se tutti i moti che si svolgono sotto l'azione di una energia potenziale $V(q)$, simmetrica rispetto al suo minimo, sono periodici e isocroni, allora esistono $k > 0$, q_0 e U_0 tali che $V(q)$ ha la forma*

$$V(q) = \frac{k}{2} (q - q_0)^2 + U_0 \quad (2.64)$$

Dim³². La costante V_0 è introdotta solo per ragioni di completezza e possiamo supporla nulla senza perdere di generalità. Analogamente, con un'ovvia traslazione delle coordinate, possiamo anche assumere che $q_0 = 0$. In altri termini, possiamo quindi limitarci a considerare energie potenziali che soddisfino

$$V(0) = 0. \quad (2.65)$$

L'ipotesi di simmetria della $V(q)$ significa che l'energia potenziale è una funzione pari, ovvero

$$V(q) = V(-q). \quad (2.66)$$

³²La dimostrazione è ripresa, nelle sue linee generali, da G. Gallavotti, *Meccanica Elementare*, Boringheri, Torino 1980

(vedremo poi cosa cambia quando togliamo questa limitazione).

Possiamo inoltre assumere che

$$q \frac{dV}{dq}(q) > 0, \quad \forall q \neq 0, \quad (2.67)$$

che implica che l'origine $q = 0$ è un punto di minimo isolato per l'energia potenziale, e inoltre che valga

$$\lim_{|q| \rightarrow \infty} V(q) = +\infty. \quad (2.68)$$

Infatti, se così non fosse, il sistema ammetterebbe o orbite illimitate, o orbite limitate non periodiche (dimostrarlo).

Fissiamo adesso $E > 0$ e denotiamo con $q(E)$ la soluzione positiva di $V(q) = E$. Ne segue che, in virtù dell'ipotesi fatte, e in particolare della (2.66), il periodo del moto è dato da

$$T(E) = 4 \int_0^{q(E)} \frac{dq}{\sqrt{2(E - V(q))}}. \quad (2.69)$$

La funzione V è una funzione crescente di q per $q > 0$, e quindi possiamo invertirla

$$v = V(q), \quad \Leftrightarrow \quad q = Q(v). \quad (2.70)$$

Possiamo cambiare coordinate in (2.69) ponendo $s = V(q)$, e quindi, sfruttando la (2.70), $dq = Q'(s) ds$. Otteniamo

$$T(E) = 4 \int_0^E \frac{Q'(s) ds}{\sqrt{2(E - s)}}. \quad (2.71)$$

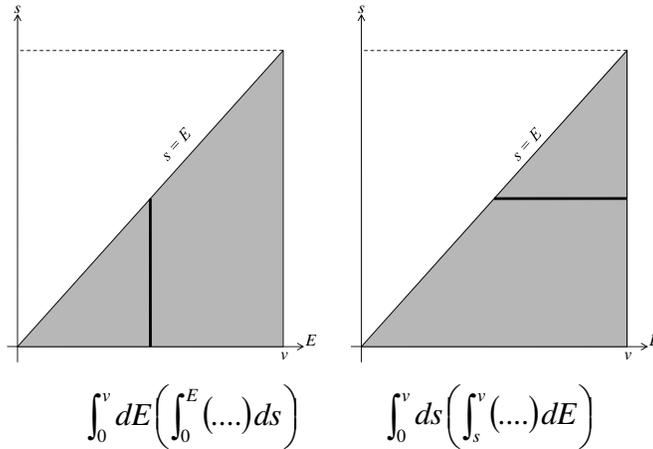
Notiamo che la funzione $T(E)$ in (2.71) è ora definita come una “trasformata” della funzione Q' . A parte questa osservazione, moltiplichiamo la (2.71) per $\frac{1}{\sqrt{v - E}}$, ed integriamo in E fra 0 e v

$$\begin{aligned} \int_0^v \frac{T(E)}{\sqrt{v - E}} dE &= \int_0^v \frac{1}{\sqrt{v - E}} \underbrace{\left(4 \int_0^E \frac{Q'(s) ds}{\sqrt{2}\sqrt{E - s}} \right)}_{T(E)} dE \\ &= \frac{4}{\sqrt{2}} \int_0^v \left[\int_0^E \frac{Q'(s)}{\sqrt{v - E}\sqrt{E - s}} ds \right] dE. \end{aligned} \quad (2.72)$$

Riferendoci alla figura 2.9, possiamo valutare l'integrale doppio che appare nella (2.72) scambiando l'ordine di integrazione: integrando cioè prima in dE fra s e v , e poi in ds fra 0 e v . Otteniamo così

$$\begin{aligned} \int_0^v \frac{T(E)}{\sqrt{v - E}} dE &= \frac{4}{\sqrt{2}} \int_0^v \left[\int_s^v \frac{Q'(s)}{\sqrt{v - E}\sqrt{E - s}} dE \right] ds \\ &= \frac{4}{\sqrt{2}} \int_0^v Q'(s) \left[\int_s^v \frac{dE}{\sqrt{v - E}\sqrt{E - s}} \right] ds. \end{aligned}$$

Figura 2.9: Integrale doppio.



Ora $\int_s^v \frac{dE}{\sqrt{v-E}\sqrt{E-s}}$, è un integrale notevole e vale π per ogni v ed s . Abbiamo quindi³³

$$\int_0^v \frac{T(E)}{\sqrt{v-E}} dE = \frac{4\pi}{\sqrt{2}} \int_0^v Q'(s) dv = \frac{4\pi}{\sqrt{2}} Q(v), \tag{2.75}$$

dal momento che $Q(0) = 0$. Sfruttando adesso l'ipotesi che $T(E) = T, \forall E$, ed il fatto che $\int_0^v \frac{dE}{\sqrt{v-E}} = 2\sqrt{v}$, abbiamo

$$T \underbrace{\int_0^v \frac{dE}{\sqrt{v-E}}}_{2\sqrt{v}} = \frac{4\pi}{\sqrt{2}} Q(v), \implies Q(v) = \frac{\sqrt{2}}{2\pi} T \sqrt{v}.$$

Quindi ricordando che $q = Q(v)$ è l'inversa di $v = V(q)$, possiamo scrivere

$$q = Q(v) = \frac{\sqrt{2}}{2\pi} T \sqrt{v}, \implies v = V(q) = \frac{2\pi^2}{T^2} q^2. \tag{2.76}$$

□

³³La (2.75) è un caso particolare dell'equazione di Abel

$$\phi(t) = \int_0^t \frac{\psi(\tau) d\tau}{\sqrt{(t-\tau)}} \tag{2.73}$$

la cui soluzione è data da

$$\psi(t) = \frac{1}{\pi} \left(\frac{\phi(0)}{t^{1/2}} + \int_0^t \frac{\phi'(\tau) d\tau}{\sqrt{(t-\tau)}} \right) \tag{2.74}$$

si veda J.R. Cannon, *The one-dimensional heat equation*, Addison-Wesley, Menlo Park, CA, 1984.

Possiamo ora vedere cosa succede se eliminiamo l'ipotesi (2.66). In questo caso $V : \mathbb{R} \rightarrow [0, +\infty)$, non è simmetrica rispetto all'origine e dunque porremo³⁴

$$V(q) = \begin{cases} V^+(q), & \text{se } q \geq 0, \\ V^-(q), & \text{se } q \leq 0, \end{cases}$$

con $V^+(0) = V^-(0) = 0$. Dato $E > 0$, indichiamo con $q^+(E)$ la soluzione di $V^+(q) = E$, mentre $q^-(E)$ denota la soluzione di $V^-(q) = E$. Ovviamente $q^+(E) > 0$, mentre $q^-(E) < 0$. In generale, $q^+(E) \neq |q^-(E)|$, dal momento che la V non è più simmetrica rispetto all'origine. Introduciamo anche le inverse di V^+ e V^- , ponendo

$$Q^+ : [0, +\infty) \rightarrow [0, +\infty), \quad v \xrightarrow{Q^+} q = Q^+(v) > 0,$$

$$Q^- : [0, +\infty) \rightarrow [0, -\infty), \quad v \xrightarrow{Q^-} q = Q^-(v) < 0.$$

Il periodo è dato da

$$\begin{aligned} T(E) &= 2 \int_{q^-(E)}^{q^+(E)} \frac{dq}{\sqrt{2(E - V(q))}} \\ &= 2 \left[\int_{q^-(E)}^0 \frac{dq}{\sqrt{2(E - V^-(q))}} + \int_0^{q^+(E)} \frac{dq}{\sqrt{2(E - V^+(q))}} \right] \\ &= 2 \left[\int_0^{q^+(E)} \frac{dq}{\sqrt{2(E - V^+(q))}} - \int_0^{|q^-(E)|} \frac{ds}{\sqrt{2(E - V^-(-s))}} \right] \end{aligned} \quad (2.77)$$

che ci consente di scrivere

$$T(E) = 2 [T^+(E) - T^-(E)],$$

dove

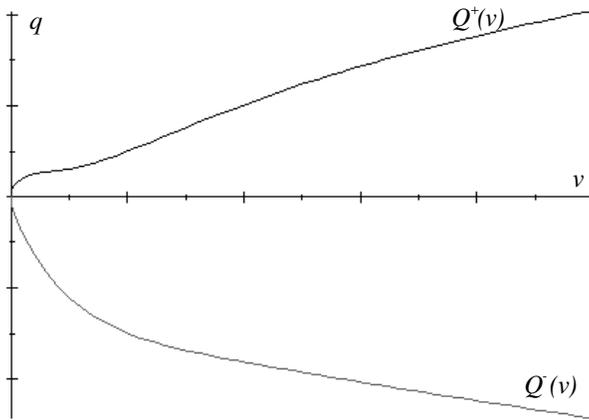
$$T^+(E) = \int_0^{q^+(E)} \frac{dq}{\sqrt{2(E - V^+(q))}}, \quad T^-(E) = \int_0^{|q^-(E)|} \frac{dq}{\sqrt{2(E - V^-(-q))}}.$$

Quindi, a parte il fattore 4, $T^+(E)$ e $T^-(E)$, altro non sono che la (2.69). Ora se imponiamo che né $T^+(E)$ né $T^-(E)$ dipendano da E , possiamo procedere esattamente come prima giungendo a

$$Q^+(v) - Q^-(v) = \frac{\sqrt{2}}{\pi} \sqrt{v}, \quad (2.78)$$

Quindi abbiamo infinite energie potenziali non simmetriche isocrone, ma ciascuna di esse è ottenuta “deformando” un'energia della forma (2.76) in modo che, $\forall v > 0$, la differenza fra $Q^+(v)$ e $Q^-(v)$ si mantenga sempre proporzionale a \sqrt{v} .

³⁴ $V^+(q)$ e $V^-(q)$ sono comunque positive e strettamente monotone.

Figura 2.10: Grafico di $Q^+(v)$ e $Q^-(v)$, dati dalla (2.79).

Esempio 2.9.1 *Se consideriamo*

$$Q^+(v) = \frac{\sqrt{2}}{2\pi} \sqrt{v} - \frac{v}{v^2+1}, \quad Q^-(v) = -\frac{\sqrt{2}}{2\pi} \sqrt{v} - \frac{v}{v^2+1}, \quad (2.79)$$

i cui grafici sono riportati nella figura 2.10, abbiamo che la (2.78) è soddisfatta, ed inoltre sia Q^+ che Q^- sono due funzioni monotone e quindi invertibili. Invertendole otteniamo

$$q = Q^+(v) = \frac{\sqrt{2}}{2\pi} \sqrt{v} - \frac{v}{v^2+1}, \quad v = V^+(q) \quad \text{definita per } q > 0,$$

$$q = Q^-(v) = -\frac{\sqrt{2}}{2\pi} \sqrt{v} - \frac{v}{v^2+1}, \quad v = V^-(q) \quad \text{definita per } q < 0,$$

da cui possiamo ricostruire $V(q)$

2.10 Sistemi lineari bidimensionali

Per lo studio dei sistemi di equazioni differenziali autonome

$$\dot{\mathbf{x}} = F(\mathbf{x})$$

è fondamentale la comprensione del caso lineare, cioè del caso in cui la funzione $F(\mathbf{x})$ è una funzione lineare.

Ci limiteremo al caso di sistemi bidimensionali, in quanto per essi è possibile scrivere delle formule generali per la soluzione ed è possibile visualizzare bene le soluzioni nel piano delle fasi.

Tuttavia il caso bidimensionale non è del tutto rappresentativo di cosa avviene per dimensioni superiori (già da $d = 3$). La principale ragione è di ordine topologico. Ricordiamo che per un sistema autonomo da ogni punto dello spazio delle fasi passa una e una sola orbita (in virtù del teorema di unicità e dell'invarianza temporale delle soluzioni). Questo implica che ogni orbita nel piano delle fasi ($d = 2$) divide il piano stesso in due regioni ("dentro" e "fuori" per un'orbita chiusa, "riva destra" e "riva sinistra" per una soluzione con orbita illimitata) che non possono essere connesse da una soluzione del sistema. Questo non è più vero già per $d = 3$.

Vediamo ora come si trova la soluzione di un sistema di equazioni differenziali della forma

$$\begin{cases} \dot{x} = ax + by, \\ \dot{y} = cx + dy, \end{cases} \quad (2.80)$$

che scriveremo in forma vettoriale

$$\dot{\mathbf{x}} = \mathbb{A}\mathbf{x}, \quad (2.81)$$

dove \mathbb{A} è la matrice

$$\mathbb{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (2.82)$$

e \mathbf{x} il vettore

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}. \quad (2.83)$$

Consideriamo adesso un **cambiamento di variabili**, dato dalla seguente trasformazione invertibile³⁵

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix}, \quad (2.84)$$

dove α, β, γ e δ sono costanti. La (2.84) può essere anche scritta nella seguente forma compatta

$$\mathbf{x}(t) = \mathbb{V}\boldsymbol{\xi}(t), \quad \Leftrightarrow \quad \boldsymbol{\xi}(t) = \mathbb{V}^{-1}\mathbf{x}(t).$$

Siccome i coefficienti della matrice \mathbb{V} sono indipendenti dal tempo

$$\dot{\mathbf{x}} = \mathbb{V}\dot{\boldsymbol{\xi}}, \quad \Leftrightarrow \quad \dot{\boldsymbol{\xi}} = \mathbb{V}^{-1}\dot{\mathbf{x}}.$$

Quindi, moltiplicando a destra per \mathbb{V}^{-1} entrambe i membri della (2.81) si ottiene

$$\underbrace{\mathbb{V}^{-1}\dot{\mathbf{x}}}_{\dot{\boldsymbol{\xi}}} = \mathbb{V}^{-1}\mathbb{A}\mathbf{x} = \mathbb{V}^{-1}\mathbb{A}\underbrace{(\mathbb{V}\mathbb{V}^{-1})}_{\mathbb{I}}\mathbf{x} = \mathbb{V}^{-1}\mathbb{A}\mathbb{V}\underbrace{(\mathbb{V}^{-1}\mathbf{x})}_{\boldsymbol{\xi}},$$

cioè

$$\dot{\boldsymbol{\xi}} = (\mathbb{V}^{-1}\mathbb{A}\mathbb{V})\boldsymbol{\xi}. \quad (2.85)$$

Se adesso scegliamo opportunamente la matrice \mathbb{V} , potremo, almeno in certi casi, fare

in modo che

$$\mathbb{V}^{-1}\mathbb{A}\mathbb{V} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

³⁵L'invertibilità della (2.84) è garantita dal fatto che $\det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \neq 0$.

Se così fosse, il sistema (2.85) diventa

$$\begin{cases} \dot{\xi} = \lambda_1 \xi, \\ \dot{\eta} = \lambda_2 \eta, \end{cases} \quad (2.86)$$

la cui integrazione è immediata³⁶

$$\begin{cases} \xi = C_1 e^{\lambda_1 t}, \\ \eta = C_2 e^{\lambda_2 t}, \end{cases} \quad (2.87)$$

dove C_1 e C_2 sono due opportune costanti di integrazione.

Ora, il fatto che $\mathbb{V}^{-1}\mathbb{A}\mathbb{V}$ possa, o non possa, esser ricondotta ad una forma diagonale dipende dalla diagonalizzabilità, o meno, della matrice \mathbb{A} . Quindi, se λ_1 e λ_2 sono i due autovalori di \mathbb{A} , e se supponiamo, per il momento, che $\det \mathbb{A} \neq 0$ (che implica³⁷ $\lambda_1 \lambda_2 \neq 0$), possiamo distinguere quattro casi:

1. la matrice \mathbb{A} ha due autovalori reali distinti λ_1 e λ_2 ;
2. la matrice \mathbb{A} ha due autovalori reali coincidenti $\lambda_1 = \lambda_2$ con molteplicità geometrica due;
3. la matrice \mathbb{A} ha due autovalori reali coincidenti $\lambda_1 = \lambda_2$ con molteplicità geometrica uno;
4. la matrice \mathbb{A} ha due autovalori complessi coniugati $\lambda_1 = \alpha + i\beta$ e $\lambda_2 = \alpha - i\beta$.

Caso 1. Siano $\mathbf{x}_{\lambda_1} = \begin{pmatrix} x_{\lambda_1} \\ y_{\lambda_1} \end{pmatrix}$, $\mathbf{x}_{\lambda_2} = \begin{pmatrix} x_{\lambda_2} \\ y_{\lambda_2} \end{pmatrix}$ i due autovettori corrispondenti a λ_1 e λ_2 . I due vettori \mathbf{x}_{λ_1} , \mathbf{x}_{λ_2} sono linearmente indipendenti³⁸ e possiamo considerarli normalizzati. Introduciamo la matrice \mathbb{V} , le cui colonne sono rispettivamente \mathbf{x}_{λ_1} e \mathbf{x}_{λ_2}

$$\mathbb{V} = [\mathbf{x}_{\lambda_1} \quad \mathbf{x}_{\lambda_2}] = \begin{pmatrix} x_{\lambda_1} & x_{\lambda_2} \\ y_{\lambda_1} & y_{\lambda_2} \end{pmatrix},$$

abbiamo $\det \mathbb{V} \neq 0$, e quindi \mathbb{V}^{-1} è ben definita. Non solo, ma è facile verificare che

$$\mathbb{A}\mathbb{V} = [\mathbb{A}\mathbf{x}_{\lambda_1} \quad \mathbb{A}\mathbf{x}_{\lambda_2}] = [\lambda_1 \mathbf{x}_{\lambda_1} \quad \lambda_2 \mathbf{x}_{\lambda_2}] = [\mathbf{x}_{\lambda_1} \quad \mathbf{x}_{\lambda_2}] \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

³⁶Si veda l'esempio 2.4.1.

³⁷Come sappiamo $\det \mathbb{A} = \lambda_1 \lambda_2$.

³⁸Se \mathbf{x}_{λ_1} , \mathbf{x}_{λ_2} non fossero linearmente indipendenti potremmo supporre $\mathbf{x}_{\lambda_1} = \beta \mathbf{x}_{\lambda_2}$. Tuttavia $\mathbb{A}\mathbf{x}_{\lambda_1} = \lambda_1 \mathbf{x}_{\lambda_1}$, $\mathbb{A}\mathbf{x}_{\lambda_2} = \lambda_2 \mathbf{x}_{\lambda_2}$, comporta $\lambda_2/\lambda_1 = 1$, che contraddice l'ipotesi $\lambda_1 \neq \lambda_2$. Infatti $\lambda_1 \mathbf{x}_{\lambda_1} = \mathbb{A}\mathbf{x}_{\lambda_1} = \beta \mathbb{A}\mathbf{x}_{\lambda_2} = \beta \lambda_2 \mathbf{x}_{\lambda_2}$, che confrontata con $\mathbf{x}_{\lambda_1} = \beta \mathbf{x}_{\lambda_2}$, implica

$$\mathbf{x}_{\lambda_1} = \underbrace{\begin{pmatrix} \beta \lambda_2 \\ \lambda_1 \end{pmatrix}}_{=\beta} \mathbf{x}_{\lambda_2}, \implies \frac{\lambda_2}{\lambda_1} = 1.$$

Di conseguenza \mathbf{x}_{λ_1} , \mathbf{x}_{λ_2} sono linearmente indipendenti.

cioè

$$\mathbb{A}\mathbb{V} = \mathbb{V} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad (2.88)$$

e quindi

$$\mathbb{V}^{-1}\mathbb{A}\mathbb{V} \stackrel{(2.88)}{=} \underbrace{\mathbb{V}^{-1}\mathbb{V}}_{\mathbb{I}} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (2.89)$$

Dunque le nuove variabili dipendenti $\xi(t)$ e $\eta(t)$ soddisfano il sistema (2.86), la cui soluzione è (2.87). In pratica il sistema (2.80) è stato disaccoppiato, in quanto l'equazione per $\dot{\xi}$ dipende solo da ξ , e quella per $\dot{\eta}$ solo da η . Per ritornare alle variabili $x(t)$ e $y(t)$ usiamo la (2.84), ottenendo

$$\underbrace{\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}}_{\mathbf{x}(t)} = \underbrace{\begin{bmatrix} \mathbf{x}_{\lambda_1} & \mathbf{x}_{\lambda_2} \end{bmatrix}}_{\mathbb{V}} \underbrace{\begin{pmatrix} C_1 e^{\lambda_1 t} \\ C_2 e^{\lambda_2 t} \end{pmatrix}}_{\boldsymbol{\xi}(t)},$$

che, scritta in forma più compatta, diventa

$$\mathbf{x}(t) = \mathbf{x}_{\lambda_1} C_1 e^{\lambda_1 t} + \mathbf{x}_{\lambda_2} C_2 e^{\lambda_2 t}. \quad (2.90)$$

E' facile verificare che (2.90) fornisce, al variare delle costanti C_1 e C_2 , **tutte** le soluzioni dell'equazione (2.81). Per dimostrare ciò bisogna infatti far vedere che, per ogni condizione iniziale \mathbf{x}_0 , è possibile determinare le costanti C_1 e C_2 in modo che

$$C_1 \mathbf{x}_{\lambda_1} + C_2 \mathbf{x}_{\lambda_2} = \mathbf{x}_0. \quad (2.91)$$

Ma l'equazione (2.91) nelle incognite C_1 e C_2 si scrive anche così

$$\mathbb{V} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \mathbf{x}_0,$$

ed ha sempre una e una sola soluzione dal momento che $\det \mathbb{V} \neq 0$.

Caso 2. Adesso consideriamo il caso $\lambda_1 = \lambda_2 = \lambda$, a cui corrispondono due autovettori indipendenti \mathbf{x}_1 e \mathbf{x}_2 (che costituiscono la base dell'autospazio relativo a λ). Consideriamo adesso il generico vettore $\mathbf{v} = \alpha \mathbf{x}_1 + \beta \mathbf{x}_2$,

$$\mathbb{A}\mathbf{v} = \mathbb{A}(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2) = \lambda(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2).$$

Ovvero, $\forall \mathbf{v} \in \mathbb{R}^2$ abbiamo $\mathbb{A}\mathbf{v} = \lambda \mathbf{v}$. Pertanto $\mathbb{A} = \lambda \mathbb{I}$, la matrice \mathbb{A} è un multiplo dell'identità. Il sistema (2.80) di fatto si scrive così

$$\begin{cases} \dot{x} = \lambda x, \\ \dot{y} = \lambda y, \end{cases} \implies \begin{cases} x = x_0 \exp\{\lambda t\}, \\ y = y_0 \exp\{\lambda t\}. \end{cases}$$

Caso 3. Sia $\mathbf{x}_\lambda = \begin{pmatrix} x_\lambda \\ y_\lambda \end{pmatrix}$ l'unico (a meno di multipli reali) autovettore corrispondente all'autovalore $\lambda = \lambda_1 = \lambda_2$. Supponiamo $|\mathbf{x}_\lambda| = 1$, e indichiamo con $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ un vettore, di modulo unitario, cioè $|\mathbf{y}| = 1$, ortogonale a \mathbf{x}_λ ,

$\mathbf{x}_\lambda \cdot \mathbf{y} = 0$. Quindi la coppia $\{\mathbf{x}_\lambda, \mathbf{y}\}$ costituiscono una base ortonormale di \mathbb{R}^2 . Al solito si introduce la matrice

$$\mathbb{V} = \begin{bmatrix} \mathbf{x}_\lambda & \mathbf{y} \end{bmatrix}, \quad (2.92)$$

il cui determinante è ovviamente non nullo perché $(\mathbf{x}_\lambda$ e \mathbf{y} sono linearmente indipendenti). Introducendo

$$\mathbf{x}_\lambda^T = (x_\lambda \quad y_\lambda), \quad \mathbf{y}^T = (y_1 \quad y_2),$$

è facile vedere che

$$\mathbb{V}^{-1} = \begin{bmatrix} \mathbf{x}_\lambda^T \\ \mathbf{y}^T \end{bmatrix} = \begin{pmatrix} x_\lambda & y_\lambda \\ y_1 & y_2 \end{pmatrix}.$$

Infatti, eseguendo il prodotto righe per colonne, si ha

$$\mathbb{V}^{-1}\mathbb{V} = \begin{bmatrix} \mathbf{x}_\lambda^T \\ \mathbf{y}^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_\lambda & \mathbf{y} \end{bmatrix} = \begin{pmatrix} \mathbf{x}_\lambda \cdot \mathbf{x}_\lambda & \mathbf{x}_\lambda \cdot \mathbf{y} \\ \mathbf{x}_\lambda \cdot \mathbf{y} & \mathbf{y} \cdot \mathbf{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Si passa quindi alle variabili $\xi(t), \eta(t)$ introducendo la solita trasformazione (2.84). Il sistema nelle nuove variabili è dato da

$$\dot{\boldsymbol{\xi}} = \mathbb{V}^{-1}\dot{\mathbf{x}} = \mathbb{V}^{-1}\mathbb{A}\mathbb{V}\boldsymbol{\xi}, \quad (2.93)$$

dove adesso dobbiamo calcolare $\mathbb{V}^{-1}\mathbb{A}\mathbb{V}$. Cominciamo con

$$\mathbb{A}\mathbb{V} = \mathbb{A} \begin{bmatrix} \mathbf{x}_\lambda & \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbb{A}\mathbf{x}_\lambda & \mathbb{A}\mathbf{y} \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x}_\lambda & \mathbb{A}\mathbf{y} \end{bmatrix}.$$

Per quanto riguarda $\mathbb{A}\mathbf{y}$, possiamo esprimerlo sulla base $\{\mathbf{x}_\lambda, \mathbf{y}\}$, scrivendo³⁹

$$\mathbb{A}\mathbf{y} = \alpha\mathbf{x}_\lambda + \beta\mathbf{y},$$

dove α e β sono numeri reali. Quindi $\mathbb{A}\mathbb{V} = \begin{bmatrix} \lambda\mathbf{x}_\lambda & \alpha\mathbf{x}_\lambda + \beta\mathbf{y} \end{bmatrix}$, e di conseguenza

$$\begin{aligned} \mathbb{V}^{-1}\mathbb{A}\mathbb{V} &= \begin{bmatrix} \mathbf{x}_\lambda^T \\ \mathbf{y}^T \end{bmatrix} \begin{bmatrix} \lambda\mathbf{x}_\lambda & \alpha\mathbf{x}_\lambda + \beta\mathbf{y} \end{bmatrix} = \begin{pmatrix} \lambda\mathbf{x}_\lambda \cdot \mathbf{x}_\lambda & \mathbf{x}_\lambda \cdot (\alpha\mathbf{x}_\lambda + \beta\mathbf{y}) \\ \mathbf{x}_\lambda \cdot \mathbf{y} & \mathbf{y} \cdot (\alpha\mathbf{x}_\lambda + \beta\mathbf{y}) \end{pmatrix} \\ &= \begin{pmatrix} \lambda & \alpha \\ 0 & \beta \end{pmatrix}. \end{aligned}$$

Ora i due autovalori di \mathbb{A} sono coincidenti, e quindi $\det \mathbb{A} = \lambda^2$. D'altra parte

$$\det(\mathbb{V}^{-1}\mathbb{A}\mathbb{V}) = \det \mathbb{A} = \lambda^2,$$

ma anche

$$\det(\mathbb{V}^{-1}\mathbb{A}\mathbb{V}) = \det \begin{pmatrix} \lambda & \alpha \\ 0 & \beta \end{pmatrix} = \lambda\beta.$$

³⁹Notiamo che se fosse $\alpha = 0$, allora \mathbf{y} sarebbe anch'esso un autovettore, linearmente indipendente con \mathbf{x}_λ , contro l'ipotesi di molteplicità geometrica uno.

Per cui, confrontando, abbiamo $\beta = \lambda$, ed allora⁴⁰

$$\mathbb{V}^{-1}\mathbb{A}\mathbb{V} = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}.$$

Possiamo allora riscrivere il sistema (2.93) nelle nuove variabili $\xi(t), \eta(t)$

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}}_{\mathbb{V}^{-1}\mathbb{A}\mathbb{V}} \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \implies \begin{cases} \dot{\xi} = \lambda\xi + \alpha\eta, \\ \dot{\eta} = \lambda\eta, \end{cases}$$

e risolverlo nella $\eta(t)$, cioè $\eta = C_2 e^{\lambda t}$. L'equazione per la ξ

$$\dot{\xi} = \lambda\xi + \alpha \underbrace{C_2 e^{\lambda t}}_{\eta(t)}, \quad (2.94)$$

rientra nella classe di quelle analizzate nell'esempio 2.4.2, la cui soluzione è⁴¹

$$\xi(t) = e^{\lambda t} (C_1 + C_2 \alpha t).$$

Tornando alle variabili dipendenti originarie, cioè $x(t)$ e $y(t)$, abbiamo

$$\begin{aligned} \underbrace{\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}}_{\mathbf{x}(t)} &= \underbrace{\begin{bmatrix} \mathbf{x}_\lambda & \mathbf{y} \end{bmatrix}}_{\mathbb{V}} \underbrace{\begin{pmatrix} e^{\lambda t} (C_2 \alpha t + C_1) \\ C_2 e^{\lambda t} \end{pmatrix}}_{\boldsymbol{\xi}(t)} \\ &= e^{\lambda t} (C_2 \alpha t + C_1) \mathbf{x}_\lambda + C_2 e^{\lambda t} \mathbf{y}. \end{aligned} \quad (2.95)$$

Caso 4. Analizziamo adesso il caso degli autovalori complessi coniugati. Siano quindi $\alpha + i\beta$ e $\alpha - i\beta$, i due autovalori di \mathbb{A} . Poiché sono due autovalori distinti ($\beta \neq 0$ ovviamente), ad essi corrisponderanno due autovettori linearmente indipendenti *una volta che si consideri la matrice \mathbb{A} come matrice complessa 2×2* . Indichiamo con \mathbf{z} l'autovettore associato a $\alpha + i\beta$: è immediato verificare se $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, con

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

vettori reali, allora il vettore $\bar{\mathbf{z}} = \mathbf{x} - i\mathbf{y}$ è un autovettore associato all'autovettore $\alpha - i\beta$. Inoltre, analogamente a quanto succede nel caso reale, i due vettori \mathbf{z} e $\bar{\mathbf{z}}$ sono linearmente indipendenti (sul campo complesso). Questo implica che i vettori reali \mathbf{x} e \mathbf{y} sono, al loro volta, linearmente indipendenti sul campo reale. Infatti, se così non

⁴⁰Questa forma della matrice $\mathbb{V}^{-1}\mathbb{A}\mathbb{V}$ viene detta *forma canonica di Jordan*.

⁴¹Anche senza "scomodare" l'esempio 2.4.2, si osserva che moltiplicando la (2.94) per $e^{-\lambda t}$, si ottiene $\dot{\xi}e^{-\lambda t} - \lambda e^{-\lambda t}\xi = \alpha C_2$, ovvero $\frac{d(\xi e^{-\lambda t})}{dt} = \alpha C_2$, la cui soluzione si ricava con una semplice integrazione

$$\xi e^{-\lambda t} = C_1 + \alpha C_2 t, \implies \xi(t) = e^{\lambda t} (C_1 + \alpha C_2 t).$$

fosse avremmo $\mathbf{y} = \gamma \mathbf{x}$ (si può ovviamente assumere che $\mathbf{x} \neq 0$) da cui si ricava $\mathbf{z} = \mathbf{x} + i\mathbf{y} = (1 + i\gamma)\mathbf{x}$, e $\bar{\mathbf{z}} = \mathbf{x} - i\mathbf{y} = (1 - i\gamma)\mathbf{x}$ e quindi $\bar{\mathbf{z}} = \frac{(1 - i\gamma)}{(1 + i\gamma)}\mathbf{z}$, in contrasto con l'indipendenza lineare (sui complessi) dei due autovettori di \mathbb{A} .

Calcolando $\mathbb{A}\mathbf{z}$ si ottiene

$$\mathbb{A}\mathbf{z} = \mathbb{A}\mathbf{x} + i\mathbb{A}\mathbf{y} = (\alpha + i\beta)(\mathbf{x} + i\mathbf{y}) = (\alpha\mathbf{x} - \beta\mathbf{y}) + i(\alpha\mathbf{y} + \beta\mathbf{x}),$$

da cui otteniamo

$$\begin{cases} \mathbb{A}\mathbf{x} = \alpha\mathbf{x} - \beta\mathbf{y}, \\ \mathbb{A}\mathbf{y} = \alpha\mathbf{y} + \beta\mathbf{x}. \end{cases} \quad (2.96)$$

Introduciamo le nuove variabili dipendenti $\xi(t)$ e $\eta(t)$, tramite la trasformazione (2.84), dove la matrice \mathbb{V} è data da⁴²

$$\mathbb{V} = \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix}.$$

Riscrivendo il sistema (2.81) nelle nuove variabili otteniamo al solito la (2.93), dove, come nel caso precedente, dobbiamo valutare $\mathbb{V}^{-1}\mathbb{A}\mathbb{V}$. Svolgendo i calcoli

$$\begin{aligned} \mathbb{A}\mathbb{V} &= \mathbb{A} \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbb{A}\mathbf{x} & \mathbb{A}\mathbf{y} \end{bmatrix} \stackrel{(2.96)}{=} \begin{bmatrix} \alpha\mathbf{x} - \beta\mathbf{y} & \alpha\mathbf{y} + \beta\mathbf{x} \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix}}_{\mathbb{V}} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \end{aligned}$$

e dunque

$$\mathbb{V}^{-1}\mathbb{A}\mathbb{V} = \mathbb{V}^{-1}\mathbb{V} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Di conseguenza, nelle nuove variabili ξ ed η , il sistema di equazioni differenziali (2.81) diventa

$$\begin{cases} \dot{\xi} = \alpha\xi + \beta\eta, \\ \dot{\eta} = -\beta\xi + \alpha\eta, \end{cases} \quad \text{ovvero} \quad \begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (2.97)$$

La soluzione di questo sistema diventa agevole passando alle coordinate polari, ovvero al sistema di variabili (ρ, θ) tale che $\xi = \rho \cos \theta$, $\eta = \rho \sin \theta$. Infatti, moltiplicando la prima equazione per ξ , la seconda per η e sommando otteniamo

$$\frac{d}{dt}(\xi^2 + \eta^2) = 2\alpha(\xi^2 + \eta^2), \quad (2.98)$$

ovvero

$$\frac{d\rho^2}{dt} = 2\alpha\rho^2, \quad (2.99)$$

che si integra immediatamente, ottenendo

$$\rho(t) = \rho_0 e^{\alpha t} = \rho_0 \exp \{(\operatorname{Re} \lambda) t\}. \quad (2.100)$$

⁴²Evidentemente \mathbb{V} è invertibile dal momento che \mathbf{x} , \mathbf{y} sono indipendenti.

L'equazione per θ si ottiene invece moltiplicando la (2.97)₁ per η , la (2.97)₂ per ξ e sottraendo. Si ottiene in questo modo

$$\dot{\xi}\eta - \dot{\eta}\xi = \beta \underbrace{(\xi^2 + \eta^2)}_{\rho^2}, \quad (2.101)$$

che possiamo riscrivere come

$$\frac{\dot{\xi}\eta - \dot{\eta}\xi}{\rho^2} = \beta. \quad (2.102)$$

E' facile verificare che il primo membro della (2.102) è la derivata, cambiata di segno, di $\theta(t)$

$$\frac{d\theta}{dt} = -\beta, \quad (2.103)$$

da cui otteniamo

$$\theta(t) = -\beta t + \theta_0 = -(\operatorname{Im} \lambda) t + \theta_0. \quad (2.104)$$

Tornando ora alle variabili $\xi(t)$ e $\eta(t)$, scriveremo la soluzione come

$$\begin{cases} \xi(t) = \rho_0 e^{\alpha t} \cos(-\beta t + \theta_0), \\ \eta(t) = \rho_0 e^{\alpha t} \sin(-\beta t + \theta_0), \end{cases} \quad (2.105)$$

dove i coefficienti ρ_0 e θ_0 sono legati alle condizioni iniziali da

$$\rho_0 = \sqrt{\xi_0^2 + \eta_0^2}, \quad \theta_0 = \operatorname{arctg} \frac{\eta_0}{\xi_0}. \quad (2.106)$$

Infine, per ritornare alle variabili originarie $x(t)$ e $y(t)$, si applica la (2.84) ottenendo

$$\begin{aligned} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \begin{bmatrix} \mathbf{x} & \mathbf{y} \end{bmatrix} \begin{pmatrix} \rho_0 e^{\alpha t} \cos(-\beta t + \theta_0) \\ \rho_0 e^{\alpha t} \sin(-\beta t + \theta_0) \end{pmatrix} \\ &= \rho_0 e^{\alpha t} [\cos(-\beta t + \theta_0) \mathbf{x} + \sin(-\beta t + \theta_0) \mathbf{y}]. \end{aligned} \quad (2.107)$$

Vediamo adesso cosa succede nel caso in cui la matrice \mathbb{A} abbia determinante nullo. Il fatto di avere determinante nullo implica che almeno un autovalore della matrice, diciamo λ_1 , sia nullo. Si hanno quindi due casi:

(A). $\lambda_1 = 0, \lambda_2 \neq 0$;

(B). $\lambda_1 = 0$, è autovalore doppio con molteplicità geometrica pari ad 1.

Caso (A). In questo caso possiamo pensare al kernel di \mathbb{A} come al sottospazio associato all'autovettore nullo. Sia quindi \mathbf{x}_{λ_1} il relativo autovettore e \mathbf{x}_{λ_2} , quello relativo a λ_2 . Introduciamo la trasformazione (2.84) e, operando come nel **Caso 1** giungiamo alla (2.86) che adesso si scrive così

$$\begin{cases} \dot{\xi} = 0, \\ \dot{\eta} = \lambda_2 \eta. \end{cases} \quad \Longrightarrow \quad \begin{cases} \xi(t) = \xi_0, \\ \eta(t) = \eta_0 e^{\lambda_2 t}. \end{cases}$$

Ricordando poi la (2.90), la soluzione è

$$\mathbf{x}(t) = \mathbf{x}_{\lambda_1} \xi_0 + \mathbf{x}_{\lambda_2} \eta_0 e^{\lambda_2 t}.$$

Caso (B). Se $\lambda = 0$ è un autovettore doppio, abbiamo ancora due possibilità. La prima, banale, è che \mathbb{A} sia la matrice nulla. Ovviamente la soluzione del sistema in questo caso è $x(t) \equiv x_0$ e $y(t) \equiv y_0$. L'altra possibilità è che la molteplicità geometrica di $\lambda = 0$ sia uno. Questo è ancora una forma particolare del **Caso 3** che abbiamo già visto. Infatti se \mathbf{x}_λ è l'autovettore unitario relativo a $\lambda = 0$, e \mathbf{y} è un vettore unitario ortogonale ad \mathbf{x}_λ , operando la solita trasformazione (2.84) con \mathbb{V} data da (2.92), giungiamo a questo sistema nelle nuove variabili $\xi(t)$ e $\eta(t)$

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \quad \Rightarrow \quad \begin{cases} \dot{\xi} = \alpha \eta, \\ \dot{\eta} = 0, \end{cases}$$

la cui soluzione è

$$\begin{cases} \xi(t) = \xi_0 + \alpha \eta_0 t, \\ \eta(t) = \eta_0. \end{cases}$$

Applicando infine la (2.95) con $\lambda = 0$, abbiamo la soluzione del sistema (2.90)

$$\mathbf{x}(t) = (\eta_0 \alpha t + \xi_0) \mathbf{x}_\lambda + \eta_0 \mathbf{y}.$$

Esempio 2.10.1 Determiniamo la soluzione generale di

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

In questo caso $\mathbb{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, ed i suoi autovalori sono

$$\det(\mathbb{A} - \lambda \mathbb{I}) = 0, \quad \Rightarrow \quad (1 - \lambda)^2 - 4 = 0, \quad \lambda = \begin{cases} -1, \\ 3. \end{cases}$$

Si determinano poi i rispettivi autovettori $\mathbb{A}\mathbf{x}_1 = -\mathbf{x}_1$, $\mathbb{A}\mathbf{x}_2 = 3\mathbf{x}_2$, richiedendo $|\mathbf{x}_1| = |\mathbf{x}_2| = 1$. Si ha

$$\mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Si introduce quindi la matrice

$$\mathbb{V} = [\mathbf{x}_1 \quad \mathbf{x}_2] = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix},$$

e si considera il cambio di variabili $\boldsymbol{\xi} = \mathbb{V}^{-1}\mathbf{x}$, $\mathbf{x} = \mathbb{V}\boldsymbol{\xi}$. Nelle nuove variabili $\xi(t)$, $\eta(t)$ l'equazione differenziale $\dot{\mathbf{x}} = \mathbb{A}\mathbf{x}$, si riscrive come

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \quad \Rightarrow \quad \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = \begin{pmatrix} C_1 e^{-t} \\ C_2 e^{3t} \end{pmatrix}.$$

Quindi, in termini di $\mathbf{x}(t)$, la soluzione generale del sistema si può riscrivere come

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \mathbb{V}\boldsymbol{\xi}(t) = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -C_1e^{-t} + C_2e^{3t} \\ C_1e^{-t} + C_2e^{3t} \end{pmatrix},$$

ovvero $\mathbf{x}(t) = C_1e^{-t}\mathbf{x}_1 + C_2e^{3t}\mathbf{x}_2$. Le costanti c_1 e c_2 si determinano in base alle condizioni iniziali.

Esempio 2.10.2 Scriviamo la soluzione generale di

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 5 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Adesso $\mathbb{A} = \frac{1}{2} \begin{pmatrix} 5 & 1 \\ -1 & 3 \end{pmatrix}$, e $\det(\mathbb{A} - \lambda\mathbb{I}) = 0$, dà $\lambda_1 = \lambda_2 = 2$. Abbiamo quindi un solo autovettore $\mathbf{x} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Si considera un vettore unitario \mathbf{y} ortogonale ad \mathbf{x} ,

$$\mathbf{y} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Si introduce la matrice $\mathbb{V} = [\mathbf{x} \quad \mathbf{y}] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, la cui inversa è

$$\mathbb{V}^{-1} = \mathbb{V}^T = \begin{bmatrix} \mathbf{x}^T \\ \mathbf{y}^T \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Al solito si introducono le nuove variabili dipendenti $\boldsymbol{\xi}(t) = \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix}$, come combinazione lineare delle $x(t)$ e $y(t)$, $\mathbf{x}(t) = \mathbb{V}\boldsymbol{\xi}(t)$, $\boldsymbol{\xi}(t) = \mathbb{V}^T\mathbf{x}(t)$. Nelle variabili $\boldsymbol{\xi}(t)$, il sistema si riscrive come

$$\dot{\boldsymbol{\xi}} = \mathbb{V}^T\mathbb{A}\mathbb{V}\boldsymbol{\xi} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix}.$$

Quindi, abbiamo questo sistema

$$\begin{cases} \dot{\xi}(t) = 2\xi(t) + \eta(t), \\ \dot{\eta}(t) = 2\eta(t), \end{cases} \quad \Rightarrow \quad \begin{cases} \dot{\xi}(t) = 2\xi(t) + C_2e^{2t}, \\ \eta(t) = C_2e^{2t}. \end{cases}$$

Moltiplicando la prima equazione per e^{-2t} , possiamo riscriverla come

$$\frac{d}{dt}(\xi e^{-2t}) = C_2, \quad \Rightarrow \quad \xi(t) = (C_1 + C_2t) e^{2t}.$$

Abbiamo quindi $\mathbf{x}(t) = (C_1 + C_2t) e^{2t}\mathbf{x} + C_2e^{2t}\mathbf{y}$.

Esempio 2.10.3 Analizziamo la soluzione di

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (2.108)$$

Gli autovalori di $\mathbb{A} = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$, sono $\lambda_1 = 2 + i$, $\lambda_2 = 2 - i$, cioè $\operatorname{Re} \lambda = 2$, e $\operatorname{Im} \lambda = 1$. Determinando un autovettore complesso, per esempio quello relativo a λ_1 , $\mathbb{A} = (2 + i) \mathbf{z}$, si ottiene

$$\mathbf{z} = \begin{pmatrix} 1 \\ i \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{\mathbf{x}} + i \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\mathbf{y}}.$$

Si introduce la matrice $\mathbb{V} = [\mathbf{x} \ \mathbf{y}] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, che quindi coincide con l'identità, cioè $\boldsymbol{\xi}(t) \equiv \mathbf{x}(t)$. Passando direttamente alle coordinate polari

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \rho(t) \cos \theta(t) \\ \rho(t) \sin \theta(t) \end{pmatrix},$$

otteniamo

$$\begin{cases} \rho(t) = \rho_o e^{(\operatorname{Re} \lambda)t} = \rho_o e^{2t}, \\ \theta(t) = -(\operatorname{Im} \lambda)t + \theta_o = -t + \theta_o. \end{cases}$$

da cui

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \rho_o e^{2t} \cos(-t + \theta_o) \\ \rho_o e^{2t} \sin(-t + \theta_o) \end{pmatrix}.$$

Esempio 2.10.4 Analizziamo adesso, facendo uso del formalismo dei sistemi lineari bidimensionali, l'equazione dell'oscillatore armonico $\ddot{x} = -\omega^2 x$. Considerando le solite variabili $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, l'equazione può essere riscritta come

$$\underbrace{\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}}_{\mathbb{A}} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\mathbf{x}}.$$

Gli autovalori di \mathbb{A} , sono $\lambda_1 = -i\omega$, $\lambda_2 = i\omega$. In particolare, un autovettore relativo a $-i\omega$ è

$$\mathbf{z} = \begin{pmatrix} 1 \\ -i\omega \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ -\omega \end{pmatrix},$$

per cui la matrice \mathbb{V} è data da

$$\mathbb{V} = \begin{pmatrix} 1 & 0 \\ 0 & -\omega \end{pmatrix}.$$

Introducendo le nuove variabili dipendenti $\xi(t)$ e $\eta(t)$, e ricordando la (2.105), abbiamo

$$\begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = \begin{pmatrix} \rho_o \cos(\omega t + \theta_o) \\ \rho_o \sin(\omega t + \theta_o) \end{pmatrix}.$$

Per cui, tornando alle variabili originali

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = \begin{pmatrix} \xi(t) \\ -\omega \eta(t) \end{pmatrix},$$

ossia

$$\begin{cases} x(t) = \rho_o \cos(\omega t + \theta_o), \\ y(t) = -\omega \rho_o \sin(\omega t + \theta_o) = \dot{x}(t), \end{cases}$$

che è proprio la (2.38).

2.11 Moto armonico smorzato

Consideriamo l'oscillatore armonico dell'esempio 2.6.1 soggetto ad una forza di attrito viscoso, cioè proporzionale, in prima approssimazione, alla velocità del punto materiale (si veda anche l'esempio 2.5.1). Denotando con $x(t)$ l'ascissa del punto materiale, l'equazione di moto è

$$m\ddot{x} = \underbrace{(-kx)}_{\text{molla}} + \underbrace{(-\eta\dot{x})}_{\text{attrito}},$$

dove η è il coefficiente di attrito viscoso. Seguendo le notazioni degli esempi 2.6.1 e 2.5.1, si introduce $\omega^2 = \frac{K}{m}$, $\mu = \frac{\eta}{m}$, e l'equazione si riscrive come

$$\ddot{x} + \mu\dot{x} + \omega^2x = 0, \quad (2.109)$$

che poi si trasforma nel seguente sistema lineare

$$\begin{cases} \dot{x} = y \\ \dot{y} = -\mu y - \omega^2 x \end{cases} \Rightarrow \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega^2 & -\mu \end{pmatrix}}_{\mathbb{A}} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (2.110)$$

Si determinano gli autovalori di \mathbb{A} , $\lambda^2 + \mu\lambda + \omega^2 = 0$, ovvero

$$\lambda = \frac{1}{2} \left[-\mu \pm \sqrt{\mu^2 - 4\omega^2} \right].$$

Si hanno quindi tre casi.

Caso I. $\mu^2 - 4\omega^2 > 0$ (smorzamento forte). I due autovalori, entrambi negativi, sono

$$\lambda_1 = \frac{1}{2} \left[-\mu - \sqrt{\mu^2 - 4\omega^2} \right], \quad \lambda_2 = \frac{1}{2} \left[-\mu + \sqrt{\mu^2 - 4\omega^2} \right].$$

Ricordando la (2.90), $x(t)$ è di questo tipo:

$$\begin{aligned} x(t) &= \underbrace{(x_{\lambda_1} C_1)}_{c_1} e^{-\frac{\mu}{2}t - \frac{1}{2}\sqrt{\mu^2 - 4\omega^2}t} + \underbrace{(x_{\lambda_2} C_2)}_{c_2} e^{-\frac{\mu}{2}t + \frac{1}{2}\sqrt{\mu^2 - 4\omega^2}t} \\ &= e^{-\frac{\mu}{2}t} \left[c_1 e^{-\frac{1}{2}\sqrt{\mu^2 - 4\omega^2}t} + c_2 e^{\frac{1}{2}\sqrt{\mu^2 - 4\omega^2}t} \right], \end{aligned}$$

dove con c_1 e c_2 sono le costanti di integrazione da determinarsi in base alle condizioni iniziali. Notiamo che i due autovalori sono negativi e quindi abbiamo due esponenziali decrescenti: $x(t) \xrightarrow[t \rightarrow \infty]{} 0$.

Caso II. $\mu^2 - 4\omega^2 = 0$ (smorzamento critico). Questo caso corrisponde al **Caso 3** della sezione 2.10: un solo autovalore con molteplicità geometrica pari a uno. La soluzione è data dalla (2.95), per cui $x(t)$ è di questo tipo

$$\begin{aligned} x(t) &= e^{-\frac{\mu}{2}t} (C_2\alpha t + C_1) x_\lambda + C_2 e^{-\frac{\mu}{2}t} y_1 \\ &= e^{-\frac{\mu}{2}t} \underbrace{(C_1 x_\lambda + C_2 y_1)}_{c_1} + e^{-\frac{\mu}{2}t} \underbrace{(C_2 \alpha x_\lambda)}_{c_2} t \\ &= e^{-\frac{\mu}{2}t} [c_1 + c_2 t], \end{aligned}$$

dove al solito c_1 e c_2 sono le costanti di integrazione. Al solito, $\lim_{t \rightarrow \infty} x(t) = 0$.

Caso III. $\mu^2 - 4\omega^2 < 0$ (oscillazioni smorzate). Gli autovalori sono complessi

$$\lambda_1 = \frac{1}{2} \left[-\mu - i\sqrt{4\omega^2 - \mu^2} \right], \quad \lambda_2 = \frac{1}{2} \left[-\mu + i\sqrt{4\omega^2 - \mu^2} \right].$$

Quindi, rifacendosi al **Caso 4** della sezione 2.10, abbiamo

$$\alpha = \operatorname{Re} \lambda_1 = -\frac{\mu}{2}, \quad \beta = \operatorname{Im} \lambda_1 = -\frac{1}{2} \sqrt{4\omega^2 - \mu^2}.$$

La forma generale della soluzione è data dalla (2.107). Di conseguenza nel nostro caso abbiamo

$$x(t) = \rho_0 e^{-\frac{\mu}{2}t} \left[x_1 \cos \left(\frac{1}{2} \sqrt{4\omega^2 - \mu^2} t + \theta_0 \right) + y_1 \sin \left(\frac{1}{2} \sqrt{4\omega^2 - \mu^2} t + \theta_0 \right) \right],$$

che, definendo opportunamente le costanti c_1 e c_2 , potremo scrivere come

$$x(t) = e^{-\frac{\mu}{2}t} \left[c_1 \cos \left(\frac{1}{2} \sqrt{4\omega^2 - \mu^2} t \right) + c_2 \sin \left(\frac{1}{2} \sqrt{4\omega^2 - \mu^2} t \right) \right].$$

Notiamo ancora la presenza dell'esponenziale negativo che quindi “smorza” l'ampiezza delle oscillazioni. Di solito ci si riferisce a questo caso come **caso periodico** in quanto la soluzione ha un comportamento periodico, ma di ampiezza smorzata dal fattore esponenziale. Per contrapposizione, ci si riferisce al primo e al secondo caso come **casi aperiodici**.

Per concludere, rimarchiamo alcune proprietà generali facilmente deducibili dalla forma esplicita della soluzione:

Primo. La soluzione generale della (2.109) è sempre data dalla combinazione lineare di due soluzioni “fondamentali”, ovvero di due qualsiasi soluzioni *linearmente indipendenti*⁴³ tra loro.

Secondo. Per t “sufficientemente grande” la soluzione si annulla, e di conseguenza anche $\lim_{t \rightarrow \infty} \dot{x} = 0$: il punto materiale tende alla quiete ed alla posizione $x = 0$, che rappresenta, nel senso della (2.56), la configurazione di equilibrio del sistema (2.110).

⁴³Due, o più, funzioni si dicono linearmente indipendenti quando l'unica loro combinazione lineare a coefficienti costanti *identicamente nulla* è quella con tutti i coefficienti nulli.

Terzo. Se calcoliamo l'energia meccanica E tramite la formula (2.31), otteniamo

$$E = \frac{\dot{x}^2}{2} + V(x), \quad \text{con} \quad V(x) = \frac{\omega^2}{2}x^2.$$

Derivando quindi E rispetto al tempo si ha

$$\frac{dE}{dt} = \dot{x} \ddot{x} + \omega^2 \dot{x}x = \dot{x} (\ddot{x} + \omega^2 x) \stackrel{(2.109)}{=} -\mu \dot{x}^2 \leq 0,$$

l'energia meccanica non si conserva più. In particolare, siccome $\lim_{t \rightarrow \infty} \dot{x} = 0$, anche l'energia meccanica si annullerà per $t \rightarrow \infty$. Concludiamo quindi che l'oscillatore armonico smorzato non è un sistema conservativo ma *dissipativo*. Ovviamente l'energia meccanica non viene "distrutta", ma viene trasformata dall'attrito in calore (cioè in energia termica).

Quarto. La configurazione di equilibrio $(x, y) \equiv (0, 0)$ del sistema (2.110) è una *configurazione di equilibrio asintoticamente stabile* nel senso del criterio di Lyapunov illustrato nella sezione 2.8.2. Infatti se consideriamo

$$\begin{aligned} \Lambda(x, y) &= \frac{y^2}{2} + \frac{\omega^2}{2}x^2 + \frac{\mu^2}{4}x^2 + \frac{\mu}{2}xy \\ &= \frac{\omega^2}{2}x^2 + \frac{1}{4}(\mu x + y)^2 + \frac{y^2}{4}, \end{aligned}$$

è banale verificare che $\Lambda > 0$, al di fuori di $x = 0$ e $y = 0$. Inoltre, siccome

$$\begin{aligned} \nabla \Lambda \cdot \mathbf{F}(x) &= \begin{pmatrix} \left(\omega^2 + \frac{\mu^2}{2} \right) x + \frac{\mu}{2} y \\ y + \frac{\mu}{2} x \end{pmatrix} \cdot \begin{pmatrix} y \\ -\mu y - \omega^2 x \end{pmatrix} \\ &= -\frac{\mu}{2} (\omega^2 x^2 + y^2), \end{aligned}$$

abbiamo $\nabla \Lambda \cdot \mathbf{F}(x) < 0$, ad eccezione del punto di equilibrio. Pertanto $(x, y) \equiv (0, 0)$ è asintoticamente stabile.

2.12 Moto armonico smorzato con forzante esterna

Consideriamo adesso un'equazione scalare del second'ordine lineare, non omogenea⁴⁴

$$\ddot{x} + \nu \dot{x} + \omega^2 x = f(t), \quad (2.111)$$

che corrisponde, dal punto di vista fisico, ad un oscillatore armonico smorzato soggetto ad una forza $f(t)$ variabile nel tempo, usualmente detta **forzante esterna**, o **termine forzante**. La soluzione generale di (2.111) si ottiene sommando una qualsiasi soluzione di (2.111) alla soluzione generale dell'equazione lineare omogenea (2.109) che abbiamo analizzato nella sezione 2.11.

⁴⁴In questa sezione adottiamo la notazione "standard" in cui il coefficiente di \dot{x} viene denotato con ν , anziché con μ .

In questo paragrafo analizzeremo il caso particolare in cui il termine forzante di tipo sinusoidale $f(t) = a \cos(\omega_f t)$, con ω_f pulsazione della forzante. Non occorre introdurre una fase specifica per il termine forzante, basta considerare $t = 0$ quando la forza raggiunge il suo massimo. Riscriviamo quindi la (2.111) nella forma

$$\ddot{x} + \nu \dot{x} + \omega^2 x = a \cos(\omega_f t). \quad (2.112)$$

Il caso speciale $\nu = 0$ corrisponde ad un oscillatore armonico soggetto ad una forza esterna.

Determiniamo ora una soluzione particolare della (2.112), cercandola nella forma

$$x_P(t) = c_1 \sin(\omega_f t) + c_2 \cos(\omega_f t).$$

Sostituendo $x_P(t)$ nella (2.112) otteniamo

$$\begin{aligned} \omega_f^2 [-c_1 \sin(\omega_f t) - c_2 \cos(\omega_f t)] + \nu \omega_f [c_1 \cos(\omega_f t) - c_2 \sin(\omega_f t)] \\ + \omega^2 [c_1 \sin(\omega_f t) + c_2 \cos(\omega_f t)] = a \cos(\omega_f t), \end{aligned}$$

che è soddisfatta se e solo se i coefficienti di $\cos(\omega_f t)$ e $\sin(\omega_f t)$ sono nulli, ovvero se i coefficienti c_1 e c_2 soddisfano il sistema

$$\begin{cases} (\omega^2 - \omega_f^2)c_1 - \nu \omega_f c_2 = 0, \\ \nu \omega_f c_1 + (\omega^2 - \omega_f^2)c_2 = a. \end{cases} \quad (2.113)$$

Il sistema (2.113) ha sempre soluzione eccettuato il caso in cui $\omega_f = \omega$ e $\nu = 0$, ovvero nel caso di un oscillatore armonico privo di attrito a cui sia applicata una forzante della stessa frequenza delle oscillazioni libere del sistema⁴⁵. La soluzione di (2.113) è data da

$$\begin{aligned} c_1 &= \frac{a\nu\omega_f}{\nu^2\omega_f^2 + (\omega^2 - \omega_f^2)^2}, \\ c_2 &= \frac{a(\omega^2 - \omega_f^2)}{\nu^2\omega_f^2 + (\omega^2 - \omega_f^2)^2}, \end{aligned}$$

da cui possiamo ricostruire la la soluzione particolare.

E' tuttavia più interessante riscrivere la soluzione particolare nella forma $A \cos(\omega_f t - \beta)$, dove A è detta ampiezza A , e β ritardo di fase. Abbiamo quindi

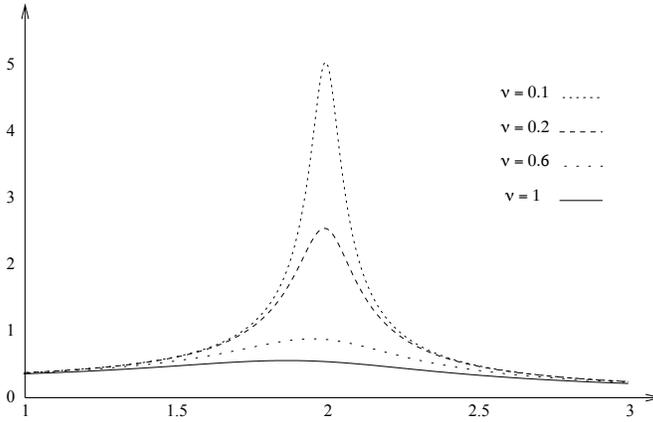
$$x_P(t) = \frac{a}{\sqrt{\nu^2\omega_f^2 + (\omega^2 - \omega_f^2)^2}} \cos\left(\omega_f t - \arctan\left(\frac{\nu\omega_f}{\omega^2 - \omega_f^2}\right)\right). \quad (2.114)$$

⁴⁵Questo fenomeno viene detto risonanza. In questo caso la soluzione particolare è data da

$$x_P(t) = Ct \cos(\omega t), \quad C = \text{costante},$$

ovvero un'oscillazione con ampiezza che cresce linearmente nel tempo.

Figura 2.11: Ampiezza A data dalla (2.115) in funzione di ω_f , per diversi valori di ν , considerando $\omega = 2$.



Di conseguenza

$$A = \frac{a}{\sqrt{\nu^2 \omega_f^2 + (\omega^2 - \omega_f^2)^2}}, \quad (2.115)$$

$$\beta = \arctan\left(\frac{\nu \omega_f}{\omega^2 - \omega_f^2}\right). \quad (2.116)$$

E' importante notare, come abbiamo sottolineato nella sezione 2.11, che la soluzione generale dell'equazione omogenea, che sommata a (2.114) ci fornisce la soluzione generale della (2.112), decade esponenzialmente a zero. Questo significa che dopo un tempo sufficientemente "lungo" (ma in realtà "breve" nella scala dei tempi in gioco⁴⁶, e tanto più breve quanto ν è "grande") la soluzione che "rimane" (ovvero quella che è "osservabile" dal punto di vista fisico) è data dalla soluzione particolare, *indipendentemente dalle condizioni iniziali del moto*.

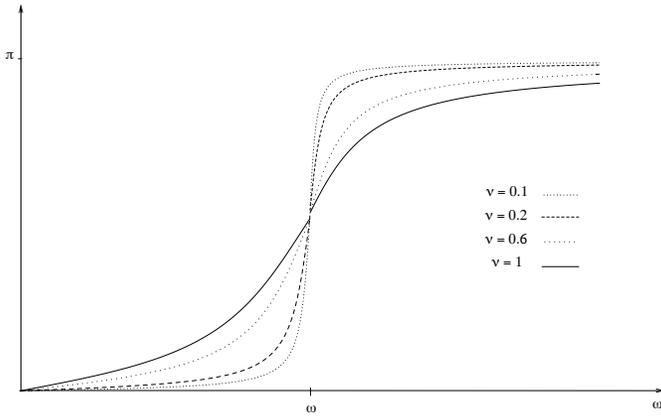
Le figure 2.11 e 2.12 mostrano rispettivamente i grafici del rapporto tra l'ampiezza della soluzione e l'ampiezza del termine forzante e del ritardo di fase al variare della frequenza forzante ω_f per un valore fissato di ω e per diversi valori del coefficiente di smorzamento ν .

Si osserva che il valore massimo dell'ampiezza A , data dalla (2.115), viene assunto per $\omega_f = \sqrt{\omega^2 - \nu^2/2}$ (che è un numero reale se ν è sufficientemente piccolo), e può superare 1 se è soddisfatta da disuguaglianza

$$1 > \nu^2 \omega_f^2 + (\omega^2 - \omega_f^2)^2,$$

ovvero se $\nu^2 \omega^2 < 1$. In questo caso l'ampiezza della "risposta" è maggiore di quella del termine forzante. In altre parole, il sistema funziona da *amplificatore* del termi-

⁴⁶Il tempo *caratteristico* del fenomeno è dato dal periodo dell'oscillazione forzante, ovvero $T = 2\pi/\omega_f$.

Figura 2.12: Ritardo di fase, per diversi valori di ν , con $\omega = 2$ 

ne forzante. Questo fenomeno, detto **risonanza**, è particolarmente interessante nelle applicazioni elettrotecniche. Infatti l'equazione (2.111) è anche l'equazione per l'intensità di corrente in un circuito RLC soggetto ad una tensione $f(t)$.

2.13 Sistemi lineari n -dimensionali

Nella sezione 2.10 abbiamo visto la soluzione dei sistemi lineari 2×2 , tuttavia la "strategia" di risoluzione si estende a sistemi di qualsiasi dimensione

$$\dot{\mathbf{X}} = \mathbb{A}\mathbf{X} \quad (2.117)$$

dove $\mathbf{X}(t) \in \mathbb{R}^n$ e $\mathbb{A} \in \mathbb{R}^{n \times n}$. In questa sezione però affrontiamo la soluzione del sistema lineare (2.117) con un approccio alternativo, che evidentemente vale anche nel caso bidimensionale (ovvero nel caso in cui \mathbb{A} è una matrice 2×2 e $\mathbf{X}(t) \in \mathbb{R}^2$).

Sia \mathbf{X}_0 il dato iniziale (a $t = 0$) per il vettore $\mathbf{X}(t)$, e integriamo la (2.117) fra 0 e t , ottenendo

$$\mathbf{X}(t) = \mathbf{X}_0 + \int_0^t \mathbb{A}\mathbf{X}(s) ds. \quad (2.118)$$

Tentiamo poi di costruire la soluzione di (2.117) "per approssimazioni successive", ossia costruendo la successione $\mathbf{X}_n(t)$ definita per ricorrenza

$$\mathbf{X}_{n+1}(t) = \mathbf{X}_0 + \int_0^t \mathbb{A}\mathbf{X}_n(s) ds, \quad n = 1, 2, 3, \dots \quad (2.119)$$

È chiaro che se la successione $\mathbf{X}_n(t)$ ammette limite, ponendo

$$\mathbf{X}(t) = \lim_{n \rightarrow \infty} \mathbf{X}_n(t),$$

allora, almeno formalmente, la $\mathbf{X}(t)$ soddisfa la (2.118). Derivando quest'ultima otteniamo (2.117): il limite $\mathbf{X}(t)$ è dunque la soluzione cercata. Dalla (2.119) si ottiene,

ricorsivamente,

$$\begin{aligned} \mathbf{X}_1(t) &= \mathbf{X}_0 + \int_0^t \mathbb{A} \mathbf{X}_0 ds = \mathbf{X}_0 + t \mathbb{A} \mathbf{X}_0, \\ \mathbf{X}_2(t) &= \mathbf{X}_0 + \int_0^t \mathbb{A} \mathbf{X}_1(s) ds = \mathbf{X}_0 + t \mathbb{A} \mathbf{X}_0 + \frac{t^2}{2} \mathbb{A}^2 \mathbf{X}_0, \\ \mathbf{X}_3(t) &= \mathbf{X}_0 + \int_0^t \mathbb{A} \mathbf{X}_2(s) ds = \mathbf{X}_0 + t \mathbb{A} \mathbf{X}_0 + \frac{t^2}{2} \mathbb{A}^2 \mathbf{X}_0 + \frac{t^3}{6} \mathbb{A}^3 \mathbf{X}_0, \\ &\dots \end{aligned}$$

ovvero

$$\mathbf{X}_n(t) = \sum_{k=0}^n \frac{t^k}{k!} \mathbb{A}^k \mathbf{X}_0.$$

Si può facilmente dimostrare che la serie *converge uniformemente* (su intervalli di tempo limitati) e che il suo limite

$$\mathbf{X}(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{A}^k \mathbf{X}_0, \quad (2.120)$$

risolve la (2.117). La matrice $\sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{A}^k$ è detta **esponenziale della matrice \mathbb{A}** e viene indicata con $e^{t\mathbb{A}}$

$$e^{t\mathbb{A}} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{A}^k. \quad (2.121)$$

Dunque possiamo concludere in generale che la soluzione del problema di Cauchy

$$\begin{cases} \dot{\mathbf{X}} = \mathbb{A} \mathbf{X}, \\ \mathbf{X}(0) = \mathbf{X}_0, \end{cases}$$

è data dalla matrice esponenziale di \mathbb{A} applicata al vettore dei dati iniziali:

$$\mathbf{X}(t) = e^{t\mathbb{A}} \mathbf{X}_0. \quad (2.122)$$

Nota 2.13.1 Osserviamo che se \mathbf{X}_λ è un autovettore di \mathbb{A} con autovalore λ allora \mathbf{X}_λ è anche autovettore di $e^{t\mathbb{A}}$ con autovalore $e^{t\lambda}$. Infatti

$$e^{t\mathbb{A}} \mathbf{X}_\lambda = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{A}^k \mathbf{X}_\lambda = \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda^k \mathbf{X}_\lambda = e^{t\lambda} \mathbf{X}_\lambda.$$

È facile ritrovare per questa via le soluzioni ottenute nel caso bidimensionale. In particolare, è istruttivo costruire la matrice esponenziale nel caso in cui \mathbb{A} abbia due autovalori complessi coniugati. Sappiamo già che in questo caso, introducendo, tramite la (2.84), le nuove variabili $\xi(t)$ e $\eta(t)$, possiamo riscrivere il sistema di partenza nella forma (2.97). In particolare, denotando con \mathbb{B} la matrice $\mathbb{V}^{-1} \mathbb{A} \mathbb{V}$, cioè

$$\mathbb{B} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

abbiamo

$$\mathbb{B} = \alpha \mathbb{I} + \beta \mathbb{J}, \quad \text{dove } \mathbb{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

e pertanto la soluzione di (2.97) potrà essere scritta come

$$\begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} = e^{t\mathbb{B}} \begin{pmatrix} \xi_o \\ \eta_o \end{pmatrix},$$

con

$$e^{t\mathbb{B}} = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{t^k}{k!} \mathbb{B}^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{t^k}{k!} (\alpha \mathbb{I} + \beta \mathbb{J})^k.$$

Applichiamo adesso la formula del binomio di Newton⁴⁷

$$(\alpha \mathbb{I} + \beta \mathbb{J})^k = \sum_{j=0}^k \underbrace{\binom{k}{j}}_{\frac{k!}{j!(k-j)!}} (\alpha \mathbb{I})^j (\beta \mathbb{J})^{k-j},$$

ottenendo

$$\begin{aligned} \sum_{k=0}^n \frac{t^k}{k!} \mathbb{B}^k &= \sum_{k=0}^n \frac{t^k}{k!} \left[\sum_{j=0}^k \frac{k!}{j!(k-j)!} (\alpha \mathbb{I})^j (\beta \mathbb{J})^{k-j} \right] \\ &= \sum_{k=0}^n \sum_{j=0}^k \frac{t^k}{j!(k-j)!} (\alpha \mathbb{I})^j (\beta \mathbb{J})^{k-j} = \sum_{j=0}^n \sum_{k=j}^n \frac{t^k}{j!(k-j)!} (\alpha \mathbb{I})^j (\beta \mathbb{J})^{k-j} \\ &= \sum_{j=0}^n \sum_{r=0}^n \frac{t^{j+r}}{j!r!} (\alpha \mathbb{I})^j (\beta \mathbb{J})^r = \sum_{j=0}^n \frac{t^j}{j!} (\alpha \mathbb{I})^j \sum_{r=0}^n \frac{t^r}{r!} (\beta \mathbb{J})^r, \end{aligned}$$

dove si è utilizzato il cambio di indice $r = k - j$. Dunque, per $n \rightarrow \infty$ si ha

$$e^{t\mathbb{B}} = e^{\alpha t \mathbb{I}} e^{t\beta \mathbb{J}} = e^{\alpha t} e^{t\beta \mathbb{J}}, \quad (2.123)$$

dal momento che

$$\lim_{n \rightarrow \infty} \sum_{j=0}^n \frac{t^j}{j!} (\alpha \mathbb{I})^j = \left(\lim_{n \rightarrow \infty} \sum_{j=0}^n \frac{\alpha^j t^j}{j!} \right) \mathbb{I} = e^{\alpha t} \mathbb{I}.$$

Quindi, per completare il calcolo di $e^{t\mathbb{B}}$ dobbiamo valutare esplicitamente $e^{t\beta \mathbb{J}}$. Come passo preliminare, osserviamo che $\mathbb{J}^2 = -\mathbb{I}$ e dunque

$$\mathbb{J}^{2k} = (\mathbb{J}^2)^k = (-1)^k \mathbb{I}, \quad \mathbb{J}^{2k+1} = \underbrace{\mathbb{J}^{2k}}_{(-1)^k \mathbb{I}} \mathbb{J} = (-1)^k \mathbb{J}, \quad k = 0, 1, 2, \dots$$

⁴⁷Si osservi che possiamo applicare la formula del binomio di Newton perché le matrici \mathbb{I} e \mathbb{J} commutano, cioè $\mathbb{I}\mathbb{J} = \mathbb{J}\mathbb{I}$. Infatti, date due matrici qualsiasi \mathbb{P} e \mathbb{Q} ,

$$(\mathbb{P} + \mathbb{Q})^2 = \mathbb{P}^2 + \mathbb{Q}^2 + \mathbb{P}\mathbb{Q} + \mathbb{Q}\mathbb{P}.$$

La classica formula del quadrato del binomio si ha solamente se $\mathbb{P}\mathbb{Q} = \mathbb{Q}\mathbb{P}$.

Pertanto

$$\begin{aligned} \sum_{k=0}^{2n} \frac{(t\beta)^k}{k!} \mathbb{J}^k &= \underbrace{\sum_{k=0}^n \frac{(t\beta)^{2k}}{(2k)!} \mathbb{J}^{2k}}_{\text{somma sugli indici pari}} + \underbrace{\sum_{k=0}^{n-1} \frac{(t\beta)^{(2k+1)}}{(2k+1)!} \mathbb{J}^{(2k+1)}}_{\text{somma sugli indici dispari}} \\ &= \sum_{k=0}^n (-1)^k \frac{(t\beta)^{(2k)}}{(2k)!} \mathbb{I} + \sum_{k=0}^{n-1} (-1)^k \frac{(t\beta)^{(2k+1)}}{(2k+1)!} \mathbb{J}. \end{aligned}$$

Consideriamo adesso il limite per $n \rightarrow \infty$, ricordando gli sviluppi in serie di Taylor delle funzioni $\sin t\beta$ e $\cos t\beta$. Abbiamo

$$\underbrace{\sum_{k=0}^{\infty} \frac{(t\beta)^k}{k!} \mathbb{J}^k}_{e^{t\beta\mathbb{J}}} = \underbrace{\left[\sum_{n=0}^{\infty} (-1)^n \frac{(t\beta)^{(2n)}}{(2n)!} \right]}_{\sin t\beta} \mathbb{I} + \underbrace{\left[\sum_{n=0}^{\infty} (-1)^n \frac{(t\beta)^{(2n+1)}}{(2n+1)!} \right]}_{\cos t\beta} \mathbb{J},$$

ovvero

$$e^{t\beta\mathbb{J}} = \cos(t\beta)\mathbb{I} + \sin(t\beta)\mathbb{J} = \begin{pmatrix} \cos(t\beta) & \sin(t\beta) \\ -\sin(t\beta) & \cos(t\beta) \end{pmatrix}.$$

Sostituendo quest'ultima nella (2.123) abbiamo

$$e^{t\mathbb{B}} = e^{\alpha t} \begin{pmatrix} \cos(t\beta) & \sin(t\beta) \\ -\sin(t\beta) & \cos(t\beta) \end{pmatrix},$$

che ci fa ritrovare la soluzione data dalla (2.105) (a meno delle costanti ρ_0 e θ_0 che sono legate ai dati iniziali, ovvero a \mathbf{X}_0).

Capitolo 3

Le Equazioni di Lagrange

Questo capitolo è dedicato all'introduzione della formulazione lagrangiana della meccanica che di fatto è un formalismo “potente” per ottenere le equazioni di moto di sistemi meccanici anche molto complessi¹. In tutto questo capitolo considereremo un solo punto materiale non soggetto ad alcun vincolo. Prima di introdurre le equazioni di Lagrange, discuteremo la cinematica del punto materiale quando i vettori velocità ed accelerazione vengono espressi rispetto ad una base locale. Passeremo poi alle forze, introducendo il concetto di forze conservative e quindi vedremo il metodo di Lagrange per giungere alle equazioni di moto. Come applicazione illustreremo il moto di un punto materiale libero soggetto ad una forza di tipo centrale, quale, per esempio, quella di attrazione gravitazionale.

3.1 Cinematica del punto

Consideriamo uno spazio affine euclideo² A con dimensione 3 in cui è definito un riferimento cartesiano ortogonale. Il moto di un punto viene rappresentato da una curva parametrizzata dal tempo t , ovvero $\mathcal{I} \ni t \rightarrow P(t) \in A$, dove $\mathcal{I} \subset \mathbb{R}$, è un intervallo temporale. Ad ogni istante t corrisponde un punto dello spazio affine. L'immagine in A di questa applicazione si chiama *traiettoria*, o *orbita*, di P . Ad ogni istante $t \in \mathcal{I}$ corrisponde un vettore $(P(t) - O)$, che denoteremo semplicemente con $\mathbf{x}(t)$. Se $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, è una base ortogonale di V , $\mathbf{x}(t)$ è individuato dal seguente vettore³ di \mathbb{R}^3

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, \quad \Leftrightarrow \quad \mathbf{x}(t) = x(t)\mathbf{e}_x + y(t)\mathbf{e}_y + z(t)\mathbf{e}_z.$$

¹Tale formalismo fu introdotto da J. L. Lagrange (Torino 1736 – Parigi 1813), matematico e astronomo italiano, che svolse la sua attività scientifica a Berlino e a Parigi. Infatti, nell'introduzione del suo trattato *Mécanique analytique*, pubblicato nel 1788, Lagrange scrive: *Je me suis proposé de réduire la théorie de cette Science, et l'art de résoudre les problèmes qui s'y rapportent, à des formules générales, dont le simple développement donne toute les équations nécessaires pour la solution de chaque problème.*

²Più precisamente con A intenderemo sempre la terna (A, V, F) dove A è l'insieme dei punti, V è lo spazio vettoriale soggiacente lo spazio affine e F è l'applicazione che alla coppie di punti associa i rispettivi vettori di V . In A è definito il prodotto scalare \cdot .

³Si assume che le funzioni $x(t)$, $y(t)$ e $z(t)$ siano almeno $C^2(\mathcal{I})$.

La **velocità** e l'**accelerazione** del punto P , sono definite come i vettori

$$\begin{aligned}(P(t) - O) &= \frac{d\mathbf{x}(t)}{dt} = \dot{x}(t) \mathbf{e}_x + \dot{y}(t) \mathbf{e}_y + \dot{z}(t) \mathbf{e}_z, \\(P(t) - O) &= \frac{d^2\mathbf{x}(t)}{dt^2} = \ddot{x}(t) \mathbf{e}_x + \ddot{y}(t) \mathbf{e}_y + \ddot{z}(t) \mathbf{e}_z.\end{aligned}$$

Supponiamo adesso che sia definito, in un dominio $\mathfrak{D}_A \subset A$, un sistema di coordinate curvilinee

$$\mathbb{R}^3 \supset Q \ni \mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \xrightarrow{x} \mathbf{x}(\mathbf{q}).$$

Quindi il moto del punto, fino a quando avviene in \mathfrak{D}_A , potrà anche essere rappresentato dando le tre funzioni $(q_1(t), q_2(t), q_3(t))$, ovvero $\mathbf{q}(t)$. Avremo pertanto

$$P - O = \mathbf{x}(\mathbf{q}(t)), \Leftrightarrow \begin{cases} x = x(\mathbf{q}(t)) = x(q_1(t), q_2(t), q_3(t)) \\ y = y(\mathbf{q}(t)) = y(q_1(t), q_2(t), q_3(t)) \\ z = z(\mathbf{q}(t)) = z(q_1(t), q_2(t), q_3(t)) \end{cases}$$

Il vettore velocità si potrà anche scrivere come segue

$$\dot{\mathbf{x}} = \underbrace{\frac{\partial \mathbf{x}}{\partial q_1}}_{\mathbf{u}_1} \dot{q}_1 + \underbrace{\frac{\partial \mathbf{x}}{\partial q_2}}_{\mathbf{u}_2} \dot{q}_2 + \underbrace{\frac{\partial \mathbf{x}}{\partial q_3}}_{\mathbf{u}_3} \dot{q}_3, \quad (3.1)$$

dove $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ è la *base locale* (v. capitolo 1, sezione 1.4). Quindi, rispetto alla base locale, il vettore velocità è rappresentato dalla terna

$$\begin{pmatrix} \dot{q}_1(t) \\ \dot{q}_2(t) \\ \dot{q}_3(t) \end{pmatrix}.$$

Per esser più precisi, riferendoci sempre alla sezione 1.4, la terna delle \dot{q}_i , $i = 1, 2, 3$, è la terna delle *componenti controvarianti* del vettore velocità rispetto alla base locale (evidentemente se la base locale è ortonormale questa precisazione è inutile).

Definizione 3.1.1 Se il punto P ha massa m , si definisce *energia cinetica* di P

$$T = \frac{m}{2} \left\| \frac{d\mathbf{x}(t)}{dt} \right\|^2.$$

Calcoliamo adesso l'energia cinetica di P considerando le coordinate curvilinee $\mathbf{q}(t)$. Otteniamo

$$\begin{aligned}T &= \frac{m}{2} \left(\sum_{h=1}^3 \mathbf{u}_h \dot{q}_h \right) \cdot \left(\sum_{k=1}^3 \mathbf{u}_k \dot{q}_k \right) \\ &= \frac{m}{2} \sum_{h,k=1}^3 (\mathbf{u}_h \cdot \mathbf{u}_k) \dot{q}_h \dot{q}_k,\end{aligned}$$

che viene anche scritta come⁴

$$T = \frac{1}{2} \sum_{h,k=1}^3 a_{hk} \dot{q}_h \dot{q}_k, \quad \text{dove } a_{hk} = m \mathbf{u}_h \cdot \mathbf{u}_k. \quad (3.2)$$

In particolare, se la base locale è ortogonale, avremo⁵

$$T = \frac{m}{2} \sum_{h=1}^3 \left\| \frac{\partial \mathbf{x}}{\partial q_h} \right\|^2 \dot{q}_h^2.$$

In generale abbiamo⁶ $a_{hk} = a_{hk}(\mathbf{q})$, e, introducendo la matrice simmetrica⁷ \mathbb{A} , le cui componenti a_{ij} , sono definite nella (3.2),

$$\mathbb{A}(\mathbf{q}) = m \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \mathbf{u}_1 \cdot \mathbf{u}_3 \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \mathbf{u}_2 \cdot \mathbf{u}_3 \\ \mathbf{u}_3 \cdot \mathbf{u}_1 & \mathbf{u}_3 \cdot \mathbf{u}_2 & \mathbf{u}_3 \cdot \mathbf{u}_3 \end{pmatrix}, \quad (3.3)$$

potremo scrivere

$$T = \frac{1}{2} \dot{\mathbf{q}}^T \mathbb{A}(\mathbf{q}) \dot{\mathbf{q}}, \quad (3.4)$$

con

$$\dot{\mathbf{q}}^T = (\dot{q}_1 \quad \dot{q}_2 \quad \dot{q}_3), \quad \text{e} \quad \dot{\mathbf{q}} = \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix}.$$

Per quanto riguarda l'accelerazione, abbiamo questo risultato fondamentale che consente di esprimere le componenti⁸ del vettore $m \frac{d^2 \mathbf{x}(t)}{dt^2}$ nella base locale $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

Teorema 3.1.1 *Se T è l'energia cinetica di P , allora*

$$m \frac{d^2 \mathbf{x}(t)}{dt^2} \cdot \mathbf{u}_k = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k}, \quad i = 1, 2, 3. \quad (3.5)$$

*L'espressione di destra nella (3.5) viene usualmente detta **binomio di Lagrange**.*

Dim. Prima di iniziare la dimostrazione è interessante osservare che la (3.5) ha una struttura completamente diversa dalla (3.1). Infatti, per calcolare nella base puntuale

⁴Evidentemente $a_{hk} = a_{kh}$.

⁵Ricordiamo che, in generale, non è detto che $\mathbf{u}_i = \frac{\partial \mathbf{x}}{\partial q_i}$, $i = 1, 2, 3$, siano normalizzati.

⁶ $a_{hk}(\mathbf{q})$ è una scrittura sintetica per esprimere $a_{hk} = a_{hk}(q_1, q_2, q_3)$.

⁷Osserviamo che la matrice \mathbb{A} è, a parte il coefficiente m , la matrice metrica \mathbb{G} della base locale, definita, nel capitolo 1, dalla (1.8).

⁸Per la precisione la formula (3.5) dà le componenti covarianti di $m \frac{d^2 \mathbf{x}(t)}{dt^2}$ rispetto alla base locale, ovvero fornisce le proiezioni ortogonali del vettore $m \frac{d^2 \mathbf{x}(t)}{dt^2}$ sui vettori della base locale (v. capitolo 1, sezione 1.4).

$\frac{d^2 \mathbf{x}(t)}{dt^2}$ dobbiamo tener presente che i vettori della base locale cambiano da punto a punto e quindi cambieranno durante il moto. Abbiamo

$$\begin{aligned} \frac{\partial T}{\partial \dot{q}_k} &= m \frac{\partial \dot{\mathbf{x}}}{\partial \dot{q}_k} \cdot \dot{\mathbf{x}} \stackrel{(3.1)}{=} m \frac{\partial}{\partial \dot{q}_k} \left(\sum_{j=1}^3 \dot{q}_j \mathbf{u}_j \right) \cdot \dot{\mathbf{x}}. \\ &= m \mathbf{u}_k \cdot \dot{\mathbf{x}}, \end{aligned}$$

per cui

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) = m \frac{d\mathbf{u}_k}{dt} \cdot \dot{\mathbf{x}} + m \mathbf{u}_k \cdot \ddot{\mathbf{x}}. \quad (3.6)$$

Adesso calcoliamo

$$\begin{aligned} \frac{\partial T}{\partial q_k} &= m \frac{\partial \dot{\mathbf{x}}}{\partial q_k} \cdot \dot{\mathbf{x}} = m \frac{\partial}{\partial q_k} \left(\frac{d\mathbf{x}}{dt} \right) \cdot \dot{\mathbf{x}} \\ &= m \underbrace{\frac{d}{dt} \left(\frac{\partial \mathbf{x}}{\partial q_k} \right)}_{\mathbf{u}_k} \cdot \dot{\mathbf{x}} = m \frac{d\mathbf{u}_k}{dt} \cdot \dot{\mathbf{x}}. \end{aligned} \quad (3.7)$$

Da un semplice confronto delle (3.6) e (3.7), si ricava

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) = \frac{\partial T}{\partial q_k} + m \mathbf{u}_k \cdot \ddot{\mathbf{x}},$$

da cui la (3.5) discende banalmente.

□

Esempio 3.1.1 Vediamo che la (3.5) fornisce le componenti del vettore accelerazione anche rispetto al riferimento cartesiano ortonormale $\{O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$. Infatti

$$\mathbf{x}(t) = x(t) \mathbf{e}_x + y(t) \mathbf{e}_y + z(t) \mathbf{e}_z, \Rightarrow \dot{\mathbf{x}}(t) = \dot{x}(t) \mathbf{e}_x + \dot{y}(t) \mathbf{e}_y + \dot{z}(t) \mathbf{e}_z,$$

per cui $T = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$, ovvero

$$T = \frac{1}{2} \left(\dot{x} \quad \dot{y} \quad \dot{z} \right) \underbrace{\begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}}_{\mathbb{A}=m\mathbb{I}} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix},$$

da cui otteniamo

$$m\ddot{\mathbf{x}} \cdot \mathbf{e}_x = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{x}} \right) - \frac{\partial T}{\partial x} = m\ddot{x},$$

$$m\ddot{\mathbf{x}} \cdot \mathbf{e}_y = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{y}} \right) - \frac{\partial T}{\partial y} = m\ddot{y},$$

$$m\ddot{\mathbf{x}} \cdot \mathbf{e}_z = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{z}} \right) - \frac{\partial T}{\partial z} = m\ddot{z}.$$

Esempio 3.1.2 Riferendoci all'esempio 1.4.1 del capitolo 1, consideriamo le coordinate polari nel piano

$$\mathbf{x}(\mathbf{q}) = \begin{pmatrix} x(r, \theta) = r \cos \theta \\ y(r, \theta) = r \sin \theta \end{pmatrix}, \Rightarrow \mathbf{u}_r = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \mathbf{u}_\theta = \begin{pmatrix} -r \sin \theta \\ r \cos \theta \end{pmatrix}.$$

Supponiamo adesso che $P = P(t)$, ovvero $r = r(t)$, $\theta = \theta(t)$. Il vettore velocità $\dot{\mathbf{x}} = (\dot{P} - O)$, può essere espresso sia rispetto alla base $\{\mathbf{e}_x, \mathbf{e}_y\}$,

$$\dot{\mathbf{x}} = (\dot{r} \cos \theta - r\dot{\theta} \sin \theta) \mathbf{e}_x + (\dot{r} \sin \theta + r\dot{\theta} \cos \theta) \mathbf{e}_y,$$

sia rispetto alla base locale $\{\mathbf{u}_r, \mathbf{u}_\theta\}$

$$\dot{\mathbf{x}} = \dot{r}\mathbf{u}_r + \dot{\theta}\mathbf{u}_\theta.$$

In particolare, introducendo la base locale normalizzata, $\boldsymbol{\kappa}_r = \mathbf{u}_r$, e $\mathbf{u}_\theta = r\boldsymbol{\kappa}_\theta$, avremo

$$\dot{\mathbf{x}} = \dot{r}\boldsymbol{\kappa}_r + r\dot{\theta}\boldsymbol{\kappa}_\theta.$$

Quindi, se m è la massa di P , l'energia cinetica sarà

$$T = \frac{m}{2} (\dot{r}\boldsymbol{\kappa}_r + r\dot{\theta}\boldsymbol{\kappa}_\theta)^2 = \frac{m}{2} (\dot{r}^2 + r^2\dot{\theta}^2),$$

ovvero, ricordando la (3.3) e la (3.4)

$$T = \frac{1}{2} \begin{pmatrix} \dot{r} & \dot{\theta} \end{pmatrix} \underbrace{\begin{pmatrix} m & 0 \\ 0 & mr^2 \end{pmatrix}}_{\mathbb{A}} \begin{pmatrix} \dot{r} \\ \dot{\theta} \end{pmatrix}.$$

Possiamo quindi determinare la componente dell'accelerazione lungo \mathbf{u}_r , ovvero la componente radiale, e quella tangenziale

$$\begin{aligned} m\ddot{\mathbf{x}} \cdot \mathbf{u}_r &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{r}} \right) - \frac{\partial T}{\partial r} = m(\ddot{r} - r\dot{\theta}^2), \\ m\ddot{\mathbf{x}} \cdot \mathbf{u}_\theta &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\theta}} \right) - \frac{\partial T}{\partial \theta} = mr^2\ddot{\theta} + 2mrr\dot{\theta}. \end{aligned}$$

Esempio 3.1.3 Consideriamo adesso l'esempio 1.4.2 delle coordinate sferiche

$$\mathbf{x}(\mathbf{q}) = \begin{pmatrix} x(r, \varphi, \theta) = r \sin \theta \cos \varphi \\ y(r, \varphi, \theta) = r \sin \theta \sin \varphi \\ z(r, \varphi, \theta) = r \cos \theta \end{pmatrix},$$

la cui base locale (ortogonale ma non normalizzata) è

$$\mathbf{u}_r = \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}, \quad \mathbf{u}_\varphi = \begin{pmatrix} -r \sin \theta \sin \varphi \\ r \sin \theta \cos \varphi \\ 0 \end{pmatrix}, \quad \mathbf{u}_\theta = \begin{pmatrix} r \cos \theta \cos \varphi \\ r \cos \theta \sin \varphi \\ -r \sin \theta \end{pmatrix},$$

mentre quella normalizzata è $\boldsymbol{\kappa}_r = \mathbf{u}_r$, $\boldsymbol{\kappa}_\varphi = \frac{\mathbf{u}_\varphi}{r \sin \theta}$, $\boldsymbol{\kappa}_\theta = \frac{\mathbf{u}_\theta}{r}$. Se r , φ , e θ dipendono dal tempo t , esprimendo il vettore velocità di P nella base locale abbiamo

$$\dot{\mathbf{x}} = \dot{r} \mathbf{u}_r + \dot{\varphi} \mathbf{u}_\varphi + \dot{\theta} \mathbf{u}_\theta = \dot{r} \boldsymbol{\kappa}_r + r \dot{\varphi} \sin \theta \boldsymbol{\kappa}_\varphi + r \dot{\theta} \boldsymbol{\kappa}_\theta, \quad (3.8)$$

da cui otteniamo la seguente espressione dell'energia cinetica

$$\begin{aligned} T &= \frac{1}{2} \begin{pmatrix} \dot{r} & \dot{\varphi} & \dot{\theta} \end{pmatrix} \underbrace{\begin{pmatrix} m & 0 & 0 \\ 0 & mr^2 \sin^2 \theta & 0 \\ 0 & 0 & mr^2 \end{pmatrix}}_A \begin{pmatrix} \dot{r} \\ \dot{\varphi} \\ \dot{\theta} \end{pmatrix} \\ &= \frac{m}{2} \left(\dot{r}^2 + r^2 \dot{\varphi}^2 \sin^2 \theta + r^2 \dot{\theta}^2 \right). \end{aligned} \quad (3.9)$$

Valutando adesso le componenti di $m\ddot{\mathbf{x}}$ rispetto alla base locale $\{\mathbf{u}_r, \mathbf{u}_\varphi, \mathbf{u}_\theta\}$, si ha

$$\begin{aligned} m\ddot{\mathbf{x}} \cdot \mathbf{u}_r &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{r}} \right) - \frac{\partial T}{\partial r} = m \left[\ddot{r} - r \left(\dot{\varphi}^2 \sin^2 \theta + \dot{\theta}^2 \right) \right], \\ m\ddot{\mathbf{x}} \cdot \mathbf{u}_\varphi &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\varphi}} \right) - \frac{\partial T}{\partial \varphi} = m \frac{d}{dt} \left(r^2 \sin^2 \theta \dot{\varphi} \right), \\ m\ddot{\mathbf{x}} \cdot \mathbf{u}_\theta &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\theta}} \right) - \frac{\partial T}{\partial \theta} = m \left[\frac{d}{dt} \left(r^2 \dot{\theta} \right) - r^2 \dot{\varphi}^2 \sin \theta \cos \theta \right]. \end{aligned}$$

3.2 Forze conservative

3.2.1 Campi scalari e gradiente

Cominciamo con l'introdurre i concetti di campi scalari, campi vettoriali e campi gradiente. Si lavora in uno spazio affine euclideo (A, V, F) dove A è l'insieme dei punti, V è lo spazio vettoriale soggiacente lo spazio affine e F è l'applicazione che alle coppie di punti associa i rispettivi vettori di V . In A è definito il prodotto scalare “ \cdot ”.

Definizione 3.2.1 Dato uno spazio affine euclideo di dimensione n si dice campo scalare su A un'applicazione f da A in \mathbb{R} , $f : A \rightarrow \mathbb{R}$, che opera così

$$P \xrightarrow{f} f(P) \in \mathbb{R}.$$

In generale un campo scalare f è una funzione di \mathbf{x} , ossia delle coordinate cartesiane. Quando sono definite coordinate curvilinee $\mathbf{q} = (q_1, q_2, \dots, q_n)$, f può essere visto come funzione delle \mathbf{q} , vale a dire $\hat{f}(\mathbf{q}) = f(\mathbf{x}(\mathbf{q}))$. Il campo si dice di classe C^k se ammette derivate parziali continue fino all'ordine k .

Definizione 3.2.2 Sia dato un campo scalare f di classe C^1 , e un riferimento cartesiano ortogonormale $\{O, \mathbf{e}_i\}$. Si dice **gradiente del campo scalare** f , o semplicemente **gradiente**, il vettore ∇f , le cui componenti rispetto alla base $\{O, \mathbf{e}_i\}$ sono

$$\nabla f = \frac{\partial f}{\partial x_1} \mathbf{e}_1 + \frac{\partial f}{\partial x_2} \mathbf{e}_2 + \dots + \frac{\partial f}{\partial x_n} \mathbf{e}_n. \quad (3.10)$$

La derivata di f lungo la direzione \mathbf{n} , è la proiezione ortogonale di ∇f lungo \mathbf{n} , cioè $\frac{\partial f}{\partial \mathbf{n}} = \nabla f \cdot \mathbf{n}$.

Sia adesso $\{\boldsymbol{\kappa}_i = \mathbf{u}_i / |\mathbf{u}_i| = \frac{\partial \mathbf{x}}{\partial q_i} / \left| \frac{\partial \mathbf{x}}{\partial q_i} \right|, i = 1, 2, \dots, n\}$ una base locale (normalizzata) associata alle coordinate curvilinee \mathbf{q} , ci chiediamo: come può essere generalizzata la definizione (3.10) alla base locale $\{\boldsymbol{\kappa}_i\}$? O meglio: quali sono le componenti del vettore ∇f nella base $\{\boldsymbol{\kappa}_i\}$? Certamente potremo scrivere $\nabla f = \sum_{i=1}^n (\nabla f \cdot \boldsymbol{\kappa}_i) \boldsymbol{\kappa}_i$, cioè

$$\nabla f = \sum_{i=1}^n (\nabla f \cdot \boldsymbol{\kappa}_i) \boldsymbol{\kappa}_i = \sum_{i=1}^n \left(\nabla f \cdot \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|} \right) \boldsymbol{\kappa}_i = \sum_{i=1}^n \frac{1}{\|\mathbf{u}_i\|} (\nabla f \cdot \mathbf{u}_i) \boldsymbol{\kappa}_i.$$

Ma se $\hat{f}(\mathbf{q}) = f(\mathbf{x}(\mathbf{q}))$, si ha

$$\begin{aligned} \frac{\partial \hat{f}(q_1, q_2, \dots, q_n)}{\partial q_i} &= \frac{\partial}{\partial q_i} f(x_1(q_1, q_2, \dots, q_n), \dots, x_n(q_1, q_2, \dots, q_n)) \\ &= \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial q_i} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial q_i} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial q_i} \\ &= \nabla f \cdot \mathbf{u}_i, \end{aligned} \quad (3.11)$$

e quindi⁹

$$\nabla f = \sum_{i=1}^n \underbrace{\left(\frac{1}{\|\mathbf{u}_i\|} \frac{\partial \hat{f}}{\partial q_i} \right)}_{\nabla f \cdot \boldsymbol{\kappa}_i} \boldsymbol{\kappa}_i. \quad (3.12)$$

Esempio 3.2.1 Consideriamo il piano con le coordinate polari, la cui base locale normalizzata è $\boldsymbol{\kappa}_r = \mathbf{u}_r$, e $\boldsymbol{\kappa}_\varphi = \frac{1}{r} \mathbf{u}_\varphi$. Sia $f(x, y)$ un campo scalare, che espresso rispetto alle coordinate polari, si scriverà come $\hat{f}(r, \varphi) = f(x(r, \varphi), y(r, \varphi))$. L'espressione del gradiente di f rispetto alla base locale è

$$\nabla f = \frac{\partial \hat{f}}{\partial r} \boldsymbol{\kappa}_r + \frac{1}{r} \frac{\partial \hat{f}}{\partial \varphi} \boldsymbol{\kappa}_\varphi,$$

⁹Osserviamo che la (3.12) fornisce un'espressione delle componenti covarianti del gradiente ∇f rispetto alla base locale (cioè delle proiezioni ortogonali del vettore ∇f sui vettori della base $\boldsymbol{\kappa}_i$).

dove $\frac{\partial \hat{f}}{\partial r}$, e $\frac{1}{r} \frac{\partial \hat{f}}{\partial \varphi}$ rappresentano le proiezioni ortogonali di ∇f rispettivamente su $\boldsymbol{\kappa}_r$ e $\boldsymbol{\kappa}_\varphi$.

Esempio 3.2.2 Consideriamo le coordinate sferiche $x = r \sin \theta \cos \varphi$, $y = r \sin \theta \sin \varphi$, $z = r \cos \theta$, la cui base locale normalizzata è $\boldsymbol{\kappa}_r = \mathbf{u}_r$, $\boldsymbol{\kappa}_\varphi = \frac{\mathbf{u}_\varphi}{r \sin \theta}$, $\boldsymbol{\kappa}_\theta = \frac{\mathbf{u}_\theta}{r}$. Al solito, se $f(x, y, z)$ è un campo scalare, $\hat{f}(r, \theta, \varphi) = f(x(r, \theta, \varphi), y(r, \theta, \varphi), z(r, \theta, \varphi))$ è la sua espressione in termini delle coordinate sferiche, si ha

$$\nabla f = \frac{\partial \hat{f}}{\partial r} \boldsymbol{\kappa}_r + \frac{1}{r} \frac{\partial \hat{f}}{\partial \theta} \boldsymbol{\kappa}_\theta + \frac{1}{r \sin \theta} \frac{\partial \hat{f}}{\partial \varphi} \boldsymbol{\kappa}_\varphi \quad (3.13)$$

3.2.2 Campi vettoriali e campi gradiente

Definizione 3.2.3 Un campo vettoriale su uno spazio affine A è un'applicazione $\boldsymbol{\chi} : A \rightarrow V$, che ad ogni punto $P \in A$, associa un vettore $\boldsymbol{\chi}(P) \in V$, applicato in P .

Se in A è dato un riferimento $\{O, \mathbf{u}_i\}$, il campo $\boldsymbol{\chi}$ può essere così rappresentato

$$\boldsymbol{\chi} = \chi_1 \mathbf{u}_1 + \chi_2 \mathbf{u}_2 + \dots + \chi_n \mathbf{u}_n,$$

dove ciascuna χ_i è un campo scalare, ovvero $\chi_i = \chi_i(\mathbf{x})$, $i = 1, 2, \dots, n$. In altri termini il campo vettoriale $\boldsymbol{\chi}$ è rappresentato dalla seguente funzione vettoriale

$$\boldsymbol{\chi}(P) \longrightarrow \boldsymbol{\chi}(\mathbf{x}) = \begin{pmatrix} \chi_1(\mathbf{x}) \\ \chi_2(\mathbf{x}) \\ \vdots \\ \chi_n(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \chi_1(x_1, x_2, \dots, x_n) \\ \chi_2(x_1, x_2, \dots, x_n) \\ \vdots \\ \chi_n(x_1, x_2, \dots, x_n) \end{pmatrix}.$$

Il campo si dirà di classe C^k se ogni componente χ_i è almeno di classe C^k .

Nello spazio affine A una curva Γ è un'applicazione $\Gamma : \mathcal{I} \subset \mathbb{R} \rightarrow A$, che al generico $\eta \in \mathcal{I}$, associa un punto $P(\eta) \in A$. Se in A è dato un sistema cartesiano $\{O, \mathbf{e}_i\}$, la curva si rappresenta come un'applicazione da $\mathcal{I} \subset \mathbb{R} \rightarrow \mathbb{R}^n$, cioè

$$\eta \xrightarrow{\Gamma} \mathbf{x}(\eta) = x_1(\eta) \mathbf{e}_1 + x_2(\eta) \mathbf{e}_2 + \dots + x_n(\eta) \mathbf{e}_n,$$

intendendo con $\mathbf{x}(\eta)$, la n -upla di funzioni $x_i(\eta)$, $i = 1, \dots, n$, che rappresentano le componenti del vettore $(P(\eta) - O)$. La funzione $\mathbf{x}(\eta)$ a valori in \mathbb{R}^n viene anche detta **parametrizzazione della curva** Γ . La curva si dice regolare se $x_i(\eta)$, $i = 1, 2, \dots, n$, è derivabile e $\|\mathbf{x}'(\eta)\| \neq 0$, dove $\mathbf{x}'(\eta) = x'_1(\eta) \mathbf{e}_1 + x'_2(\eta) \mathbf{e}_2 + \dots + x'_n(\eta) \mathbf{e}_n$.

Definizione 3.2.4 Data una curva regolare¹⁰ Γ ed un campo vettoriale $\boldsymbol{\chi}$, si definisce integrale curvilineo del campo lungo la curva Γ

$$\int_{\Gamma} \boldsymbol{\chi} = \int_{\mathcal{I}} \boldsymbol{\chi}(\mathbf{x}(\eta)) \cdot \mathbf{x}'(\eta) d\eta. \quad (3.14)$$

¹⁰E' sufficiente richiedere che Γ sia regolare a tratti, ossia regolare eccetto che in un numero finito di punti.

E' facile provare che $\int_{\Gamma} \chi$ è indipendente dalla parametrizzazione (a meno del segno) della curva.

Definizione 3.2.5 Un campo vettoriale χ si dice **campo gradiente** se esiste un campo scalare f di classe C^1 , tale che $\chi = \nabla f$. f viene usualmente detta **potenziale** del campo vettoriale χ .

Evidentemente non tutti i campi vettoriali sono anche campi gradienti. Le condizioni che caratterizzano i campi gradienti sono riassunte nel seguente

Teorema 3.2.1 Sia χ un campo vettoriale continuo definito in un aperto connesso $U \subseteq A$. Le seguenti affermazioni sono equivalenti:

- (a) χ è un campo gradiente.
- (b) L'integrale curvilineo lungo una qualunque curva che collega i punti P_0 e P_1 è indipendente dalla curva, ma dipende soltanto dagli estremi.
- (c) L'integrale curvilineo lungo una qualunque curva chiusa regolare a tratti è nullo.

Dim. Dimostriamo solo l'implicazione (a) \Rightarrow (b), da cui poi discende (c). Sia infatti $\mathbf{x}(\eta)$, $\eta \in (\eta_0, \eta_1)$, la parametrizzazione di una curva regolare Γ che collega P_0 con P_1 . Abbiamo

$$\begin{aligned} \int_{\Gamma} \chi &= \int_{\eta_0}^{\eta_1} \chi(\mathbf{x}(\eta)) \cdot \mathbf{x}'(\eta) d\eta = \int_{\eta_0}^{\eta_1} \nabla f(\mathbf{x}(\eta)) \cdot \mathbf{x}'(\eta) d\eta \\ &= \int_{\eta_0}^{\eta_1} \sum_{i=1}^n \frac{\partial f(\mathbf{x}(\eta))}{\partial x_i} \cdot x'_i(\eta) d\eta = \int_{\eta_0}^{\eta_1} \frac{df}{d\eta} d\eta \\ &= \underbrace{f(\mathbf{x}(\eta_1))}_{f(P_1)} - \underbrace{f(\mathbf{x}(\eta_0))}_{f(P_0)}. \end{aligned} \quad (3.15)$$

□

Vediamo una caratteristica dei campi vettoriali gradiente. Supponiamo di fissare un riferimento cartesiano ortogonale $\{O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, in cui χ si possa scrivere come $\chi = \chi_x(x, y, z) \mathbf{e}_x + \chi_y(x, y, z) \mathbf{e}_y + \chi_z(x, y, z) \mathbf{e}_z$. Se $f(x, y, z)$ è il potenziale abbiamo

$$\begin{cases} \frac{\partial f}{\partial x} = \chi_x(x, y, z), \\ \frac{\partial f}{\partial y} = \chi_y(x, y, z), \\ \frac{\partial f}{\partial z} = \chi_z(x, y, z). \end{cases} \quad (3.16)$$

Ora, se χ è di classe C^1 , applicando il Teorema di Schwarz, si ha

$$\frac{\partial}{\partial y} \underbrace{\left(\frac{\partial f}{\partial x} \right)}_{\chi_x} = \frac{\partial}{\partial x} \underbrace{\left(\frac{\partial f}{\partial y} \right)}_{\chi_y}, \Leftrightarrow \frac{\partial \chi_x}{\partial y} = \frac{\partial \chi_y}{\partial x}$$

$$\frac{\partial}{\partial z} \underbrace{\left(\frac{\partial f}{\partial x} \right)}_{\chi_x} = \frac{\partial}{\partial x} \underbrace{\left(\frac{\partial f}{\partial z} \right)}_{\chi_z}, \Leftrightarrow \frac{\partial \chi_x}{\partial z} = \frac{\partial \chi_z}{\partial x}$$

$$\frac{\partial}{\partial z} \underbrace{\left(\frac{\partial f}{\partial y} \right)}_{\chi_y} = \frac{\partial}{\partial y} \underbrace{\left(\frac{\partial f}{\partial z} \right)}_{\chi_z}, \Leftrightarrow \frac{\partial \chi_y}{\partial z} = \frac{\partial \chi_z}{\partial y}$$

Ma allora, definendo il **rotore del campo** χ come

$$\nabla \wedge \chi = \left(\frac{\partial \chi_z}{\partial y} - \frac{\partial \chi_y}{\partial z} \right) \mathbf{e}_1 + \left(\frac{\partial \chi_x}{\partial z} - \frac{\partial \chi_z}{\partial x} \right) \mathbf{e}_2 + \left(\frac{\partial \chi_y}{\partial x} - \frac{\partial \chi_x}{\partial y} \right) \mathbf{e}_3,$$

che può essere ottenuto formalmente tramite il seguente “determinate simbolico”

$$\nabla \wedge \chi = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \chi_x & \chi_y & \chi_z \end{vmatrix},$$

si ha

Proposizione 3.2.1 *Se χ è un campo gradiente allora $\nabla \wedge \chi = 0$, ovvero, come si usa spesso dire, χ è irrotazionale*

□

In generale il viceversa non è vero. Non è vero cioè che se $\nabla \wedge \chi = 0$, allora χ è un campo gradiente. La prova di questo fatto è data dal seguente

Esempio 3.2.3 *Consideriamo il campo vettoriale*

$$\chi = \left(\frac{-y}{x^2 + y^2} \right) \mathbf{e}_1 + \left(\frac{x}{x^2 + y^2} \right) \mathbf{e}_2.$$

definito ovunque eccetto che in $x = 0, y = 0$. E' facile verificare che $\nabla \wedge \chi = 0$, ma che non è un campo gradiente. E' infatti sufficiente calcolare $\int_{\Gamma} \chi$ dove Γ è la circonferenza centrata nell'origine di raggio R

$$\Gamma : [0, 2\pi) \longrightarrow A, \quad \mathbf{x}(\eta) = R \cos \eta \mathbf{e}_1 + R \sin \eta \mathbf{e}_2,$$

e rendesi conto che $\int_{\Gamma} \chi \neq 0$. Infatti

$$\begin{aligned} \int_{\Gamma} \chi &= \int_0^{2\pi} \chi(\mathbf{x}(\eta)) \cdot \mathbf{x}'(\eta) d\eta \\ &= \int_0^{2\pi} \left(-\frac{\sin \eta}{R} \mathbf{e}_1 + \frac{\cos \eta}{R} \mathbf{e}_2 \right) \cdot (-R \sin \eta \mathbf{e}_1 + R \cos \eta \mathbf{e}_2) d\eta \\ &= \int_0^{2\pi} d\eta = 2\pi. \end{aligned}$$

Allora χ non soddisfa il punto (c) del teorema 3.2.1, e quindi non è un campo gradiente.

In effetti, la possibilità di invertire la proposizione

$$\chi \text{ campo gradiente} \implies \nabla \wedge \chi = 0,$$

risiede nelle caratteristiche del dominio dove è definito χ . A tal proposito abbiamo il seguente

Teorema. Sia χ un campo vettoriale definito in un dominio semplicemente connesso¹¹ e sia $\nabla \wedge \chi = 0$. Allora χ è un campo gradiente, ovvero esiste un potenziale f definito in tutto il dominio per cui $\chi = \nabla f$.

□

Dunque in un **dominio semplicemente connesso**

$$\text{campo gradiente} \Leftrightarrow \text{campo irrotazionale.}$$

In generale, per determinare la funzione potenziale f di un campo gradiente (alternativamente alla risoluzione di (3.16)) conviene scegliere un percorso regolare a tratti Γ particolarmente semplice che unisca un punto $P_0 \equiv (x_0, y_0, z_0)$, fissato una volta per tutte, con il generico punto $P \equiv (x, y, z)$.

3.2.3 Forza posizionale e forza conservativa

Dato uno spazio affine A , una **forza posizionale**, o **campo di forza**, è un campo vettoriale definito su un aperto connesso $U \subseteq A$, ovvero un'applicazione $\mathbf{F} : U \rightarrow V$, che alla posizione P associa la forza $\mathbf{F}(P)$. Un punto materiale che si trova nella posizione P è dunque soggetto alla forza $\mathbf{F}(P)$. Se $\{O, e_i\}$ è un riferimento cartesiano A , scriveremo

$$\mathbf{F} = F_1(\mathbf{x}) e_1 + F_2(\mathbf{x}) e_2 + F_3(\mathbf{x}) e_3,$$

dove, al solito, con \mathbf{x} indichiamo le coordinate del punto¹² P . Se \mathbf{F} è un *campo gradiente* allora, ricordando la definizione 3.2.5, esiste un campo scalare $V(\mathbf{x})$, usualmente detto **energia potenziale**, per cui¹³

$$\mathbf{F} = -\nabla V. \quad (3.17)$$

Una forza posizionale \mathbf{F} di tipo campo gradiente, viene detta **forza conservativa**.

Adesso, sfruttando il teorema 3.2.1 vogliamo leggere le proprietà dei campi gradienti in termini più fisici. A tal scopo introduciamo il concetto di lavoro compiuto da una forza \mathbf{F} su un punto materiale ad essa soggetto.

¹¹Un aperto connesso C di uno spazio topologico si dice semplicemente connesso se, fissati comunque due punti A, B in C e scelte arbitrariamente due curve $\Gamma_1(\lambda), \Gamma_2(\lambda)$, con $\lambda \in [\lambda_-, \lambda_+]$, che collegano $A = \Gamma_1(\lambda_-) = \Gamma_2(\lambda_-)$ con $B = \Gamma_1(\lambda_+) = \Gamma_2(\lambda_+)$, esiste una trasformazione continua (omotopia) di una curva nell'altra. Formalmente, deve esistere una funzione continua $\psi : [\lambda_-, \lambda_+] \times [0, 1] \rightarrow C$, tale che $\psi(\lambda, 0) = \Gamma_1(\lambda)$, $\psi(\lambda, 1) = \Gamma_2(\lambda)$, $\psi(\lambda_-, \nu) = A$, $\psi(\lambda_+, \nu) = B$, per ogni $\lambda \in [\lambda_+, \lambda_-]$, $\nu \in [0, 1]$.

¹²Vale a dire le componenti del vettore $(P - O)$.

¹³L'energia potenziale $V(\mathbf{x})$ corrisponde al potenziale, cambiato di segno, della definizione 3.2.5.

Definizione 3.2.6 Data una forza¹⁴ \mathbf{F} , il lavoro compiuto da \mathbf{F} nell'unità di tempo su punto materiale, ovvero la potenza, è definito come

$$W = \mathbf{F} \cdot \dot{\mathbf{x}},$$

dove $\dot{\mathbf{x}}$ è la velocità del punto materiale. Il lavoro compiuto da \mathbf{F} nell'intervallo di tempo (t_0, t_1) è dato da¹⁵

$$L(t_0, t_1) = \int_{t_0}^{t_1} \mathbf{F}(\mathbf{x}(\tau)) \cdot \dot{\mathbf{x}}(\tau) d\tau. \quad (3.18)$$

Nella formula (3.18) si riconosce l'integrale curvilineo di \mathbf{F} lungo la traiettoria parametrizzata col tempo. Quindi, considerando le forze posizionali ed applicando semplicemente il teorema 3.2.1, abbiamo la ben nota caratterizzazione delle forze conservative

Teorema 3.2.2 Una forza posizionale \mathbf{F} definita in $U \subseteq A$, U aperto connesso, è conservativa se e sole se il lavoro compiuto da \mathbf{F} in un qualsiasi intervallo di tempo (t_0, t_1) è indipendente dalla traiettoria.

□

Sia ora \mathbf{F} una forza conservativa che, per semplicità, supponiamo definita su tutto lo spazio, ci chiediamo: come possiamo determinarne la relativa energia potenziale? Se $\mathbf{F} = -\nabla V(\mathbf{x})$, si fissa un punto P_o , le cui coordinate sono identificate dal vettore \mathbf{x}_o , dopo di che si considera il generico punto P le cui coordinate sono date dal vettore \mathbf{x} . Sia poi $\Gamma_{(P_o, P)}$ una generica curva regolare a tratti che connette P_o con P , cioè

$$[\eta_0, \eta_1] \xrightarrow{\Gamma_{(P_o, P)}} \mathbf{x}(\eta), \quad \text{tale che } \mathbf{x}(\eta_0) = \mathbf{x}_o, \quad \mathbf{x}(\eta_1) = \mathbf{x}.$$

Abbiamo

$$\begin{aligned} \int_{\Gamma_{(P_o, P)}} \mathbf{F} &= - \int_{\Gamma_{(P_o, P)}} \nabla V \stackrel{(3.14)}{=} - \int_{\eta_0}^{\eta_1} \nabla V(\mathbf{x}(\eta)) \cdot \mathbf{x}'(\eta) d\eta \\ &= - \int_{\eta_0}^{\eta_1} \frac{dV(\mathbf{x}(\eta))}{d\eta} d\eta = - [\underbrace{V(\mathbf{x}(\eta_0))}_{V(\mathbf{x}_o)} - \underbrace{V(\mathbf{x}(\eta_1))}_{V(\mathbf{x})}], \end{aligned}$$

e quindi $V(\mathbf{x}) = - \int_{\Gamma_{(P_o, P)}} \mathbf{F} + V(\mathbf{x}_o)$. Chiaramente $V(\mathbf{x}_o)$ è una costante che quindi possiamo porre uguale a zero (l'energia potenziale è definita a meno di una costante). Pertanto, se $\Gamma(P, P_o)$ è una qualsiasi curva regolare a tratti che, partendo da P_o giunge sino al punto P , si ha

$$V(\mathbf{x}) = - \int_{\Gamma_{(P, P_o)}} \mathbf{F}. \quad (3.19)$$

¹⁴Osserviamo che la definizione di lavoro prescinde dal fatto che la forza \mathbf{F} sia o meno posizionale: la definizione si applica a qualsiasi forza.

¹⁵E' ovviamente sottinteso che $\mathbf{x}(t) \in U, \forall t \in [t_0, t_1]$, essendo U il dominio dove è definita \mathbf{F} .

Concludiamo questa sezione osservando come il senso fisico della definizione 3.2.6 può essere chiarito pensando a cosa succede quando solleviamo un corpo pesante (di massa m). Supponiamo di sollevare il peso di una altezza h lungo la verticale e supponiamo di farlo in modo da mantenere la velocità del peso costante per la maggior parte della ascesa. Questo significa che dovremo applicare, nell'intervallo di tempo in cui il corpo sale con velocità costante, una forza \mathbf{F} uguale e opposta alla forza peso $m\mathbf{g}$ agente sul corpo (la forza dovrà essere inizialmente maggiore del peso per mettere in moto il corpo, e sarà poi minore del peso se riportiamo il corpo allo stato di quiete alla nuova quota). In questo caso, se T è il tempo impiegato per sollevare il peso, il lavoro effettuato dalla forza \mathbf{F} sarà

$$L = \int_0^T \mathbf{F} \cdot \dot{\mathbf{x}} dt = \|\mathbf{F}\| \underbrace{\|\dot{\mathbf{x}}\| T}_h = \|\mathbf{F}\| h,$$

dove h è proprio l'innalzamento subito dal corpo (forza e velocità costante sono paralleli e concordi). Il *lavoro* misura in questo caso “quanta fatica si è dovuta fare per alzare il peso”.

La definizione di lavoro è strettamente legata allo studio delle *macchine semplici* (la leva, il piano inclinato, le pulegge). Vediamo cosa succede se solleviamo il solito peso servendoci di un piano inclinato. Supponendo di poter trascurare l'attrito del peso con il piano inclinato, abbiamo ora che la forza necessaria per sollevare il peso (se esercitata nella direzione parallela al piano inclinato stesso) deve compensare la sola *componente* del peso parallela al piano: la componente normale verrà fornita “gratuitamente” dalla reazione di appoggio sul piano inclinato. Questo fa sì che la forza necessaria per sollevare il peso sia minore di quella necessaria per sollevarlo lungo la verticale. In compenso il *cammino* compiuto sarà maggiore. E' facile vedere che nei due casi è uguale il lavoro compiuto (e quindi la “fatica spesa”).

Esempio 3.2.4 Vediamo alcuni esempi di forze posizionali conservative nello spazio affine tridimensionale.

Forza uniforme, $\mathbf{F} = \mathbf{p}$, campo vettoriale costante. Questo è il caso della **forza peso** che è data da $m\mathbf{g}$, dove m è la massa del corpo e \mathbf{g} l'accelerazione di gravità definita su tutto lo spazio. Consideriamo un riferimento cartesiano ortogonale $\{O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, e supponiamo che $\mathbf{p} = p\mathbf{e}_z$, con p costante. Sia $\mathbf{x} : [t_0, t_1] \rightarrow \mathbb{R}^3$,

$$\mathbf{x}(t) = x(t)\mathbf{e}_x + y(t)\mathbf{e}_y + z(t)\mathbf{e}_z, \quad (3.20)$$

una generica curva regolare che connette i punti P_0 e P , cioè

$$\mathbf{x}(t_0) = \mathbf{x}_0 = x_0\mathbf{e}_x + y_0\mathbf{e}_y + z_0\mathbf{e}_z, \quad \mathbf{x}(t_1) = \mathbf{x} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z. \quad (3.21)$$

Applichiamo la (3.18) e calcoliamo il lavoro $\mathbf{F} = p\mathbf{e}_z$, lungo la curva

$$\begin{aligned} L(t_0, t_1) &= \int_{t_0}^{t_1} p\mathbf{e}_z \cdot (\dot{x}(t)\mathbf{e}_x + \dot{y}(t)\mathbf{e}_y + \dot{z}(t)\mathbf{e}_z) dt = p \int_{t_0}^{t_1} \dot{z}(t) dt \\ &= p(z - z_0). \end{aligned} \quad (3.22)$$

Dal teorema 3.2.2 segue che \mathbf{F} è conservativa.

Per calcolarne l'energia potenziale è sufficiente considerare $\nabla V(\mathbf{x}) = -\mathbf{F} = -p\mathbf{e}_z$, cioè

$$\left. \begin{aligned} \frac{\partial V}{\partial x} &= 0, \\ \frac{\partial V}{\partial y} &= 0, \\ \frac{\partial V}{\partial z} &= -p, \end{aligned} \right\} V(x, y, z) = -pz + C,$$

dove C è una qualsiasi costante che, per semplicità, poniamo uguale a zero. Otteniamo dunque $V(x, y, z) = -pz$. Notiamo che saremmo potuti giungere allo stesso risultato ricorrendo alla (3.19) e sfruttando la (3.22). Infatti

$$V(x, y, z) - V(x_o, y_o, z_o) \stackrel{(3.19)}{=} -L(t_0, t_1) \stackrel{(3.22)}{=} -p(z - z_o),$$

da cui, ponendo $V(x_o, y_o, z_o) = 0$ e scegliendo il punto P_o in modo che $z_o = 0$, otteniamo di nuovo $V(x, y, z) = -pz$. In particolare, se $\mathbf{F} = -mg\mathbf{e}_z$, cioè se $p = -mg$, l'energia potenziale dovuta alla forza peso è

$$V = mgz.$$

Forza di richiamo elastica verso il punto O , $\mathbf{F} = -k(P - O)$. Tale forza è quella dovuta ad una molla di lunghezza a riposo trascurabile, rigidezza (o costante elastica) k , con un estremo fisso nel punto O . \mathbf{F} è definita su tutto lo spazio e, scrivendola in componenti, abbiamo

$$\mathbf{F} = -k\mathbf{x} = -k(x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z),$$

Si dimostra che la forza è conservativa applicando il teorema 3.2.2. Infatti, se $\mathbf{x}(t)$ è la curva (3.20) che connette i punti P_o e P_1 , dati da (3.21), si ha

$$\begin{aligned} L(t_0, t_1) &= \int_{t_0}^{t_1} \mathbf{F}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}} dt \\ &= -k \int_{t_0}^{t_1} (x\dot{x} + y\dot{y} + z\dot{z}) dt = -\frac{k}{2} \int_{t_0}^{t_1} \frac{d}{dt} (x^2 + y^2 + z^2) dt \\ &= -\frac{k}{2} (|\mathbf{x}_1|^2 - |\mathbf{x}_o|^2). \end{aligned}$$

Il lavoro è pertanto indipendente dalla curva che connette P_o e P_1 . La relativa energia potenziale può essere calcolata come nel caso precedente ottenendo

$$V(x, y, z) = \frac{k}{2} (x^2 + y^2 + z^2).$$

L'espressione di V in coordinate sferiche è banalmente $V(r) = \frac{k}{2}r^2$. In particolare, ricordando la (3.13) abbiamo

$$\mathbf{F} = -\frac{\partial V(r)}{\partial r} \boldsymbol{\kappa}_r = -kr\boldsymbol{\kappa}_r.$$

Forza centrale di tipo $\mathbf{F}(P) = f(\|P - O\|) \text{vers}(P - O)$, dove O è un punto fisso dello spazio. Tale forza è analoga, in forma, a quella di richiamo elastica, ma adesso la costante elastica non è uniforme, dipendendo dalla distanza di P dall'origine. In più, la funzione f potrebbe non essere definita nell'origine: questo è il caso, per esempio, della forza di attrazione gravitazionale o della forza elettrostatica. In ogni caso, anche escludendo l'origine, il dominio dove la forza è definita è comunque semplicemente connesso. Per mostrare che la \mathbf{F} è conservativa facciamo vedere che \mathbf{F} soddisfa il punto (c) del teorema 3.2.1. Infatti, ricordando la (1.13) dell'esempio 1.4.2, possiamo scrivere $\mathbf{F} = f(r) \boldsymbol{\kappa}_r$. Fissiamo quindi un punto P_o , cioè $(P_o - O) = r_o \boldsymbol{\kappa}_r$, e consideriamo un generico $(P - O) = r \boldsymbol{\kappa}_r$. Adesso, lavorando in coordinate polari sferiche, consideriamo una generica curva $\Gamma_{(P_o, P)}$ regolare a tratti che connette P_o e P , cioè

$$\mathbf{x}(\eta) = r(\eta) \sin \theta(\eta) \cos \phi(\eta) \mathbf{e}_x + r(\eta) \sin \theta(\eta) \sin \phi(\eta) \mathbf{e}_y + r(\eta) \cos \theta(\eta) \mathbf{e}_z.$$

Dalla (3.14) abbiamo

$$\int_{\Gamma_{(P_o, P)}} \mathbf{F} = \int_{\eta_o}^{\eta_1} \mathbf{F}(\mathbf{x}(\eta)) \cdot \mathbf{x}'(\eta) d\eta.$$

Siccome $\mathbf{x}'(\eta) = r'(\eta) \boldsymbol{\kappa}_r + \theta'(\eta) \mathbf{u}_\theta + \phi'(\eta) \mathbf{u}_\phi$, il precedente integrale diventa

$$\int_{\Gamma_{(P_o, P)}} \mathbf{F} = \int_{\eta_o}^{\eta_1} f(r(\eta)) r'(\eta) d\eta = - \int_{\eta_o}^{\eta_1} \frac{d}{d\eta} [V(r(\eta))] ds,$$

dove

$$V(r) = - \int_{r_o}^r f(s) ds. \quad (3.23)$$

Per cui, se la curva è chiusa, cioè $r = r_o$, si ha $\int_{\Gamma_{(P_o, P_o)}} \mathbf{F} = 0$, e dunque la forza è conservativa. Si noti che la (3.23) è l'espressione per l'energia potenziale V (che dipende soltanto dalla coordinata radiale r). In particolare, ricordando la (3.13), abbiamo

$$\nabla V(r) = \frac{\partial}{\partial r} \left[- \int_{r_o}^r f(s) ds \right] \boldsymbol{\kappa}_r = -f(r) \boldsymbol{\kappa}_r, \Rightarrow -\nabla V(r) = \mathbf{F}.$$

Nel caso della forza di attrazione gravitazionale, la forza con cui il punto materiale P_1 di massa m_1 , attrae il punto materiale P_2 di massa m_2 è

$$\mathbf{F} = -G \frac{m_1 m_2}{r^2} \boldsymbol{\kappa}_r,$$

dove G è la costante di gravitazione universale, $G = 6.67 \times 10^{-11} \text{ N m}^2/\text{kg}^2$. Quindi focalizzandoci sul punto P_2 , la forza cui lo stesso è soggetto è scrivibile come $\mathbf{F} = -\frac{m_2 K}{r^2} \boldsymbol{\kappa}_r$, dove $K = G m_1$, la cui energia potenziale, in coordinate polari sferiche, è

$$V(r) = - \int_{r_o}^r -\frac{mK}{s^2} ds = -\frac{mK}{r} + \frac{mK}{r_o}.$$

Considerando $r_o \rightarrow \infty$, abbiamo $V(r) = -\frac{mK}{r}$.

3.3 Equazioni di Lagrange per un punto materiale

Consideriamo adesso il moto di un punto materiale che si muove nello spazio soggetto ad una forza posizionale conservativa \mathbf{F} , la cui energia potenziale è V . Il moto, in accordo con la seconda legge di Newton, è descritto dall'equazione

$$m \ddot{\mathbf{x}} = \mathbf{F} = -\nabla V(\mathbf{x}). \quad (3.24)$$

L'equazione vettoriale (3.24) si può scrivere in un sistema di coordinate cartesiane ortogonali e da essa si ottiene il sistema

$$\ddot{x}_i = F_i(x_1, x_2, x_3) = -\frac{\partial V}{\partial x_i}, \quad i = 1, 2, 3, \quad (3.25)$$

In molti casi è però preferibile usare un sistema di coordinate locali, per esempio le coordinate sferiche se il campo di forza ha una simmetria centrale rispetto a un punto fisso dello spazio.

La riscrittura di (3.24) in un diverso sistema di coordinate locali richiede una scomposizione diversa dalla (3.25). Vediamo come sia possibile costruire una “ricetta” per la scrittura “automatica” di un sistema di equazioni scalari di moto equivalente alla (3.25) in un qualsiasi sistema di coordinate. Tale procedura inoltre si generalizza a sistemi meccanici composti da un numero qualsiasi di punti materiali interagenti tra loro e con l'esterno ed eventualmente soggetti a limitazioni al loro moto “naturale” espresse da “vincoli olonomi lisci” (il senso esatto di queste espressioni verrà chiarito nel capitolo 4).

Supponiamo che siano date le coordinate curvilinee $\mathbf{q} = (q_1, q_2, q_3)$,

$$\mathbf{x} = \mathbf{x}(\mathbf{q}), \quad \Leftrightarrow \quad \begin{cases} x = x(q_1, q_2, q_3), \\ y = y(q_1, q_2, q_3), \\ z = z(q_1, q_2, q_3), \end{cases}$$

e che $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ sia la relativa base locale. Sia inoltre

$$\hat{V}(q_1, q_2, q_3) = V(x(q_1, q_2, q_3), y(q_1, q_2, q_3), z(q_1, q_2, q_3)).$$

Adesso proiettiamo¹⁶ la (3.24) lungo i vettori della base locale

$$m \ddot{\mathbf{x}} \cdot \mathbf{u}_i = -\nabla V(\mathbf{x}) \cdot \mathbf{u}_i, \quad i = 1, 2, 3. \quad (3.26)$$

Ricordando (3.11) abbiamo

$$\frac{\partial \hat{V}(q_1, q_2, q_3)}{\partial q_i} = \nabla V \cdot \mathbf{u}_i, \quad i = 1, 2, 3,$$

mentre dal teorema 3.1.1 otteniamo

$$m \ddot{\mathbf{x}} \cdot \mathbf{u}_i = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i}, \quad i = 1, 2, 3.$$

¹⁶Proiettare l'equazione (3.24) lungo $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ significa determinare le componenti covarianti dell'accelerazione e della forza rispetto alla base locale.

Quindi la (3.26) si riscrive così

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = \frac{\partial \hat{V}(q_1, q_2, q_3)}{\partial q_i}, \quad i = 1, 2, 3,$$

o anche come

$$\frac{d}{dt} \left(\frac{\partial (T - \hat{V})}{\partial \dot{q}_i} \right) - \frac{\partial (T - \hat{V})}{\partial q_i} = 0, \quad i = 1, 2, 3,$$

dal momento che \hat{V} non dipende dalle \dot{q}_i . Ma allora, se introduciamo la **funzione di Lagrange**, o semplicemente **Lagrangiana**,

$$\mathcal{L}(q, \dot{q}) = T(q, \dot{q}) - \hat{V}(q), \quad (3.27)$$

la scrittura dell'equazione $m\ddot{x} = \mathbf{F}$ rispetto alla base locale diventa

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \quad i = 1, 2, 3. \quad (3.28)$$

A questo punto dovrebbe essere chiara la ricetta “automatica”: si scelgono le variabili (q_1, q_2, q_3) , dette anche **coordinate lagrangiane**, si esprime la velocità v , e quindi l'energia cinetica T e l'energia potenziale V rispetto a queste variabili, si somma T a $-V$ e si inserisce la funzione Lagrangiana così ottenuta nella (3.28). Il risultato sono le equazioni di moto nel sistema di coordinate scelto.

Notiamo infine che, ricordando l'espressione di T tramite la (3.4), con \mathbb{A} data dalla (3.3), possiamo riscrivere la Lagrangiana (3.27) nella seguente forma “compatta”

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2} \dot{q}^T \mathbb{A}(q) \dot{q} - \hat{V}(q).$$

3.3.1 Conservazione dell'energia, variabili cicliche e funzione di Routh

Si definisce **funzione di Hamilton**¹⁷, o semplicemente **Hamiltoniana**,

$$\mathcal{H}(q, \dot{q}) = \sum_{i=1}^3 \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L}. \quad (3.29)$$

¹⁷Sir William Rowan Hamilton (Dublino 1805 - Dublino 1865). Fisico matematico e professore di astronomia al Trinity College di Dublino.

Ricordando la (3.2) abbiamo¹⁸

$$\begin{aligned}
 \sum_{i=1}^3 \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} &= \sum_{i=1}^3 \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = \sum_{i=1}^3 \dot{q}_i \frac{\partial}{\partial \dot{q}_i} \left(\frac{1}{2} \sum_{h,k=1}^3 a_{h\ k}(\mathbf{q}) \dot{q}_h \dot{q}_k \right) \\
 &= \sum_{i=1}^3 \dot{q}_i \left[\frac{1}{2} \sum_{h,k=1}^3 a_{h\ k}(\mathbf{q}) \left(\frac{\partial \dot{q}_h}{\partial \dot{q}_i} \dot{q}_k + \dot{q}_h \frac{\partial \dot{q}_k}{\partial \dot{q}_i} \right) \right] \\
 &= \sum_{i=1}^3 \dot{q}_i \left[\frac{1}{2} \sum_{h,k=1}^3 a_{h\ k}(\mathbf{q}) (\delta_{hi} \dot{q}_k + \dot{q}_h \delta_{ki}) \right] \\
 &= \underbrace{\frac{1}{2} \sum_{i=1}^3 \sum_{k=1}^3 a_{i\ k}(\mathbf{q}) \dot{q}_i \dot{q}_k}_T + \underbrace{\frac{1}{2} \sum_{i=1}^3 \sum_{h=1}^3 a_{h\ i}(\mathbf{q}) \dot{q}_i \dot{q}_h}_T = 2T,
 \end{aligned}$$

per cui

$$\mathcal{H}(\mathbf{q}, \dot{\mathbf{q}}) = 2T - \mathcal{L} \stackrel{(3.27)}{=} T + \hat{V}. \quad (3.30)$$

L'Hamiltoniana di un punto materiale soggetto soltanto a forze conservative del tipo (3.17) è la somma dell'energia cinetica di quella potenziale: coincide quindi con l'**energia meccanica totale del punto materiale**.

Proposizione 3.3.1 *L'Hamiltoniana $\mathcal{H}(\mathbf{q}, \dot{\mathbf{q}})$ è costante durante il moto o, come si usa dire, è un'integrale primo del moto.*

Dim. Dobbiamo provare che $\frac{d\mathcal{H}}{dt} = 0$. Infatti dalla (3.29)

$$\begin{aligned}
 \frac{d\mathcal{H}}{dt} &= \sum_{i=1}^3 \ddot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} + \dot{q}_i \underbrace{\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right)}_{\substack{= \frac{\partial \mathcal{L}}{\partial q_i} \\ \text{v. formula} \\ (3.28)}} - \frac{d\mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{dt} \\
 &= \sum_{i=1}^3 \ddot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} + \dot{q}_i \frac{\partial \mathcal{L}}{\partial q_i} - \sum_{j=1}^3 \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \ddot{q}_j + \frac{\partial \mathcal{L}}{\partial q_j} \dot{q}_j = 0.
 \end{aligned}$$

□

¹⁸Il risultato $\sum_{i=1}^3 \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T$, altro non è che il teorema di Eulero per una funzione omogenea di n -esimo grado $f(x_1, x_2, \dots, x_k)$,

$$\sum_{i=1}^k \frac{\partial f}{\partial x_i} x_i = n f.$$

La proposizione appena illustrata mette in luce, dato il significato energetico della \mathcal{H} , una proprietà molto importante: **in particolari condizioni (punto materiale non vincolato e soggetto soltanto a forze conservative) l'energia meccanica si conserva.**

In particolare, nella sezione 5.7 daremo la definizione di Hamiltoniana per un sistema di N punti materiali, soggetto a m vincoli e vedremo che la relazione $\mathcal{H} = T + \hat{V}$, **non è vera in generale, lo è soltanto nel caso di vincoli fissi.** Analizzeremo inoltre i casi in cui $\frac{d\mathcal{H}}{dt} = 0$.

Oltre all'energia può accadere che esistano anche altre quantità che rimangono costanti durante il moto: è il caso, per esempio, delle *variabili cicliche* o *ignorabili*.

Definizione 3.3.1 *La variabile q_i si dice **ciclica**¹⁹, o **ignorabile**, se*

$$\frac{\partial \mathcal{L}}{\partial q_i} = 0,$$

cioè q_i non compare esplicitamente nella Lagrangiana \mathcal{L} , mentre compare la \dot{q}_i .

E' immediato vedere che se q_i è ciclica e il moto è governato dalle equazioni (3.28), allora

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = 0, \implies \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \text{costante}.$$

Definizione 3.3.2 *Si dice **momento coniugato** o **momento generalizzato**, la quantità*

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}. \quad (3.31)$$

Abbiamo quindi la seguente

Proposizione 3.3.2 *Se q_i è ciclica e il moto è governato dal sistema (3.28), allora p_i è costante durante il moto ed è quindi un integrale primo.*

□

Il punto adesso è il seguente: se sono presenti variabili cicliche come possono essere utilizzate per semplificare le equazioni di moto (3.28)? Una possibilità evidentemente è questa: si scrivono esplicitamente quelle equazioni del sistema (3.28) non relative alle variabili cicliche e si cerca di semplificarle sfruttando il fatto che $p_i = \text{costante}$. Questo, per esempio, è il metodo che adoteremo nel paragrafo 3.4.1, quando studieremo l'equazione che regola il moto di un punto materiale dall'origine sotto l'azione di una forza conservativa di tipo centrale.

Tuttavia questa non è la sola possibilità: si possono infatti utilizzare le variabili cicliche per "semplificare" la funzione Lagrangiana. La procedura però è delicata e merita qualche commento. Infatti, supponiamo che la q_1 sia ciclica, cioè $\mathcal{L}(\dot{q}_1, \dot{q}_2, \dot{q}_3, q_2, q_3)$. Se siamo in grado di invertire la (3.31), scriveremo

$$\dot{q}_1 = \hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3),$$

¹⁹L'espressione *coordinate cicliche* fu coniata da Helmholtz; come sinonimo si trova spesso l'espressione *variabili ignorabili* e nel testo del Lanczos si parla di *kinosthenic coordinates* che non osiamo tradurre.

dove sappiamo che p_1 è costante (e quindi è determinata dalle condizioni iniziali). Quindi, sostituendo nella \mathcal{L} otteniamo

$$\mathcal{L}(\hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3) = \hat{\mathcal{L}}(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3),$$

dove, lo rammentiamo ancora, p_1 è **una costante e non una variabile**. Parrebbe quindi di aver eliminato una variabile e, di conseguenza, di aver semplificato il sistema. Tuttavia la Lagrangiana $\hat{\mathcal{L}}$ così ottenuta non conduce alle equazioni del moto corrette: cioè

$$\frac{d}{dt} \left(\frac{\partial \hat{\mathcal{L}}}{\partial \dot{q}_i} \right) - \frac{\partial \hat{\mathcal{L}}}{\partial q_i} = 0, \quad i = 2, 3,$$

non sono le giuste equazioni di moto. Infatti

$$\begin{aligned} \frac{\partial \hat{\mathcal{L}}}{\partial \dot{q}_2} &= \frac{\partial \mathcal{L}}{\partial \hat{q}_1} \frac{\partial \hat{q}_1}{\partial \dot{q}_2} + \frac{\partial \mathcal{L}}{\partial \dot{q}_2} = p_1 \frac{\partial \hat{q}_1}{\partial \dot{q}_2} + \frac{\partial \mathcal{L}}{\partial \dot{q}_2}, \\ \frac{\partial \hat{\mathcal{L}}}{\partial q_2} &= \frac{\partial \mathcal{L}}{\partial \hat{q}_1} \frac{\partial \hat{q}_1}{\partial q_2} + \frac{\partial \mathcal{L}}{\partial q_2} = p_1 \frac{\partial \hat{q}_1}{\partial q_2} + \frac{\partial \mathcal{L}}{\partial q_2}, \end{aligned}$$

e analoghe equazioni si trovano per \dot{q}_3 , e q_3 . Di conseguenza

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial \hat{\mathcal{L}}}{\partial \dot{q}_2} \right) - \frac{\partial \hat{\mathcal{L}}}{\partial q_2} &= p_1 \left[\frac{d}{dt} \left(\frac{\partial \hat{q}_1}{\partial \dot{q}_2} \right) - \frac{\partial \hat{q}_1}{\partial q_2} \right] + \underbrace{\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_2} \right) - \frac{\partial \mathcal{L}}{\partial q_2}}_{=0} \\ &= p_1 \left[\frac{d}{dt} \left(\frac{\partial \hat{q}_1}{\partial \dot{q}_2} \right) - \frac{\partial \hat{q}_1}{\partial q_2} \right], \end{aligned}$$

ed analogamente per q_3 . Si conclude quindi che $\hat{\mathcal{L}}$ non è la Lagrangiana effettiva del sistema semplificato. La funzione che invece soddisfa le equazioni di moto, poste nella forma $\frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} - \frac{\partial}{\partial q_i} = 0$, $i = 2, 3$, è la **funzione di Routh**²⁰, o **Routhiana**

$$\begin{aligned} \mathcal{R}(\dot{q}_2, \dot{q}_3, q_2, q_3) &= \hat{\mathcal{L}}(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3) - p_1 \hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3) \\ &= \mathcal{L}(\hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3) - \\ &\quad p_1 \hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3). \end{aligned} \tag{3.32}$$

Proposizione 3.3.3 *La funzione Routhiana soddisfa le equazioni*

$$\frac{d}{dt} \left(\frac{\partial \mathcal{R}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{R}}{\partial q_i} = 0, \quad i = 2, 3. \tag{3.33}$$

Dim. La dimostrazione è una semplice riscrittura dei passaggi sopra illustrati.

□

²⁰Edward John Routh (Québec, 1831 – Cambridge, 1907) è stato un matematico inglese.

Generalizzando la procedura²¹ al caso in cui ci siano due variabili cicliche, per esempio q_1 e q_2 (cioè al caso in cui né q_1 né q_2 compaiono esplicitamente nella Lagrangiana \mathcal{L}) si procede introducendo i momenti coniugati p_i , $i = 1, 2$, tramite

$$\begin{cases} p_1 = \frac{\partial \mathcal{L}(\dot{q}_1, \dot{q}_2, \dot{q}_3, q_3)}{\partial \dot{q}_1}, \\ p_2 = \frac{\partial \mathcal{L}(\dot{q}_1, \dot{q}_2, \dot{q}_3, q_3)}{\partial \dot{q}_2}, \end{cases}$$

che si sfruttano per esprimere \dot{q}_1 e \dot{q}_2 , in funzione di p_1 , p_2 , \dot{q}_3 e q_3 ,

$$\dot{q}_i = \widehat{q}_i(p_1, p_2, \dot{q}_3, q_3), \quad i = 1, 2.$$

La funzione Routhiana è dunque

$$\mathcal{R}(\dot{q}_3, q_3) = \widehat{\mathcal{L}}(p_1, p_2, \dot{q}_3, q_3) - \sum_{i=1,2} p_i \widehat{q}_i(p_1, p_2, \dot{q}_3, q_3), \quad (3.34)$$

dove $\widehat{\mathcal{L}}(p_1, p_2, \dot{q}_3, q_3) = \mathcal{L}(\widehat{q}_1(p_1, p_2, \dot{q}_3, q_3), \widehat{q}_2(p_1, p_2, \dot{q}_3, q_3), \dot{q}_3, q_3)$.

Concludiamo la sezione osservando che la presenza di una, o più variabili cicliche non ha alcun effetto sulla funzione Hamiltoniana. Infatti, se per esempio la variabile q_1 non compare nella \mathcal{L} (cioè è ciclica), data la definizione (3.29), non comparirà neppure nell'Hamiltoniana \mathcal{H} , cioè $\mathcal{H} = \mathcal{H}(\dot{q}_1, \dot{q}_2, \dot{q}_3, q_2, q_3)$. Di conseguenza se introduciamo

$$\widehat{\mathcal{H}}(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3) = \mathcal{H}(\widehat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3), \quad (3.35)$$

è facile vedere che

$$\widehat{\mathcal{H}}(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3) = \sum_{i=2}^3 \frac{\partial \mathcal{R}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{R},$$

dove \mathcal{R} è data dalla (3.34). Svolgendo i calcoli sul membro di destra si ottiene

$$\begin{aligned} \frac{\partial \mathcal{R}}{\partial \dot{q}_i} &= \frac{\partial}{\partial \dot{q}_i} \left[\mathcal{L}(\widehat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3) - p_1 \widehat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3) \right] \\ &= \underbrace{\frac{\partial \mathcal{L}}{\partial \widehat{q}_1}}_{p_1} \frac{\partial \widehat{q}_1}{\partial \dot{q}_i} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - p_1 \frac{\partial \widehat{q}_1}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}, \quad i = 2, 3, \end{aligned} \quad (3.36)$$

²¹Per ulteriori approfondimenti si rimanda al capitolo 3, del testo: P. Biscari, C. Poggi, E. G. Virga, *Mechanics Notebook*, Liguori Editore, Napoli 1999.

da cui

$$\begin{aligned}
 \sum_{i=2}^3 \left(\frac{\partial \mathcal{R}}{\partial \dot{q}_i} \dot{q}_i \right) - \mathcal{R} & \stackrel{(3.34)}{=} \sum_{i=2}^3 \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i \right) - \mathcal{L} + \underbrace{p_1}_{\frac{\partial \mathcal{L}}{\partial \dot{q}_1}} \dot{q}_1 \\
 & = \sum_{i=1}^3 \frac{\partial \mathcal{L} \left(\hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3 \right)}{\partial \dot{q}_i} \dot{q}_i - \\
 & \quad \mathcal{L} \left(\hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3 \right) \\
 & = \mathcal{H} \left(\hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3), \dot{q}_2, \dot{q}_3, q_2, q_3 \right) = \hat{\mathcal{H}}(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3).
 \end{aligned}$$

Quindi per ottenere una versione “semplificata” dell’energia meccanica è sufficiente inserire direttamente l’espressione di $q_1 = \hat{q}_1(p_1, \dot{q}_2, \dot{q}_3, q_2, q_3)$ nella Hamiltoniana \mathcal{H} , come mostrato dalla (3.35).

La generalizzazione al caso di N punti materiali sarà illustrata nella sezione 5.6, dove considereremo l variabili q , cioè (q_1, \dots, q_l) , di cui $v \leq l$, cicliche.

Nota 3.3.1 Come già osservato la sostituzione di una variabile ciclica nella funzione di Lagrange conduce ad una $\hat{\mathcal{L}}$ non “corretta”, cioè **una Lagrangiana che genera equazioni di moto non corrette**. Per rendere ancor più evidente questo fatto, tentiamo di sfruttare la conservazione dell’energia meccanica. Ricordando la (3.30) e la proposizione 3.3.1, abbiamo $\mathcal{H} = T + \hat{V} = E$, con E costante. Pertanto, sfruttando tale relazione nella Lagrangiana, potremmo scrivere

$$\mathcal{L} = T - \hat{V} = T - (E - T) = 2T - E,$$

e le equazioni di Lagrange (3.28) si ridurrebbero a

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = 0, \quad i = 1, 2, 3,$$

ovviamente **non corrette** (basti pensare al semplice oscillatore armonico). Se, nell’ambito del formalismo lagrangiano, si vuole tener conto del fatto che $T + \hat{V} = \text{costante}$, bisogna introdurre una opportuna funzione di Routh. Tale problema verrà discusso nella sezione 8.5: l’ambito naturale dove sviluppare questo tema è infatti quello dei metodi variazionali.

3.4 Il moto centrale

Vediamo ora un’applicazione importante delle equazioni di Lagrange: il moto di un punto materiale di massa m sotto la sola azione di una forza centrale del tipo

$$\mathbf{F} = f(\|P - O\|) \mathbf{vers}(P - O), \quad \Leftrightarrow \quad \mathbf{F} = f(r) \mathbf{u}_r, \quad (3.37)$$

come quella già vista nell’esempio 3.2.4. Introduciamo il sistema di coordinate sferiche (r, θ, φ) (ovviamente centrate in O). Sappiamo che una forza della forma (3.37) è

conservativa e che la sua energia potenziale è data dalla (3.23). Per scrivere la funzione Lagrangiana basta quindi ricordare l'esempio 3.1.3, dove si trova l'espressione esplicita dell'energia cinetica del punto materiale (v. formula (3.9)). La funzione di Lagrange è quindi

$$\mathcal{L} = T - V(r) = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2 \right) + \int_{r_0}^r f(s) ds.$$

Si noti che la φ è una coordinata ciclica (v. definizione 3.3.1). Questo implica, in virtù della proposizione 3.3.2, che il momento coniugato

$$p_\varphi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = m r^2 \sin^2 \theta \dot{\varphi},$$

è costante durante il moto, ovvero un suo integrale primo.

Per integrare le equazioni è bene scegliere in modo opportuno il sistema di riferimento (si noti che le equazioni che abbiamo scritto valgono per qualsiasi riferimento purché abbia origine in O ; ci resta quindi la libertà di ruotare gli assi come più ci fa comodo). Posizioniamo quindi l'asse delle x in modo che P si trovi, all'istante iniziale, su questo asse. Ciò significa che al tempo iniziale (che possiamo prendere, senza perdere in generalità, come $t = 0$) si ha $\varphi(0) = 0$ e $\theta(0) = \frac{\pi}{2}$ (la prima condizione iniziale è ovvia, per la seconda si osservi che l'asse delle x appartiene al piano $z = 0$ comunque siano posizionati gli assi). Ora ruotiamo il sistema di riferimento attorno all'asse x finché la velocità iniziale v_0 si trovi nel piano $z = 0$. Questa scelta è univoca se la velocità iniziale non è parallela al vettore $P_0 - O$ che individua la posizione iniziale, $P(t=0) = P_0$, ovvero se $v_0 \wedge (P_0 - O) \neq 0$. Questa è detta **condizione di non degenerazione** del problema del moto in un campo centrale. È facile vedere che se questa condizione è violata, il moto si riduce a un moto unidimensionale sulla retta individuata dai punti P_0 e O . Notiamo poi che la condizione $P_0 \equiv O$ deve essere esclusa in molti casi perché la forza non è definita in O . Inoltre se $P_0 \equiv O$, il sistema di coordinate lagrangiane non è definito²².

Con questa scelta degli assi si ha che, nel punto P_0 , $\mathbf{u}_r = \mathbf{e}_1$, $\mathbf{u}_\theta = -\mathbf{e}_3$, $\mathbf{u}_\varphi = \mathbf{e}_2$, dove \mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3 , sono i versori degli assi x , y e z , rispettivamente. Di conseguenza abbiamo $\dot{\theta}(0) = 0$.

Vogliamo ora far vedere che $\theta(t) \equiv \frac{\pi}{2}$ è compatibile con le equazioni di moto. Scriviamo infatti l'equazione di Lagrange per la variabile θ

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) - \frac{\partial \mathcal{L}}{\partial \theta} = m \ddot{\theta} + 2mr \dot{r} \dot{\theta}^2 - mr^2 \dot{\varphi}^2 \cos \theta \sin \theta = 0,$$

ed osserviamo che è sempre soddisfatta prendendo $\theta(t) = \frac{\pi}{2}$ costantemente²³, qualunque siano le funzioni (incognite) $r(t)$ e $\varphi(t)$. Ne segue che il moto avviene nel piano

²²Si ricordi l'esempio 1.4.2, $\det \mathbb{J} = -r^2 \sin \theta$, si annulla per $r = 0$.

²³Il fatto che l'angolo $\theta(t)$ sia sempre uguale a $\frac{\pi}{2}$ implica che il moto avviene nel piano $z = 0$, con la nostra scelta degli assi. La cosa fisicamente significativa è che il moto avviene su un piano, non che questo sia il piano $z = 0$! Ma questo potevamo dedurlo a priori senza far intervenire alcun sistema di coordinate: infatti poiché il moto è centrale, abbiamo sempre $\ddot{\mathbf{x}} \wedge (P - O) = 0$ (la forza, e quindi l'accelerazione, è diretta come la congiungente di P con O) da cui si deduce che $\mathbf{v} \wedge (P - O) = \mathbf{L}$, è un vettore costante. Di conseguenza il moto avviene nel piano (fisso) passante per O e perpendicolare a \mathbf{L} .

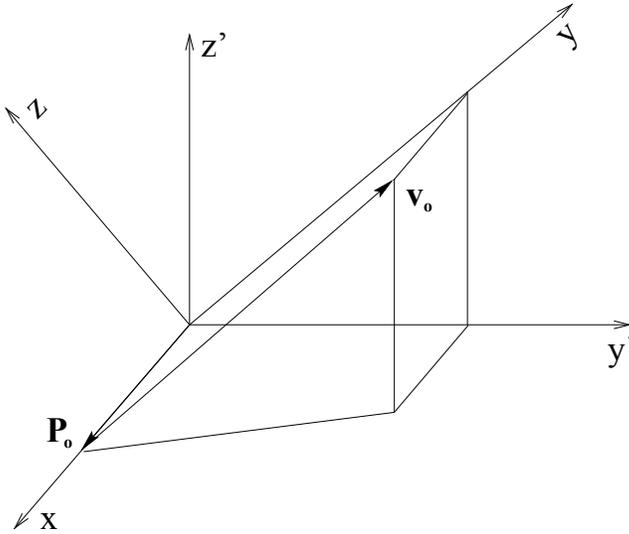


Figura 3.1: Posizionamento del piano coordinato

$\theta = \pi/2$, ovvero dal piano determinato dal centro del moto O , dalla posizione iniziale P_0 e dalla direzione della velocità iniziale v_0 . Inoltre, posto $\theta = \pi/2$ nelle altre due equazioni, otteniamo un sistema di due equazioni differenziali per le sole r e φ , che possiamo risolvere con le opportune condizioni iniziali, ottenendo la soluzione del sistema nella forma $r = r(t)$, $\theta = \pi/2$ e $\varphi = \varphi(t)$.

Riscriviamo quindi la Lagrangiana del problema tenendo conto del fatto che $\theta \equiv \pi/2$ (cioè del fatto che il moto avviene su di un piano)

$$\mathcal{L} = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\varphi}^2) - V(r).$$

Ricordando che φ è ciclica, cioè che

$$p_\varphi = m r^2 \dot{\varphi} = \text{costante}, \quad (3.38)$$

possiamo introdurre la quantità

$$\dot{A} = \frac{1}{2} r^2 \dot{\varphi}, \quad \Leftrightarrow \quad p_\varphi = 2m\dot{A}, \quad (3.39)$$

detta **velocità areolare**, e immediatamente dedurre che

$$m\dot{A} = \text{cost.} \quad \implies \quad \dot{A} = \text{cost.}$$

Tale risultato, cioè la conservazione della quantità \dot{A} , si può interpretare osservando che $\frac{1}{2} r^2 \dot{\varphi}$ rappresenta la variazione rispetto al tempo dell'*area spazzata dal raggio vettore* $P - O$, ovvero dell'area del settore di piano compreso tra i vettori $P(t_1) - O$ e $P(t_0) - O$ e la traiettoria sul piano del punto $P(t)$ per $t \in [t_0, t_1]$. Abbiamo così

dimostrato la **seconda legge di Keplero**²⁴ (che vale per qualsiasi forza centrale del tipo (3.37)):

Il raggio vettore spazza aree uguali in tempi uguali.

Nota 3.4.1 Ricordando la definizione (3.29), o meglio la (3.30), calcoliamo l'Hamiltoniana del sistema

$$\mathcal{H} = T + V = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\varphi}^2) + V(r).$$

Adesso, tenendo conto del fatto che la φ è ciclica, l'Hamiltoniana "semplificata" è

$$\hat{\mathcal{H}} = \frac{m \dot{r}^2}{2} + m \frac{2\dot{A}^2}{r^2} + V(r), \quad (3.40)$$

ovvero $\hat{\mathcal{H}} = \frac{m \dot{r}^2}{2} + \frac{1}{2r^2} \frac{p_\varphi^2}{m} + V(r).$

Esempio 3.4.1 Vediamo come si giunge alla conservazione della velocità areolare per punti materiali in campi centrali tramite il metodo del momento angolare del punto materiale rispetto ad O , definito come

$$\mathbf{L}_O = \mathbf{x} \wedge m \dot{\mathbf{x}}.$$

Se l'unica forza agente sul punto materiale P è di tipo centrale come la (3.37), è facile provare che $\dot{\mathbf{L}}_O = 0$. Infatti

$$\dot{\mathbf{L}}_O = \underbrace{\dot{\mathbf{x}} \wedge m \dot{\mathbf{x}}}_{=0} + \mathbf{x} \wedge m \ddot{\mathbf{x}} \stackrel{(3.37)}{=} \mathbf{x} \wedge f(r) \mathbf{u}_r = 0,$$

dal momento che, in coordinate sferiche, $\mathbf{x} = r \mathbf{u}_r$. Pertanto $\mathbf{L}_O(t) = \mathbf{L}_O(0)$, $\forall t \geq 0$. Ci sono quindi due possibilità:

1. $\mathbf{L}_O(0) \neq 0$.
2. $\mathbf{L}_O(0) = 0$.

Nel primo caso il punto materiale non transita mai per l'origine (se infatti P transitasse per O si avrebbe $\mathbf{L}_O(0) = 0$). Il moto inoltre avviene su di un piano che contiene il punto O , dal momento che, $\forall t \geq 0$, $\mathbf{x} \cdot \mathbf{L}_O(t) = \mathbf{x} \cdot \mathbf{L}_O(0) = 0$. Ma $\mathbf{x} \cdot \mathbf{L}_O(0) = 0$, scritta in coordinate cartesiane, è proprio l'equazione di un piano che passa per l'origine O , ed è perpendicolare al vettore $\mathbf{L}_O(0)$. Quindi, operando in coordinate sferiche, orientiamo il sistema di riferimento affinché il piano del moto coincida con $\theta = \pi/2$. Ricordando la (3.8), scriviamo il momento angolare in coordinate sferiche

$$\mathbf{L}_O(0) = m r \boldsymbol{\kappa}_r \wedge (\dot{r} \boldsymbol{\kappa}_r + r \dot{\varphi} \sin \theta \boldsymbol{\kappa}_\varphi) = m r^2 \dot{\varphi} \boldsymbol{\kappa}_r \wedge \boldsymbol{\kappa}_\varphi,$$

e quindi

$$\|\mathbf{L}_O\| = \text{costante}, \implies m r^2 \dot{\varphi} = \text{costante},$$

²⁴Johannes von Kepler (Weil der Stadt, 1571 – Ratisbona, 1630) fu un astronomo, matematico e musicista tedesco.

che è esattamente la (3.38).

Nel secondo caso abbiamo che $\mathbf{x}(t)$ e $\dot{\mathbf{x}}(t)$ sono costantemente paralleli. Utilizzando ancora le coordinate sferiche si ottien

$$0 = mr\mathbf{u}_r \wedge (\dot{r}\mathbf{u}_r + \dot{\varphi}\mathbf{u}_\varphi + \dot{\theta}\mathbf{u}_\theta), \implies \begin{cases} mr\dot{\varphi}\mathbf{u}_r \wedge \mathbf{u}_\varphi = 0, & \implies \dot{\varphi} = 0, \\ mr\dot{\theta}\mathbf{u}_r \wedge \mathbf{u}_\theta = 0, & \implies \dot{\theta} = 0, \end{cases}$$

e quindi $\varphi(t) = \varphi_0$, $\theta(t) = \theta_0$. Il moto avviene quindi sulla retta che passa per l'origine ed è caratterizzata dagli angoli φ_0 , e θ_0 .

3.4.1 L'equazione per r

Scriviamo ora l'equazione per la variabile $r(t)$ tenendo conto del fatto che $\theta \equiv \pi/2$,

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{r}} \right) - \frac{\partial \mathcal{L}}{\partial r} = m\ddot{r} - mr\dot{\varphi}^2 + \frac{\partial V(r)}{\partial r} = 0. \quad (3.41)$$

Possiamo ora riscrivere la (3.41) eliminando la dipendenza da $\dot{\varphi}$ tramite la (3.39). Si ha quindi

$$m\ddot{r} - m \frac{(2\dot{A})^2}{r^3} + \frac{\partial V(r)}{\partial r} = 0, \implies m\ddot{r} = -\frac{\partial}{\partial r} \left[m \frac{(2\dot{A})^2}{2r^2} + V(r) \right] \quad (3.42)$$

che ci dice che, in un moto centrale, il raggio evolve come la posizione di un punto materiale che si muove su una retta sotto l'azione di una energia potenziale (spesso detta **energia potenziale efficace**) data da

$$\tilde{V}(r) = m \frac{2\dot{A}^2}{r^2} + V(r). \quad (3.43)$$

Possiamo quindi studiare l'evoluzione di $r(t)$ usando le tecniche sviluppate per i moti unidimensionali nella sezione 2.6 del capitolo 2. Infatti moltiplicando la (3.42) per \dot{r} , si ottiene

$$\frac{d}{dt} \left[\frac{m}{2} \dot{r}^2 + \tilde{V}(r) \right] = 0, \implies \frac{m}{2} \dot{r}^2 + \tilde{V}(r) = E, \quad (3.44)$$

dove E è l'energia, che si mantiene costante durante il moto. Notiamo che tale risultato discende dalla proposizione 3.3.1, dal momento che $\left[\frac{m}{2} \dot{r}^2 + \tilde{V}(r) \right]$, altro non è che l'Hamiltoniana "semplificata" $\hat{\mathcal{H}}$ data dalla (3.40).

Nota 3.4.2 Possiamo ricavare l'equazione (3.42) anche col metodo della funzione di Routh illustrato nella sezione 3.3.1. Dalla (3.38) abbiamo

$$\dot{\varphi}(p_\varphi, r) = \frac{p_\varphi}{m r^2},$$

per cui

$$\begin{aligned} \mathcal{R}(\dot{r}, r) &= \underbrace{\mathcal{L}(\dot{\varphi}(p_\varphi, r), \dot{r}, r)}_{\hat{\mathcal{L}}(p_\varphi, \dot{r}, r)} - p_\varphi \dot{\varphi} \\ &= \frac{m}{2} \dot{r}^2 + \frac{m}{2} r^2 \left(\frac{p_\varphi}{m r^2} \right)^2 - V(r) - \frac{p_\varphi^2}{m r^2} \\ &= \frac{m}{2} \dot{r}^2 - \frac{1}{2r^2} \frac{p_\varphi^2}{m} - V(r). \end{aligned}$$

Dalla formula (3.33) deduciamo

$$m\ddot{r} - \frac{1}{r^3} \frac{p_\varphi^2}{m} + \frac{\partial V}{\partial r} = 0, \Leftrightarrow m\ddot{r} - m \frac{(2\dot{A})^2}{r^3} + \frac{\partial V}{\partial r} = 0,$$

che è proprio l'equazione (3.42).

3.4.2 Il problema di Keplero

Il problema di Keplero è il moto di una massa m l'azione di una forza centrale di questo tipo

$$\mathbf{F} = -\frac{mK}{\|P-O\|^2} \text{vers}(P-O), \quad \text{con } K = \text{costante} > 0.$$

che scritta in coordinate sferiche diventa

$$\mathbf{F} = -\frac{mK}{r^2} \mathbf{u}_r.$$

L'energia potenziale è dunque (si veda l'esempio 3.2.4)

$$V(r) = -\frac{mK}{r}. \quad (3.45)$$

L'energia potenziale efficace per r è quindi data da

$$\tilde{V}(r) = m \frac{2\dot{A}^2}{r^2} - \frac{mK}{r}, \quad (3.46)$$

il cui grafico (v. figura 3.2) presenta un asintoto verticale in $r = 0$ ($\lim_{r \rightarrow 0} \tilde{V} = +\infty$), ha limite 0^- per $r \rightarrow +\infty$ ed ha un minimo negativo per $r = \frac{(2\dot{A})^2}{K}$, con valore dell'energia

$$E_0 = \tilde{V}\left(\frac{(2\dot{A})^2}{K}\right) = -\frac{1}{2} \frac{mK^2}{(2\dot{A})^2} < 0. \quad (3.47)$$

Si noti che il valore E_0 dipende da \dot{A}

Al livello di energia minimo corrisponde un'orbita circolare, in quanto r resta costante durante il moto. La risultante soluzione $(r(t), \varphi(t))$ del problema di Keplero è dunque periodica nel tempo poiché, dalla (3.38), anche $\dot{\varphi} = \text{cost.}$, cioè il punto materiale ruota con velocità angolare costante attorno al centro d'attrazione O .

Ai livelli di energia E compresi tra E_0 e 0 corrispondono intervalli $r_m \leq r(t) \leq r_M$, limitati di soluzioni della disequaglianza $\tilde{V} < E$, quindi soluzioni $r(t)$ periodiche rispetto a t . Ricordiamo poi che, fissato $E \in (E_0, 0)$, i limiti r_m, r_M dell'intervallo dove varia $r(t)$ si trovano risolvendo l'equazione algebrica²⁵

$$\tilde{V}(r) = E, \implies E = \frac{2m\dot{A}^2}{r^2} - \frac{mK}{r}, \quad (3.48)$$

²⁵Nota che $E < 0$ e quindi le due radici della (3.48) sono positive.

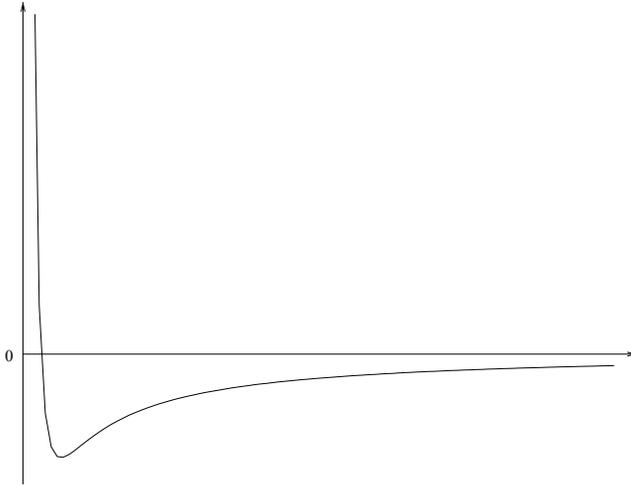


Figura 3.2: Grafico della funzione energia potenziale efficace $\tilde{V}(r)$ per il problema di Keplero.

mentre il periodo di “oscillazione” di $r(t)$ (distanza dal centro di attrazione), che è dato da (v. sezione 2.6.1, formula (2.42))

$$T = 2 \int_{r_m}^{r_M} \frac{dr}{\sqrt{2 \frac{E}{m} - \frac{(2\dot{A})^2}{r^2} + \frac{2K}{r}}}.$$

Ora, il fatto che $r(t)$ sia periodica non è sufficiente per affermare che **la soluzione** $(r(t), \varphi(t))$ **dell'intero problema sia periodica**. Infatti durante il periodo l'angolo subisce una variazione $\Delta\varphi$ che può essere calcolata a partire dalla conservazione della velocità angolare espressa dalla (3.39)

$$\dot{\varphi}(t) = \frac{2\dot{A}}{r^2(t)}, \Rightarrow \underbrace{\varphi(T) - \varphi(0)}_{\Delta\varphi} = \int_0^T \frac{2\dot{A}}{r^2(t)} dt. \quad (3.49)$$

Il moto risulterà periodico solo se l'incremento $\Delta\varphi$ subito dall'angolo nel singolo periodo T (che lo ricordiamo ancora è il periodo di oscillazione della distanza r dal centro) è *un multiplo razionale di 2π* . Infatti se

$$\Delta\varphi = \frac{2k}{N}\pi,$$

allora dopo un tempo NT (cioè dopo N periodi) abbiamo $r(NT) = r(0)$ e

$$\varphi(NT) = \varphi(0) + N\Delta\varphi = \varphi(0) + 2k\pi,$$

ovvero il punto materiale ha compiuto k giri attorno ad O . Il punto si trova nella stessa posizione (e con la stessa velocità) che aveva al tempo $t = 0$.

L'integrale (3.49) tramite cui si calcola $\Delta\varphi$ ha il difetto di essere un integrale nella variabile temporale, e quindi per calcolarlo si deve conoscere l'espressione esplicita di $r(t)$, che però non abbiamo. Quindi, anziché $\Delta\varphi$ calcoliamo $\Delta\varphi/2$

$$\frac{\Delta\varphi}{2} = \int_0^{T/2} \frac{2\dot{A}}{r^2(t)} dt, \quad (3.50)$$

e nell'integrale e operiamo la seguente sostituzione di variabile

$$\rho = r(t), \Rightarrow d\rho = \dot{r}(t) dt.$$

Se sfruttiamo la (3.44) per esprimere $\dot{r}dt$ in termini di $r(t)$, e di conseguenza in termini di ρ

$$d\rho \stackrel{(3.44)}{=} \underbrace{\sqrt{\frac{2}{m} (E - \tilde{V}(\rho))}}_{\dot{r}} dt, \Rightarrow dt = \frac{d\rho}{\sqrt{\frac{2}{m} (E - \tilde{V}(\rho))}}.$$

la (3.50) può essere così riscritta²⁶

$$\frac{\Delta\varphi}{2} = \int_{r_m}^{r_M} \frac{2\dot{A}}{\rho^2} \frac{d\rho}{\sqrt{\frac{2}{m} (E - \tilde{V}(\rho))}} = \int_{r_m}^{r_M} \frac{2\dot{A}}{\rho^2} \frac{d\rho}{\sqrt{\frac{2E}{m} - \frac{(2\dot{A})^2}{\rho^2} + \frac{2K}{\rho}}}. \quad (3.51)$$

L'integrale (3.51) si calcola facilmente cambiando variabile $z = \frac{2\dot{A}}{\rho}$, da cui si ottiene

$$\Delta\varphi = 2 \int_{\frac{2\dot{A}}{r_M}}^{\frac{2\dot{A}}{r_m}} \frac{dz}{\sqrt{\frac{2E}{m} - z^2 + \frac{K}{\dot{A}}z}}, \quad (3.52)$$

Osserviamo che i limiti di integrazione (ovvero $\frac{2\dot{A}}{r_M}$ e $\frac{2\dot{A}}{r_m}$) sono le due radici del trinomio che compare a denominatore. Infatti, siccome r_M e r_m sono le due radici della (3.48), abbiamo

$$2 \left(\frac{E}{m} - \frac{2\dot{A}^2}{r_m^2} - \frac{K}{r_m} \right) = 0, \Rightarrow \frac{2E}{m} - \left(\frac{2\dot{A}}{r_m} \right)^2 - \frac{K}{\dot{A}} \left(\frac{2\dot{A}}{r_m} \right) = 0,$$

e analogamente

$$2 \left(\frac{E}{m} - \frac{2\dot{A}^2}{r_M^2} - \frac{K}{r_M} \right) = 0, \Rightarrow \frac{2E}{m} - \left(\frac{2\dot{A}}{r_M} \right)^2 - \frac{K}{\dot{A}} \left(\frac{2\dot{A}}{r_M} \right) = 0.$$

Pertanto, ponendo $z_M = \frac{2\dot{A}}{r_M}$, e $z_m = \frac{2\dot{A}}{r_m}$, con $z_M < z_m$, possiamo così riscrivere la (3.52)

$$\Delta\varphi = 2 \int_{z_M}^{z_m} \frac{dz}{\sqrt{(z_m - z)(z - z_M)}}.$$

²⁶Ovviamente $r(0) = r_m$, mentre $r(T/2) = r_M$.

Ora, come già osservato nella sezione 2.9, $\int_{z_M}^{z_m} \frac{dz}{\sqrt{z_m - z} \sqrt{z - z_M}}$, è un integrale notevole e vale π . Quindi

$$\Delta\varphi = 2\pi, \quad (3.53)$$

Possiamo quindi affermare che il moto, per valori negativi di E , è un moto periodico, e che la traiettoria è una curva semplice (senza auto-intersezioni) che contiene il centro del moto al suo interno²⁷.

Esempio 3.4.2 Consideriamo il caso di una forza di richiamo elastica del tipo $\mathbf{F} = -kr\mathbf{u}_r$, dove k è la costante elastica. L'energia potenziale è $V(r) = -\frac{k}{2}r^2$, mentre, ricordando la (3.43), il potenziale efficace

$$\tilde{V}(r) = m\frac{2\dot{A}^2}{r^2} + m\frac{\omega^2}{2}r^2, \quad \text{dove, al solito, } \omega^2 = \frac{k}{m}. \quad (3.54)$$

In figura 3.3 è riportato l'andamento di $\tilde{V}(r)$. Si può notare che $\tilde{V}(r)$ ha un asintoto verticale per $r = 0$, e presenta un minimo in $r_e = \sqrt[4]{\frac{(2\dot{A})^2}{\omega^2}}$, dove vale $V_e = 2|\dot{A}|\omega$.

Quindi se $E > V_e$ il moto in r è periodico e si svolge all'interno dell'intervallo $[r_m, r_M]$, dove r_m e r_M sono le due soluzioni di

$$2(E - \tilde{V}(r)) = 0, \implies \frac{2E}{m} - \frac{(2\dot{A})^2}{r^2} - \omega^2 r^2 = 0, \quad (3.55)$$

cioè

$$r_m = \sqrt{\frac{E}{m\omega^2} \left[1 - \sqrt{1 - \left(\frac{2\dot{A}\omega}{E/m}\right)^2} \right]},$$

$$r_M = \sqrt{\frac{E}{m\omega^2} \left[1 + \sqrt{1 - \left(\frac{2\dot{A}\omega}{E/m}\right)^2} \right]}.$$

²⁷Nel dimostrare questo risultato abbiamo usato lo stesso argomento che ci aveva permesso di dimostrare l'isocronia del moto armonico. In quel caso avevamo anche mostrato un risultato di "unicità" dei potenziali isocroni: il solo potenziale (simmetrico) isocrono è quello elastico.

Anche per i moti centrali vale un teorema analogo, noto come **Teorema di Bertrand** che afferma che i soli potenziali centrali che ammettano moti periodici per ogni condizione iniziale sono quello **elastico** e quello **kepleriano** (in questo caso soltanto per $E < 0$). Per la dimostrazione, se ci limitiamo ad energie potenziali della forma $V(r) = -k r^\alpha$, $\alpha \neq 0$ e $k\alpha < 0$ (per avere una forza attrattiva) oppure $V(r) = -k \ln r$, $k > 0$, basta osservare che il caso $\alpha = -1$ e $\alpha = 2$ sono i soli in cui si possa riportare l'integrale

$$\Delta\varphi = 2 \int_{r_m}^{r_M} \frac{2\dot{A}}{r^2} \frac{dr}{\sqrt{2E/m - \frac{(2\dot{A})^2}{r^2} - \frac{2}{m}V(r)}}$$

alla forma (3.53) tramite un opportuno cambiamento di variabile.

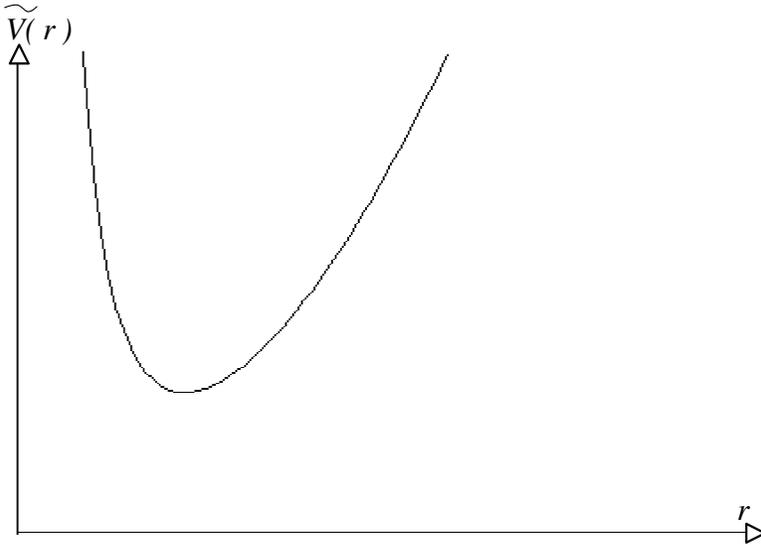


Figura 3.3: Grafico di $\tilde{V}(r)$, data dalla (3.54).

Calcoliamo adesso, sfruttando la (3.51), la variazione dell'angolo φ in un semiperiodo

$$\frac{\Delta\varphi}{2} = \int_{r_m}^{r_M} \frac{2\dot{A}}{\rho^2} \frac{d\rho}{\sqrt{\frac{2E}{m} - \frac{(2\dot{A})^2}{\rho^2} - \omega^2 \rho^2}}.$$

Introducendo al solito $z = \frac{2\dot{A}}{\rho}$, otteniamo

$$\frac{\Delta\varphi}{2} = \int_{\frac{2\dot{A}}{r_M}}^{\frac{2\dot{A}}{r_m}} \frac{z dz}{\sqrt{\frac{2E}{m} z^2 - z^4 - \omega^2 (2\dot{A})^2}},$$

che a sua volta, ponendo $u = z^2/2$, si riscrive così

$$\frac{\Delta\varphi}{2} = \int_{u_M}^{u_m} \frac{du}{\sqrt{\frac{4E}{m} u - 4u^2 - \omega^2 (2\dot{A})^2}},$$

dove $u_M = \frac{1}{2} \left(\frac{2\dot{A}}{r_M} \right)^2$, $u_m = \frac{1}{2} \left(\frac{2\dot{A}}{r_m} \right)^2$. E' facile vedere che quest'ultimi sono le radici del trinomio che appare sotto radice. Infatti sostituendo si ottiene proprio la

(3.55). Quindi possiamo scrivere

$$\Delta\varphi = 2 \underbrace{\int_{u_M}^{u_m} \frac{du}{\sqrt{(u-u_m)(u_M-u)}}}_{\pi} = 2\pi,$$

e dunque concludere che il moto del punto materiale è periodico.

3.4.3 L'orbita del problema di Keplero

Vogliamo adesso determinare l'equazione dell'orbita del punto materiale. Cerchiamo quindi di esprimere r in funzione dell'angolo φ , cioè $r = r(\varphi)$. La conservazione della velocità areolare ci permette di eliminare il tempo dalle due equazioni del moto (la (3.42) e la (3.39)). Infatti abbiamo che $\dot{\varphi}$ è **sempre diversa da zero e quindi la φ è una funzione monotona del tempo**. Possiamo quindi invertire $\varphi = \varphi(t)$ ed esprimere il tempo, e tutte le funzioni del tempo, e quindi la stessa $r(t)$, come funzioni di φ . Cominciando con r , cerchiamo $r = r(t(\varphi))$

$$\frac{dr}{dt} = \frac{dr}{d\varphi} \frac{d\varphi}{dt} \stackrel{(3.39)}{=} \frac{2\dot{A}}{r^2} \frac{dr}{d\varphi} = -2\dot{A} \frac{d}{d\varphi} \left(\frac{1}{r} \right),$$

e

$$\frac{d^2r}{dt^2} = \frac{d}{dt} \left(\frac{dr}{dt} \right) = \frac{d}{d\varphi} \left(-2\dot{A} \frac{d}{d\varphi} \left(\frac{1}{r} \right) \right) = -\frac{(2\dot{A})^2}{r^2} \frac{d^2}{d\varphi^2} \left(\frac{1}{r} \right).$$

Sostituendo nella (3.42) dove $V(r)$ è data dalla (3.45), otteniamo, dopo ovvie semplificazioni,

$$\frac{d^2}{d\varphi^2} \left(\frac{1}{r} \right) + \frac{1}{r} - \frac{K}{(2\dot{A})^2} = 0, \quad (3.56)$$

cioè

$$\frac{d^2}{d\varphi^2} \left(\frac{1}{r} - \frac{K}{(2\dot{A})^2} \right) = - \left(\frac{1}{r} - \frac{K}{(2\dot{A})^2} \right),$$

che, posto

$$y(\varphi) = \frac{1}{r(\varphi)} - \frac{K}{(2\dot{A})^2},$$

è un'equazione del tipo $y'' = -y$: l'equazione dell'oscillatore armonico, la cui generica soluzione è

$$y(\varphi) = y_0 \cos(\varphi + \beta).$$

In particolare, ponendo $y_0 = \frac{K}{(2\dot{A})^2} e$, avremo

$$\frac{1}{r(\varphi)} = \frac{K}{(2\dot{A})^2} [1 + e \cos(\varphi + \beta)], \quad (3.57)$$

dove $e > 0$ e $\beta \in [0, 2\pi)$ sono due costanti determinate dalle condizioni iniziali $r(t=0) = r_0$ e $\dot{r}(0) = \dot{r}_0$. Per quanto riguarda φ ricordiamo che abbiamo già

scelto $\varphi(t=0) = 0$, quindi $\varphi = 0$, sarà l'angolo iniziale. Inserendo quindi le condizioni iniziali $\varphi_0 = 0$, r_0 , \dot{r}_0 e $\dot{A} = \frac{r_0^2 \dot{\varphi}_0}{2}$ e la definizione di E data dalla (3.48), cioè $E = \frac{2m\dot{A}^2}{r_0^2} - \frac{mK}{r_0}$, otteniamo

$$\tan \beta = \frac{\dot{r}_0}{2\dot{A}} \frac{1}{\frac{1}{r_0} - \frac{K}{(2\dot{A})^2}}, \quad (3.58)$$

$$e = \sqrt{1 + 2 \frac{E}{m} \frac{(2\dot{A})^2}{K^2}}. \quad (3.59)$$

Osserviamo poi che possiamo sempre orientare l'asse x in maniera da avere $\beta = 0$. Pertanto, ponendo

$$\frac{1}{p} = \frac{K}{(2\dot{A})^2}, \quad (3.60)$$

l'equazione dell'orbita (3.57) si riscrive così

$$\frac{p}{r(\varphi)} = 1 + e \cos \varphi. \quad (3.61)$$

Al variare di $e > 0$, abbiamo tre possibili casi.

Primo caso. $0 < e < 1$.

Riscrivendo la (3.61) otteniamo

$$r(\varphi) = \frac{p}{1 + e \cos \varphi}, \quad (3.62)$$

che implica $r_m \leq r \leq r_M$, dove

$$\frac{r_m}{p} = \frac{1}{1+e}, \quad \frac{r_M}{p} = \frac{1}{1-e}. \quad (3.63)$$

E' facile convincersi che, ricordando la sezione 3.4.2, questo caso corrisponde al caso $-|E_0| < E < 0$, dove E_0 è dato dalla (3.47). Infatti dalla (3.59) e dalla (3.47) otteniamo²⁸

$$e = \sqrt{1 - \frac{|E|}{|E_0|}} < 1. \quad (3.64)$$

Notiamo inoltre che risolvendo la (3.48), si ottengono r_m e r_M , cioè

$$\frac{1}{r_m} = \frac{K}{(2\dot{A})^2} \left[1 + \sqrt{1 - 2 \frac{|E|}{m} \frac{(2\dot{A})^2}{K^2}} \right] \stackrel{(3.47)}{=} \frac{1}{p} \left[1 + \sqrt{1 - \frac{|E|}{|E_0|}} \right] \stackrel{(3.64)}{=} \frac{1}{p} (1 - e), \quad (3.60)$$

$$\frac{1}{r_M} = \frac{K}{(2\dot{A})^2} \left[1 - \sqrt{1 - 2 \frac{|E|}{m} \frac{(2\dot{A})^2}{K^2}} \right] \stackrel{(3.47)}{=} \frac{1}{p} \left[1 - \sqrt{1 - \frac{|E|}{|E_0|}} \right] \stackrel{(3.64)}{=} \frac{1}{p} (1 + e), \quad (3.60)$$

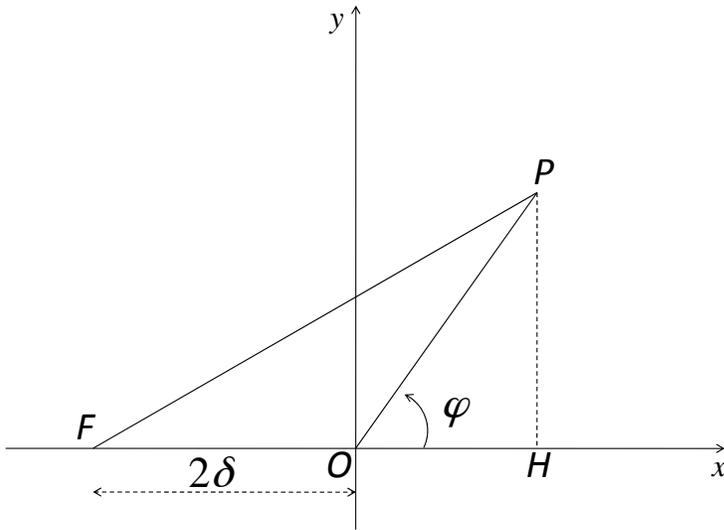


Figura 3.4: Ellisse in coordinate polari.

che coincidono esattamente con (3.63).

Vediamo adesso che l'equazione (3.62), nel caso $e < 1$, rappresenta un'ellisse in cui il punto O , centro attrattore, occupa uno dei due fuochi. Infatti, osservando la figura 3.4, l'ellisse con fuochi F e O è il luogo dei punti P tali che²⁹

$$\overline{PF} + \overline{PO} = 2d, \quad (3.65)$$

dove d è un parametro fissato maggiore di δ , distanza fra i fuochi. Ora dalla figura 3.4 deduciamo

$$\begin{aligned} \overline{PF} &= \sqrt{(\overline{FO} + \overline{OH})^2 + \overline{PH}^2} = \sqrt{(2\delta + r \cos \varphi)^2 + r^2 \sin^2 \varphi} \\ &= \sqrt{4\delta^2 + 4\delta r \cos \varphi + r^2}. \end{aligned}$$

Quindi sostituendo nella (3.65)

$$\sqrt{4\delta^2 + 4\delta r \cos \varphi + r^2} = 2d - r,$$

ed elevando al quadrato si ha

$$r(d + \delta \cos \varphi) = d^2 - \delta^2, \Rightarrow r = \frac{p}{1 + e \cos \varphi},$$

²⁸Si ricordi che $E < 0$.

²⁹Ovviamente con questo calcolo otteniamo un'ellisse il cui asse maggiore coincide con l'asse polare. Quindi l'afelio, ovvero il punto più distante dall'origine, corrisponde all'angolo $\varphi = \pi$, come nella figura, oppure $\varphi = 0$ se invertiamo i fuochi. In altre parole questa rappresentazione corrisponde a un angolo $\beta = \pi$ ($\beta = 0$) nella (3.57). Questo in generale non sarà vero per il moto in quanto, avendo scelto $\varphi(0) = 0$, non possiamo anche assumere che $\dot{r}(0) = 0$ come si ha nei punti di massima distanza dal centro del moto

se poniamo $e = \delta/d$, e $p = \frac{d^2 - \delta^2}{d} = \frac{1}{d}(1 - e^2)$. Possiamo quindi enunciare la **prima legge di Keplero**:

Il pianeta percorre un'orbita ellittica di cui il Sole occupa uno dei fuochi.

Secondo caso. $e = 1$.

Ricordando la (3.64), $e = 1 \Rightarrow E = 0$, e quindi un'orbita aperta, $r_M \rightarrow \infty$ (v. figura 3.2). In questo caso l'orbita è una parabola. Infatti dalla (3.62) si deduce

$$r + r \cos \varphi = p. \quad (3.66)$$

Se adesso si passa a coordinate cartesiane, cioè

$$P - O = x\mathbf{e}_1 + y\mathbf{e}_2, \implies r = \sqrt{x^2 + y^2}, \text{ e } r \cos \varphi = x.$$

Quindi elevando al quadrato la (3.66) si ottiene

$$x = -\frac{y^2}{2p} + \frac{p}{2},$$

cioè una parabola il cui asse coincide con l'asse x e vertice in $x = p/2$. E' abbastanza semplice verificare che la (3.57) è l'equazione di una conica (basta esprimere r e $\cos \varphi$, $\sin \varphi$ in coordinate cartesiane e razionalizzare).

Terzo caso. $e > 1$.

Ricordando la (3.59), questo caso corrisponde ad $E > 0$, e quindi ad orbite aperte. Se $e > 1$, esistono angoli "proibiti" che sono tutti quelli che soddisfano la disuguaglianza

$$\cos \varphi > -\frac{1}{e}.$$

L'orbita (3.62) si riduce ad un ramo di iperbole (la dimostrazione viene lasciata per esercizio).

3.4.4 La terza legge

A partire dall'equazione per l'orbita possiamo calcolare le lunghezze dei semiassi maggiore a e minore b dell'ellisse nel caso $e < 1$. Abbiamo infatti, posto r_m e r_M il minimo e il massimo raggio,

$$a = \frac{1}{2}(r_m + r_M) \stackrel{(3.63)}{=} \frac{1}{2} \frac{p}{(1 - e^2)}, \quad (3.67)$$

$$b = a\sqrt{1 - e^2}. \quad (3.68)$$

Possiamo ora calcolare in termini dei parametri fisici l'area spazzata dal raggio vettore in una rivoluzione, ovvero l'area dell'ellisse

$$\text{area ellisse} = \pi ab = \int_0^T \dot{A} dt = \dot{A}T, \quad (3.69)$$

siccome la velocità areolare è costante. Nella (3.69) possiamo eliminare b

$$\pi a^2 \sqrt{1 - e^2} = \dot{A} T,$$

da cui elevando al quadrato

$$T^2 = \frac{\pi^2}{\dot{A}^2} a^4 (1 - e^2) \stackrel{(3.67)}{=} \frac{\pi^2}{\dot{A}^2} a^3 \frac{p}{2} \stackrel{(3.60)}{=} a^3 \frac{4\pi^2}{K},$$

cioè

$$\frac{T^2}{a^3} = \frac{4\pi^2}{K},$$

che è la **terza legge di Keplero:**

I quadrati dei periodi di rivoluzione dei pianeti stanno tra loro nel rapporto dei cubi dei semiassi maggiori delle orbite.

Notiamo infine che la terza legge afferma (indirettamente) che il rapporto $\frac{T^2}{a^3}$ è indipendente dal pianeta considerato e dalle condizioni iniziali.

Capitolo 4

Sistemi vincolati e coordinate lagrangiane

Molti esempi di sistemi meccanici quali il moto di un punto su una traiettoria assegnata, il moto di un pendolo semplice, etc. prevedono una descrizione in questi termini: il “corpo” è rappresentato tramite un punto materiale; ad esso sono applicate delle forze specificate in funzione della sua posizione e velocità; si richiede infine che il moto soddisfi a delle limitazioni a priori sulle sue possibili posizioni e velocità. Per esempio, nel caso del pendolo semplice, il moto deve avvenire in modo che il punto mantenga invariata la sua distanza da un punto fisso.

Nel caso dei sistemi rigidi le limitazioni al moto consistono nell’assumere che le mutue distanze dei punti restino costanti durante il moto del corpo.

Vogliamo qui dare uno schema matematico per il trattamento di un sistema (finito) di punti materiali soggetti a tali limitazioni, che chiameremo **vincoli**. Per questo è necessario iniziare dallo studio della cinematica.

4.1 Sistemi olonomi

Consideriamo uno spazio affine euclideo A di dimensione 3 in cui è definito un riferimento cartesiano ortogonale $\{O, e_x, e_y, e_z\}$. E’ poi dato un sistema di N punti materiali $\{P_i, i = 1, 2, \dots, N\}$ i cui vettori posizione sono $\mathbf{x}_i, i = 1, \dots, N$,

$$\mathbf{x}_i = x_i \mathbf{e}_x + y_i \mathbf{e}_y + z_i \mathbf{e}_z .$$

Al solito indichiamo con $\dot{\mathbf{x}}_i = \frac{d\mathbf{x}_i}{dt}$ il vettore velocità del punto P_i .

Definizione 4.1.1 Si dice **vincolo cinematico** per il sistema una qualsiasi relazione tra le posizioni, le velocità e il tempo del tipo

$$F(\mathbf{x}_1, \dots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_N, t) = 0, \quad (4.1)$$

che chiameremo **equazione vincolare**.

Definizione 4.1.2 Diremo che un sistema di punti è **vincolato** se i vettori posizione e velocità dei punti che lo compongono sono soggetti a uno o più vincoli della forma (4.1), ovvero sono assegnate m equazioni vincolari

$$g_k(\mathbf{x}_1, \dots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_N, t) = 0, \quad k = 1, \dots, m. \quad (4.2)$$

Nella quasi totalità dei sistemi di interesse fisico, le equazioni in (4.2) hanno una forma semplificata in quanto le velocità compaiono linearmente nelle funzioni g_k ,

$$g_k(\mathbf{x}_1, \dots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_N, t) = h_k(\mathbf{x}_1, \dots, \mathbf{x}_N, t) + \sum_{i=1}^n \mathbf{b}_i(\mathbf{x}_1, \dots, \mathbf{x}_N, t) \cdot \dot{\mathbf{x}}_i. \quad (4.3)$$

In molti casi, poi, è possibile **integrare** queste equazioni vincolari, determinando un insieme di funzioni f_k , $k = 1, \dots, m$, per cui sia abbia

$$\left\{ \begin{array}{l} f_1(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = 0, \\ f_2(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = 0, \\ \vdots \\ f_m(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = 0, \end{array} \right. \quad (4.4)$$

e le equazioni (4.3) siano conseguenza (tramite una semplice derivazione rispetto al tempo) delle (4.4). Un esempio di vincolo imposto sulla velocità, che però può essere “integrato” e quindi si traduce in un vincolo sulle coordinate, è il vincolo di rotolamento puro che verrà analizzato nell’esempio 4.1.2.

Definizione 4.1.3 Se tutti gli m vincoli cui il sistema è soggetto sono espressi nella forma (4.4) allora diremo che il sistema è **olonomo**. Quando invece i vincoli sono nella forma generale (4.2) e non possono esser tutti ricondotti alla forma (4.4), parleremo di vincoli **anolonomi**.

Diremo infine che un **vincolo è fisso** se il vincolo non dipende dal tempo¹, cioè se la funzione vincolare non dipende esplicitamente dal tempo

$$f_k(\mathbf{x}_1, \dots, \mathbf{x}_N) = 0, \quad \forall k = 1, \dots, m.$$

Da ora in poi ci limitiamo a trattare di sistemi di vincoli olonomi, ovvero del tipo (4.4), eventualmente dipendenti dal tempo.

¹I vincoli fissi sono di gran lunga i più comuni. Inoltre la dipendenza esplicita delle equazioni vincolari dal tempo complica un po’ la trattazione: si consiglia di “leggere” tutto questa trattazione almeno una volta *sopponendo i vincoli fissi*, ovvero eliminando le dipendenze esplicite dal tempo.

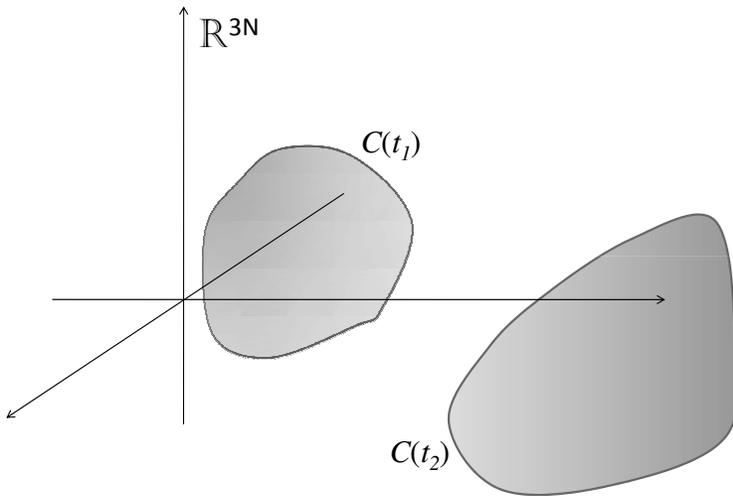


Figura 4.1: Rappresentazione schematica dello spazio delle configurazioni $\mathcal{C}(t)$, a due diversi istanti di tempo t_1 e t_2 .

Nel seguito, per ragioni di brevità, denoteremo con $\tilde{\mathbf{x}}$ la $3N$ -upla costituita dalle coordinate degli N punti, cioè

$$\tilde{\mathbf{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_N) = (\underbrace{x_1, y_1, z_1}_{\mathbf{x}_1}, \dots, \underbrace{x_N, y_N, z_N}_{\mathbf{x}_N}) \in \mathbb{R}^{3N}. \quad (4.5)$$

Iniziamo con il formulare alcune **richieste sulle funzioni** f_k :

1. Il sistema (4.4), inteso appunto come sistema di m equazioni in $3N$ incognite, (le coordinate dei punti P_i), abbia soluzioni per ogni tempo t , o in altri termini, che ad ogni istante t , l'insieme delle $\tilde{\mathbf{x}} \in \mathbb{R}^{3N}$, per cui (4.4) è soddisfatto, sia non vuoto. Questa è l'ovvia richiesta che il sistema di punti “possa stare da qualche parte!”. Indicheremo con $\mathcal{C}(t) \subset \mathbb{R}^{3N}$, l'insieme delle soluzioni del sistema (4.4), e lo chiameremo **spazio delle configurazioni** all'istante t

$$\mathcal{C}(t) = \{ \tilde{\mathbf{x}} \in \mathbb{R}^{3N} : \tilde{\mathbf{x}} \text{ è soluzione di (4.4) al tempo } t \}.$$

Gli elementi di $\mathcal{C}(t)$, ovvero le $3N$ -uple soluzioni di (4.4), vengono detti **configurazioni**.

Ovviamente l'insieme $\mathcal{C}(t)$ evolve nel tempo poiché i vincoli variano in t . Se i vincoli sono tutti fissi, allora \mathcal{C} è un insieme fissato indipendente dal tempo.

2. Le funzioni f_k siano “abbastanza” regolari in $\mathcal{C}(t) \subset \mathbb{R}^{3N}$, per ogni t , e altrettanto come funzioni di t . In pratica, chiederemo tutte le f_k abbiano derivate seconde continue.

3. I vincoli che abbiamo imposto siano tutti **effettivi**, ovvero che non ci siano nel sistema (4.4) equazioni vincolari implicate da altre (per esempio, non ci siano equazioni che sono combinazioni lineari di altre). Il teorema della funzione implicita (o teorema del Dini²) ci fornisce la condizione per cui tale richiesta è soddisfatta. Costruiamo la matrice jacobiana delle derivate prime delle f_k rispetto a tutte le $3N$ variabili (che possiamo pensare di aver ordinato $(x_1, y_1, z_1, \dots, x_N, y_N, z_N)$). Abbiamo una matrice di m righe e $3N$ colonne

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial x_N} & \frac{\partial f_1}{\partial y_N} & \frac{\partial f_1}{\partial z_N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial y_1} & \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial x_N} & \frac{\partial f_m}{\partial y_N} & \frac{\partial f_m}{\partial z_N} \end{pmatrix}. \quad (4.6)$$

I vincoli saranno tutti effettivi se, per ogni tempo t , il rango di questa matrice è massimo³.

Denotiamo, d'ora in poi, con

$$\frac{\partial f_k}{\partial \mathbf{x}_j} = \left(\frac{\partial f_k}{\partial x_j}, \frac{\partial f_k}{\partial y_j}, \frac{\partial f_k}{\partial z_j} \right), \quad (4.7)$$

il vettore gradiente di f_k rispetto alle coordinate (x_j, y_j, z_j) , coordinate del j -esimo punto materiale ed introduciamo, in \mathbb{R}^{3N} , i vettori gradiente

$$\begin{aligned} \tilde{\nabla} f_k &= \left(\frac{\partial f_k}{\partial \mathbf{x}_1}, \dots, \frac{\partial f_k}{\partial \mathbf{x}_N} \right) \\ &= \left(\underbrace{\frac{\partial f_k}{\partial x_1}, \frac{\partial f_k}{\partial y_1}, \frac{\partial f_k}{\partial z_1}}_{\frac{\partial f_k}{\partial \mathbf{x}_1}}, \dots, \underbrace{\frac{\partial f_k}{\partial x_N}, \frac{\partial f_k}{\partial y_N}, \frac{\partial f_k}{\partial z_N}}_{\frac{\partial f_k}{\partial \mathbf{x}_N}} \right), \quad k = 1, 2, \dots, m. \end{aligned} \quad (4.8)$$

Osserviamo quindi che la condizione di rango massimo, ovvero la condizione 3, equivale ad affermare che gli m vettori gradiente

$$\tilde{\nabla} f_1, \tilde{\nabla} f_2, \dots, \tilde{\nabla} f_m,$$

sono **linearmente indipendenti** per ogni punto di $\mathcal{C}(t)$ e per ogni tempo t .

Se le condizioni 1, 2 e 3 sono soddisfatte, il teorema della funzione implicita ci garantisce che fissato un tempo t_0 ed una generica configurazione $\tilde{\mathbf{x}}_0 \in \mathcal{C}(t_0)$, possiamo trovare un intorno $B(\tilde{\mathbf{x}}_0) \in \mathcal{C}(t_0)$, di $\tilde{\mathbf{x}}_0$, e un intervallo I contenente t_0 , tale che, limitatamente a $B(\tilde{\mathbf{x}}_0)$ e all'intervallo I , è possibile isolare

$$l = 3N - m, \quad (4.9)$$

coordinate indipendenti ed esprimere tutte le altre in termini di tali l coordinate. Per essere più precisi, è possibile, almeno per $t \in I$, esprimere le configurazioni $\tilde{\mathbf{x}} \in$

²Ulisse Dini (Pisa 1845 – Pisa 1918) è stato un matematico italiano.

³Ovviamente il rango non può essere maggiore del massimo tra m e $3N$, quindi abbiamo implicitamente assunto che $m \leq 3N$. Se $m = 3N$ la condizione di rango massimo implica che \mathcal{C} si riduce a un insieme di punti isolati di \mathbb{R}^{3N} e di conseguenza il sistema non possiede libertà di movimento. Nel seguito assumeremo sempre che $m < 3N$.

$B(\tilde{\mathbf{x}}_0)$ soltanto in **funzione di l variabili indipendenti**, le quali, a loro volta, variano in un opportuno aperto $U \subset \mathbb{R}^l$, e non necessariamente coincidono con le coordinate cartesiane.

Quindi, in generale, il teorema della funzione implicita garantisce l'esistenza di un aperto $U \subset \mathbb{R}^l$, di un intervallo I contenente t_0 , e di un'applicazione differenziabile da $U \times I$ in \mathbb{R}^{3N} che, fissato $t \in I$, opera così⁴

$$\left(\underbrace{(q_1, q_2, \dots, q_l)}_{\in U}, \underbrace{t}_{\in I} \right) \longrightarrow \tilde{\mathbf{x}} = \tilde{\mathbf{x}}(q_1, q_2, \dots, q_l, t),$$

ovvero

$$\left(\underbrace{(q_1, q_2, \dots, q_l, t)}_{\in U \times I} \right) \longrightarrow (\mathbf{x}_1(q_1, q_2, \dots, q_l, t), \dots, \mathbf{x}_N(q_1, q_2, \dots, q_l, t)), \quad (4.10)$$

e gode delle seguenti proprietà:

- per ogni $t \in I$, l'applicazione tra U e $B(\tilde{\mathbf{x}}_0) \cap \mathcal{C}(t)$ è **biunivoca**;
- le $3N$ funzioni

$$\mathbf{x}_i(q_1, \dots, q_l, t) = (x_i(q_1, \dots, q_l, t), y_i(q_1, \dots, q_l, t), z_i(q_1, \dots, q_l, t)),$$

$i = 1, \dots, N$, sono di classe C^1 nelle variabili q_1, q_2, \dots, q_l , e t ;

- $\forall (q_1, q_2, \dots, q_l) \in U$, e $\forall t \in I$,

$$f_k(\mathbf{x}_1(q_1, \dots, q_l, t), \dots, \mathbf{x}_N(q_1, \dots, q_l, t), t) = 0, \quad \forall k = 1, 2, \dots, m.$$

Diremo che le (q_1, \dots, q_n) sono un sistema di **coordinate locali** adatte a $\mathcal{C}(t)$, e le chiameremo **coordinate lagrangiane**, o **parametri lagrangiani**. Per brevità denoteremo con $\tilde{\mathbf{q}}$ la l -upla costituita dalle coordinate lagrangiane, cioè

$$\tilde{\mathbf{q}} = (q_1, q_2, \dots, q_l).$$

Il numero l di parametri lagrangiani viene anche detto **numero di gradi di libertà del sistema**.

Nota 4.1.1 Il fatto che le coordinate $\tilde{\mathbf{q}}$ stiano in un aperto U di \mathbb{R}^l significa che le stesse possono variare arbitrariamente in U mantenendo, al tempo stesso, i punti $\mathbf{x}_1(q_1, q_2, \dots, q_l, t), \dots, \mathbf{x}_N(q_1, q_2, \dots, q_l, t)$, nello spazio delle configurazioni $\mathcal{C}(t)$ per ogni $t \in I$. In altri termini, avviene che, $\forall t \in I$, comunque si scelga $\tilde{\mathbf{q}} \in U$, si ha

⁴La notazione $\tilde{\mathbf{x}}(q_1, \dots, q_l, t)$ sta ad indicare che le coordinate $\tilde{\mathbf{x}}$ degli N punti materiale dipendono sia dai parametri lagrangiani (q_1, \dots, q_l) (i quali a loro volta dipenderanno dal tempo) sia esplicitamente dal tempo t , a causa del moto dei vincoli. Infatti, anche se le coordinate lagrangiane (q_1, \dots, q_l) rimangono costanti nel tempo, le posizioni $\mathbf{x}_1, \dots, \mathbf{x}_N$, dei punti materiali possono evolvere a causa del moto dei vincoli.

che il sistema di vincoli (4.4) è soddisfatto, cioè

$$\begin{cases} f_1(\tilde{\mathbf{x}}(\tilde{\mathbf{q}}, t), t) = 0, \\ f_2(\tilde{\mathbf{x}}(\tilde{\mathbf{q}}, t), t) = 0, \\ \vdots \\ f_m(\tilde{\mathbf{x}}(\tilde{\mathbf{q}}, t), t) = 0. \end{cases}$$

Nota 4.1.2 È importante rimarcare il fatto che l'applicazione che consente di esprimere le coordinate $\mathbf{x}_1, \dots, \mathbf{x}_N$, in termini dei parametri lagrangiani q_1, \dots, q_l , cioè $\tilde{\mathbf{x}}(\tilde{\mathbf{q}}, t)$, è **locale**. Non è detto, in generale, che esista un'unica applicazione che permetta di esprimere tutte le configurazioni $\tilde{\mathbf{x}} \in \mathcal{C}(t)$, in funzione delle stesse $\tilde{\mathbf{q}}$.

Nota 4.1.3 È opportuno osservare che le coordinate lagrangiane q_1, \dots, q_l , sono **indipendenti**, cioè possono variare ad arbitrio all'interno dell'aperto $U \subset \mathbb{R}^l$. Non sarebbe così se, per esempio, il sistema fosse soggetto ad un ulteriore vincolo olonomo del tipo $f_{m+1}(\mathbf{x}_1, \dots, \mathbf{x}_l) = 0$. È evidente che potremmo sempre utilizzare tale ulteriore vincolo per ridurre i parametri lagrangiani⁵, portandoli ad $(l-1)$, tuttavia in talune circostanze (che, per esempio, vedremo nella sezioni 5.3.2 e 8.4.1) questa strada può non essere la migliore. In questo caso allora le q_1, \dots, q_l , variabili lagrangiane **non sono indipendenti**, ma devono soddisfare una condizione del tipo

$$\Phi(q_1, \dots, q_l) = f_{m+1}(\mathbf{x}_1(\tilde{\mathbf{q}}), \dots, \mathbf{x}_l(\tilde{\mathbf{q}})) = 0. \quad (4.11)$$

Dunque q_1, \dots, q_l , non possono variare ad arbitrio nell'aperto $U \subset \mathbb{R}^l$, dovendo essere soddisfatta la (4.11).

Esempio 4.1.1 Sono dati due punti materiali P_1 e P_2 . Il primo punto materiale P_1 è vincolato sull'asse z , mentre il secondo sulla circonferenza giacente sul piano (x, y) , centrata nell'origine, il cui raggio r varia nel tempo secondo la legge $r(t) = \alpha t$. Si vuole determinare l'equazione dei vincoli, il numero dei gradi di libertà e selezionare i parametri lagrangiani.

Cominciando dalla (4.5), abbiamo $N = 2$ punti materiali

$$\tilde{\mathbf{x}} = (\mathbf{x}_1, \mathbf{x}_2) = (x_1, y_1, z_1, x_2, y_2, z_2).$$

Quindi, prescindendo dai vincoli, abbiamo bisogno di $3N = 6$ parametri per individuare la posizione dei due punti materiali rispetto al sistema di riferimento prescelto. Adesso consideriamo i vincoli che, nello stile della (4.21), scriveremo come

$$\begin{cases} \begin{cases} f_1(\tilde{\mathbf{x}}, t) = f_1(\mathbf{x}_1, \mathbf{x}_2, t) = 0, & \implies x_1 = 0, \\ f_2(\tilde{\mathbf{x}}, t) = f_2(\mathbf{x}_1, \mathbf{x}_2, t) = 0, & \implies y_1 = 0, \end{cases} & P_1 \text{ vincolato} \\ & \text{asse } z \\ \begin{cases} f_3(\tilde{\mathbf{x}}, t) = f_4(\mathbf{x}_1, \mathbf{x}_2, t) = 0, & \implies x_2^2 + y_2^2 - \alpha^2 t^2 = 0, \\ f_4(\tilde{\mathbf{x}}, t) = f_3(\mathbf{x}_1, \mathbf{x}_2, t) = 0, & \implies z_2 = 0, \end{cases} & P_2 \text{ vincolato} \\ & \text{circonferenza} \\ & \text{piano } (x, y) \end{cases}$$

⁵Si sottintende che gli $m+1$ vincoli soddisfino sempre le sopraelencate proprietà 1, 2 e 3.

I vincoli dunque sono 4, e banalmente soddisfano i requisiti 1, 2. Dobbiamo verificare se i vincoli sono effettivi, cioè se il requisito 3 è soddisfatto. In questo caso la matrice (4.6) è

$$\begin{pmatrix} \tilde{\nabla} f_1 \\ \tilde{\nabla} f_2 \\ \tilde{\nabla} f_3 \\ \tilde{\nabla} f_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2x_2 & 2y_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

il cui rango è 4. Infatti, poiché $x_2^2 + y_2^2 = \alpha^2 t^2$, sicuramente uno fra x_2 e y_2 , è non nullo. Supponendo che $x_2 \neq 0$, dalla matrice 4×6 , possiamo estrarre il minore 4×4

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2x_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

il cui determinante è non nullo. I 4 vincoli sono dunque effettivi e allora, applicando (4.9), possiamo individuare

$$l = \underbrace{3N}_6 - \underbrace{m}_4 = 2,$$

coordinate lagrangiane. Quindi, come parametri lagrangiani, possiamo utilizzare:

(i) $z_1 = q_1 \in \mathbb{R}$, e $x_2 = q_2$, con $q \in (-\alpha t, \alpha t)$, da cui

$$y_2 = \begin{cases} \sqrt{\alpha^2 t^2 - q_2^2}, & \text{nel semipiano } y_2 > 0, \\ -\sqrt{\alpha^2 t^2 - q_2^2} & \text{nel semipiano } y_2 < 0. \end{cases}$$

Osserviamo, come rimarcato nella nota 4.1.2, la parametrizzazione non è globale. Infatti non abbiamo un'unica espressione per y_2 in funzione x_2 , ma ne abbiamo due che si applicano "localmente".

(ii) $z_1 = q_1 \in \mathbb{R}$, e $y_2 = q_2$, con $q_2 \in (-\alpha t, \alpha t)$, da cui

$$x_2 = \begin{cases} \sqrt{\alpha^2 t^2 - q_2^2}, & \text{nel semipiano } x_2 > 0, \\ -\sqrt{\alpha^2 t^2 - q_2^2} & \text{nel semipiano } x_2 < 0. \end{cases}$$

(iii) $z_1 = q_1 \in \mathbb{R}$, e l'angolo $q_2 \in [0, 2\pi)$, che il vettore $(P_2 - O)$ forma con l'asse delle ascisse. In tal caso

$$\begin{aligned} x_2 &= \alpha t \cos q_2, \\ y_2 &= \alpha t \sin q_2. \end{aligned}$$

Pertanto, utilizzando la parametrizzazione (iii), quando scriviamo esplicitamente la (4.10) abbiamo $U = \mathbb{R} \times [0, 2\pi)$, $I = (0, +\infty)$,

$$\underbrace{(q_1, q_2, t)}_{\in U \times I} \longrightarrow \tilde{\mathbf{x}}(q_1, q_2, t) = (\mathbf{x}_1(q_1, q_2, t), \mathbf{x}_2(q_1, q_2, t)),$$

che nello specifico vuol dire

$$\begin{aligned} \mathbf{x}_1(q_1, q_2, t) &= (x_1(q_1, q_2, t), y_1(q_1, q_2, t), z_1(q_1, q_2, t)) \\ &= (0, 0, q_1), \end{aligned}$$

$$\begin{aligned} \mathbf{x}_2(q_1, q_2, t) &= (x_2(q_1, q_2, t), y_2(q_1, q_2, t), z_2(q_1, q_2, t)) \\ &= (\alpha t \cos q_2, \alpha t \sin q_2, 0). \end{aligned}$$

E quindi, se $\tilde{\mathbf{q}} = (q_1, q_2)$, si ha

$$\tilde{\mathbf{x}}(\tilde{\mathbf{q}}, t) = (0, 0, q_1, \alpha t \cos q_2, \alpha t \sin q_2, 0). \quad (4.12)$$

Esempio 4.1.2 Un esempio di vincolo (4.3) che coinvolge le derivate rispetto al tempo ma che si può integrare è il vincolo di **rotolamento puro**. Consideriamo un disco di raggio R che rotola lungo una guida rettilinea. Il sistema, a priori, ha solo due gradi di libertà: l'ascissa x del centro O , e l'angolo di rotolamento φ . Quest'ultimo,

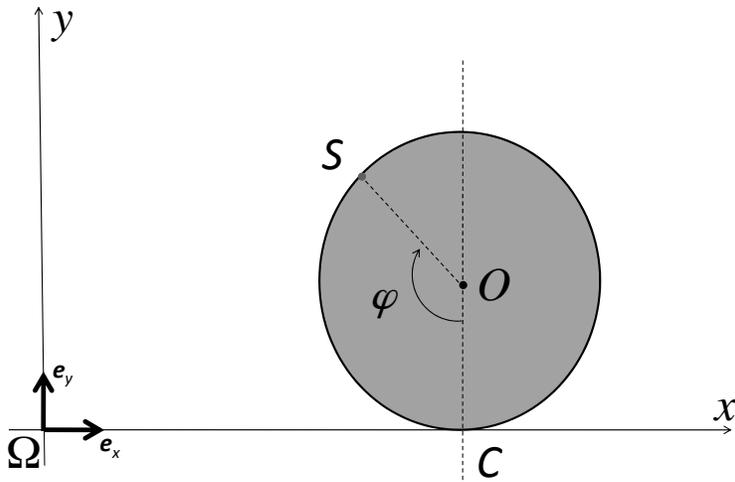


Figura 4.2: Disco di raggio R che rotola senza strisciare (rotolamento puro) sulla guida orizzontale x .

riferendoci alla figura 4.2, si misura così: si fissa sul bordo del disco un punto S (identificandolo, per esempio, con un “puntino di vernice”). L'angolo φ è quello formato dal raggio SO con una retta fissata (per esempio la retta verticale). In particolare φ sarà positivo se corrisponde a rotazioni antiorarie, negativo altrimenti (è il caso dell'angolo rappresentato in figura). Se C denota il punto di contatto fra disco e guida, diremo che il disco **rotola senza strisciare** o **rotola di rotolamento puro**, se $v_C = 0$,

cioè se, ad ogni istante, la velocità del punto di contatto è nulla. Per calcolare v_C , applichiamo la formula dei moti rigidi (6.32) del paragrafo 6.3

$$\underbrace{(C - \dot{\Omega})}_{v_C} = \underbrace{(O - \dot{\Omega})}_{v_O} + \boldsymbol{\omega} \wedge (C - O) .$$

Ora, $(O - \dot{\Omega}) = x\mathbf{e}_x + R\dot{\varphi}\mathbf{e}_y$, $\implies v_O = \dot{x}\mathbf{e}_x$, $(C - O) = -R\mathbf{e}_y$, $e \boldsymbol{\omega} = \dot{\varphi}\mathbf{e}_z$, da cui, imponendo $v_C = 0$, otteniamo

$$(\dot{x} + R\dot{\varphi})\mathbf{e}_x = 0, \implies \dot{x} = -R\dot{\varphi}, \quad (4.13)$$

cioè una relazione vincolare che coinvolge le derivate rispetto al tempo dei due gradi di libertà. La (4.13) può però essere integrata,

$$x = x_o - R(\varphi - \varphi_o),$$

dove x_o è l'ascissa di O che corrisponde all'angolo di rotolamento iniziale φ_o . Siccome possiamo sempre scegliere la configurazione iniziale in modo che $x_o = 0$ e $\varphi_o = 0$, abbiamo

$$x = -R\varphi, \quad (4.14)$$

che è una relazione che lega l'ascissa x del centro O e l'angolo di rotazione φ . E' quindi sufficiente un solo parametro ad identificare le configurazioni del disco.

Notiamo che la (4.14) è coerente dal punto di vista dei segni: se $\varphi < 0$, il disco ruota un senso orario e $x > 0$. Se invece il disco ruota in senso antiorario l'ascissa x è negativa. Non solo ma la (4.14) ha anche un chiaro significato fisico: infatti, prescindendo dal segno, $R\varphi$ rappresenta la lunghezza dell'arco CS , che è uguale alla lunghezza del segmento $C\Omega$.

Ricaviamo infine la curva $x(\varphi(t))$, $y(\varphi(t))$ descritta dal punto S nel suo moto. Sfruttando ancora formula dei moti rigidi (6.32) abbiamo

$$\begin{aligned} \mathbf{v}_S &= \mathbf{v}_O + \dot{\varphi}\mathbf{e}_z \wedge (S - O) \\ &= -R\dot{\varphi}\mathbf{e}_x + \dot{\varphi}\mathbf{e}_z \wedge (R\sin\varphi\mathbf{e}_x - R\cos\varphi\mathbf{e}_y), \end{aligned}$$

da cui, supponendo $\varphi(0) = 0$, otteniamo

$$\begin{cases} \frac{dx(\varphi(t))}{dt} = -R\dot{\varphi} + R\dot{\varphi}\cos\varphi, \\ x(0) = 0, \end{cases} \implies x(\varphi(t)) = -R[\varphi(t) - \sin\varphi(t)], \quad (4.15)$$

$$\begin{cases} \frac{dy(\varphi(t))}{dt} = R\dot{\varphi}\sin\varphi, \\ y(0) = 0, \end{cases} \implies y(\varphi(t)) = R[1 - \cos\varphi(t)]. \quad (4.16)$$

Notiamo infine che, ponendo $\psi = -\varphi(t)$ e considerando $-\pi \leq \psi \leq \pi$, la curva descritta dal sistema (4.15), (4.16) può scriversi come

$$x(\psi) = R(\psi + \sin\psi), \quad z(\psi) = R(1 - \cos\psi), \quad -\pi \leq \psi \leq \pi. \quad (4.17)$$

Tale curva viene detta **cicloide**.

Esempio 4.1.3 Consideriamo un disco di raggio r che, come in figura 4.3, si mantiene verticale e rotola senza strisciare sulla circonferenza del piano x, y , centrata in O e avente raggio R . Il sistema, a priori, possiede due gradi di libertà identificati

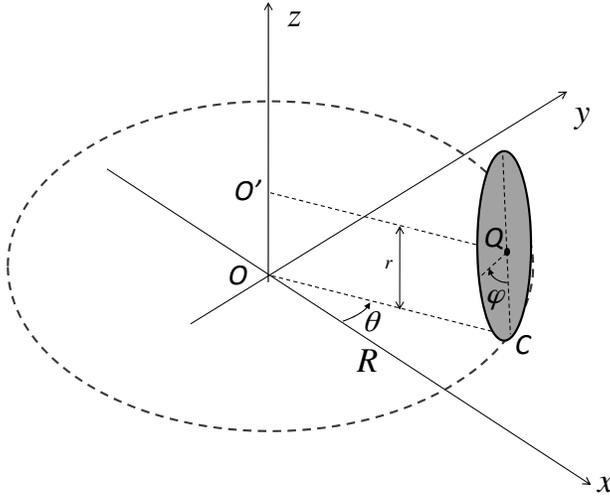


Figura 4.3: Disco di raggio r che, mantenendosi verticale rotola senza strisciare sulla circonferenza di raggio R .

dall'angolo θ , e dall'angolo di rotolamento φ . Le coordinate del centro del disco sono

$$Q - O = R \cos \theta \mathbf{e}_x + R \sin \theta \mathbf{e}_y + R \mathbf{e}_z,$$

si che

$$\mathbf{v}_Q = -R\dot{\theta} \sin \theta \mathbf{e}_x + R\dot{\theta} \cos \theta \mathbf{e}_y. \quad (4.18)$$

Le componenti della velocità angolare nel SdR della figura 4.3, sono⁶

$$\boldsymbol{\omega} = \dot{\varphi} \cos \theta \mathbf{e}_x + \dot{\varphi} \sin \theta \mathbf{e}_y + \dot{\theta} \mathbf{e}_z, \quad (4.19)$$

dove φ è l'angolo di rotolamento (misurato come nell'esempio 4.1.2). Applicando (6.32) (vedi sezione 6.3) ed imponendo $\mathbf{v}_C = 0$ (vincolo di rotolamento puro) si ottiene

$$\begin{aligned} 0 &= \mathbf{v}_C = \mathbf{v}_Q + \boldsymbol{\omega} \wedge (C - Q) \\ &= (-R\dot{\theta} - r\dot{\varphi}) \sin \theta \mathbf{e}_x + (R\dot{\theta} + r\dot{\varphi}) \cos \theta \mathbf{e}_y, \end{aligned}$$

cioè

$$\dot{\theta} = -\frac{r}{R}\dot{\varphi}, \implies \theta = -\frac{r}{R}\varphi, \quad (4.20)$$

⁶La formula (4.19) viene dimostrata nell'esempio 6.8.1.

dove, al solito, abbiamo orientato gli assi x e y e in modo che $\theta_o = \varphi_o = 0$. Ancora una volta, quindi, un vincolo che coinvolge derivate rispetto al tempo si traduce in un vincolo olonomo.

4.1.1 Spostamenti virtuali in funzione delle coordinate lagrangiane

Nella formulazione della meccanica dei sistemi vincolati è essenziale caratterizzare tutte le possibili variazioni di configurazione che un sistema di punti può effettuare a partire da una configurazione data. Per fissare le idee supponiamo per il momento che il sistema sia soggetto soltanto a m vincoli fissi

$$\left\{ \begin{array}{l} f_1(\mathbf{x}_1, \dots, \mathbf{x}_N) = 0, \\ f_2(\mathbf{x}_1, \dots, \mathbf{x}_N) = 0, \\ \vdots \\ f_m(\mathbf{x}_1, \dots, \mathbf{x}_N) = 0. \end{array} \right. \quad (4.21)$$

Sia $\mathcal{C} \subset \mathbb{R}^{3N}$, lo spazio delle configurazioni. Sia $\mathcal{B} \subseteq \mathcal{C}$, un sottoinsieme di \mathcal{C} , tale che le coordinate di ogni suo punto possano essere espresse univocamente in termini di $\tilde{\mathbf{q}} = (q_1, q_2, \dots, q_l)$ parametri lagrangiani, i quali, a loro volta, variano nell'aperto $U \subset \mathbb{R}^l$. Quindi, ricordando quanto visto nella sezione 1.4 del capitolo 1, abbiamo che l'applicazione

$$U \ni \tilde{\mathbf{q}} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_l \end{pmatrix} \longrightarrow \tilde{\mathbf{x}}(\tilde{\mathbf{q}}) = \begin{pmatrix} \mathbf{x}_1(q_1, q_2, \dots, q_l) \\ \mathbf{x}_2(q_1, q_2, \dots, q_l) \\ \vdots \\ \mathbf{x}_N(q_1, q_2, \dots, q_l) \end{pmatrix} \in \mathcal{B} \subseteq \mathcal{C} \subset \mathbb{R}^{3N},$$

gode delle seguenti proprietà:

1. $\tilde{\mathbf{x}}(\tilde{\mathbf{q}}) \in C^1(U)$.
2. $\forall \tilde{\mathbf{q}} \in U$, l'applicazione è invertibile.

Sappiamo allora che $\tilde{\mathbf{x}}(\tilde{\mathbf{q}})$ definisce un sistema di coordinate curvilinee su \mathcal{B} , e vale il seguente

Teorema 4.1.1 *Gli l vettori di \mathbb{R}^{3N} ,*

$$\begin{aligned} \tilde{\mathbf{u}}_1 = \frac{\partial \tilde{\mathbf{x}}}{\partial q_1} &= \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_1}(q_1, q_2, \dots, q_l) \\ \frac{\partial \mathbf{x}_2}{\partial q_1}(q_1, q_2, \dots, q_l) \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_1}(q_1, q_2, \dots, q_l) \end{pmatrix} \\ &\vdots \\ \tilde{\mathbf{u}}_l = \frac{\partial \tilde{\mathbf{x}}}{\partial q_l} &= \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_l}(q_1, q_2, \dots, q_l) \\ \frac{\partial \mathbf{x}_2}{\partial q_l}(q_1, q_2, \dots, q_l) \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_l}(q_1, q_2, \dots, q_l) \end{pmatrix}, \end{aligned} \quad (4.22)$$

sono linearmente indipendenti e **costituiscono quindi la base locale relativa alle variabili** q_1, q_2, \dots, q_l . Non solo, ma ogni vettore $\tilde{\mathbf{u}}_i$ è ortogonale ad ogni vettore gradiente $\tilde{\nabla} f_k$, cioè $\forall i = 1, 2, \dots, l$, e $\forall k = 1, 2, \dots, m$, vale⁷

$$\begin{aligned} \tilde{\mathbf{u}}_i \bullet \tilde{\nabla} f_k &= \sum_{s=1}^N \frac{\partial \mathbf{x}_s}{\partial q_i} \cdot \frac{\partial f_k}{\partial \mathbf{x}_s} \\ &= \sum_{s=1}^N \left(\frac{\partial x_s}{\partial q_i} \frac{\partial f_k}{\partial x_s} + \frac{\partial y_s}{\partial q_i} \frac{\partial f_k}{\partial y_s} + \frac{\partial z_s}{\partial q_i} \frac{\partial f_k}{\partial z_s} \right) = 0. \end{aligned} \quad (4.23)$$

Dim. Cominciamo col provare che i vettori $\tilde{\mathbf{u}}_i$, $i = 1, 2, \dots, l$, sono linearmente indipendenti. Supponiamo per semplicità che l sia multiplo di 3 e, sempre per semplicità, supponiamo che, applicando il teorema di funzione implicita, le coordinate lagrangiane

⁷Con “ \bullet ” indichiamo il prodotto scalare in \mathbb{R}^{3N} , che è appunto la somma di N prodotti scalari “ \cdot ” in \mathbb{R}^3 .

q_1, \dots, q_l , siano le prime l coordinate cartesiane dei primi $l/3$ punti materiali, cioè

$$\begin{aligned} x_1 &= q_1, \\ y_1 &= q_2, \\ &\vdots \\ y_{l/3} &= q_{l-1}, \\ z_{l/3} &= q_l \\ x_{l/3+1} &= x_{l/3+1}(q_1, q_2, \dots, q_l), \\ y_{l/3+2} &= y_{l/3+2}(q_1, q_2, \dots, q_l), \\ &\vdots \\ z_N &= z_N(q_1, q_2, \dots, q_l). \end{aligned}$$

Consideriamo adesso la matrice $l \times 3N$, le cui righe sono le componenti dei vettori

$$\tilde{\mathbf{u}}_i = \frac{\partial \tilde{\mathbf{x}}}{\partial q_i}, \quad i = 1, \dots, l,$$

$$\begin{aligned} \mathbb{M} &= \begin{pmatrix} \frac{\partial \tilde{\mathbf{x}}}{\partial q_1} \\ \frac{\partial \tilde{\mathbf{x}}}{\partial q_2} \\ \vdots \\ \frac{\partial \tilde{\mathbf{x}}}{\partial q_l} \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial q_1} & \frac{\partial y_1}{\partial q_1} & \frac{\partial z_1}{\partial q_1} & \dots & \frac{\partial z_{l/3}}{\partial q_1} & \frac{\partial x_{l/3+1}}{\partial q_1} & \dots & \frac{\partial z_N}{\partial q_1} \\ \frac{\partial x_1}{\partial q_2} & \frac{\partial y_1}{\partial q_2} & \frac{\partial z_1}{\partial q_2} & \dots & \frac{\partial z_{l/3}}{\partial q_2} & \frac{\partial x_{l/3+1}}{\partial q_2} & \dots & \frac{\partial z_N}{\partial q_2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial x_1}{\partial q_l} & \frac{\partial y_1}{\partial q_l} & \frac{\partial z_1}{\partial q_l} & \dots & \frac{\partial z_{l/3}}{\partial q_l} & \frac{\partial x_{l/3+1}}{\partial q_l} & \dots & \frac{\partial z_N}{\partial q_l} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \frac{\partial x_{l/3+1}}{\partial q_1} & \dots & \frac{\partial z_N}{\partial q_1} \\ 0 & 1 & 0 & \dots & 0 & \frac{\partial x_{l/3+1}}{\partial q_2} & \dots & \frac{\partial z_N}{\partial q_2} \\ \vdots & \vdots & \vdots & & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & \frac{\partial x_{l/3+1}}{\partial q_l} & \dots & \frac{\partial z_N}{\partial q_l} \end{pmatrix}. \end{aligned}$$

Siccome è immediato verificare che le prime l colonne corrispondono alla matrice identità deduciamo che matrice ha rango massimo, cioè l . Ma questo vuol dire che gli l vettori $\frac{\partial \tilde{\mathbf{x}}}{\partial q_1}, \dots, \frac{\partial \tilde{\mathbf{x}}}{\partial q_l}$, ovvero $\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l$, sono linearmente indipendenti.

Questo si estende anche al caso, più generale, in cui le variabili lagrangiane non coincidano con le prime l coordinate dei primi $l/3$ punti materiali, ma comunque esista una corrispondenza biunivoca fra (q_1, q_2, \dots, q_l) e $(x_1, y_1, \dots, z_{l/3})$. Questo, per

esempio, accade se (q_1, q_2, \dots, q_l) sono le coordinate polari sferiche dei primi $l/3$ punti. In tal caso la matrice jacobiana $l \times l$

$$\mathbb{J} = \begin{pmatrix} \frac{\partial x_1}{\partial q_1} & \frac{\partial y_1}{\partial q_1} & \frac{\partial z_1}{\partial q_1} & \dots & \frac{\partial y_{l/3}}{\partial q_1} & \frac{\partial z_{l/3}}{\partial q_1} \\ \frac{\partial x_1}{\partial q_2} & \frac{\partial y_1}{\partial q_2} & \frac{\partial z_1}{\partial q_2} & \dots & \frac{\partial y_{l/3}}{\partial q_2} & \frac{\partial z_{l/3}}{\partial q_2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \frac{\partial x_1}{\partial q_l} & \frac{\partial y_1}{\partial q_l} & \frac{\partial z_1}{\partial q_l} & \dots & \frac{\partial y_{l/3}}{\partial q_l} & \frac{\partial z_{l/3}}{\partial q_l} \end{pmatrix}$$

è invertibile, ovvero $\det \mathbb{J} \neq 0$. E' quindi immediato rendersi conto che le prime l colonne della matrice \mathbb{M} coincidono con \mathbb{J} . Siamo allora di nuovo in grado di isolare in \mathbb{M} un minore $l \times l$ con determinante non nullo e pertanto concludere che gli l vettori di \mathbb{R}^{3N} , $\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l$, sono linearmente indipendenti. Evidentemente tutto ciò si generalizza anche al caso in cui l non sia multiplo di 3 e le coordinate (q_1, \dots, q_l) non siano le l coordinate cartesiane dei primi punti materiali.

Proviamo adesso la (4.23). Partendo dalla k -esima equazione vincolare (4.21), abbiamo

$$f_k(\mathbf{x}_1(\tilde{\mathbf{q}}), \dots, \mathbf{x}_N(\tilde{\mathbf{q}})) = f_k(\mathbf{x}_1(q_1, q_2, \dots, q_l), \dots, \mathbf{x}_N(q_1, q_2, \dots, q_l)) = 0,$$

per ogni $\tilde{\mathbf{q}} = (q_1, q_2, \dots, q_l) \in U$. Ma allora se deriviamo rispetto alla generica q_i otteniamo

$$0 = \frac{\partial f_k(\mathbf{x}_1(q_1, q_2, \dots, q_l), \dots, \mathbf{x}_N(q_1, q_2, \dots, q_l))}{\partial q_i} = \sum_{s=1}^N \frac{\partial f_k}{\partial \mathbf{x}_s} \cdot \frac{\partial \mathbf{x}_s}{\partial q_i} = \tilde{\nabla} f_k \bullet \tilde{\mathbf{u}}_i.$$

□

Il precedente teorema si estende con facilità anche al caso di vincoli mobili. In tal caso, fissata la configurazione, i vettori $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \dots, \tilde{\mathbf{u}}_l$, variano nel tempo ma, ad ogni istante t , sono linearmente indipendenti.

Definizione 4.1.4 Si dice *spostamento virtuale*⁸ del sistema di N punti materiale re-

⁸Una piccola divagazione storica per spiegare l'origine del nome "spostamenti virtuali". Il nome è legato al cosiddetto *Principio dei lavori virtuali*, condizione per caratterizzare le configurazioni di equilibrio di un sistema vincolato; ecco come lo riassume Lagrange nella *Mécanique Analytique*: *Se un sistema qualsiasi di corpi o punti soggetti a delle forze qualsiasi, è in equilibrio, e si impone a questo sistema un piccolo spostamento qualsiasi, in virtù del quale ogni punto percorre uno spazio infinitamente piccolo che sarà espresso dal suo spostamento virtuale, la somma delle forze, ognuna moltiplicata per lo spazio su cui essa è applicata, percorso secondo la direzione stessa della forza, sarà sempre uguale a zero, prendendo per positivi gli spazi percorsi nel senso della forza per negativi quelli percorsi nel senso opposto.* (ovviamente possiamo semplificare la frase di Lagrange usando i prodotti scalari tra le forze e gli spostamenti, ovvero il lavoro delle forze.) L'idea che c'è sotto è che si "saggia" una posizione di equilibrio "provando a spostare" il sistema dall'equilibrio stesso, in modo "infinitesimo". Questo è ciò che si fa quando vogliamo farsi un'idea di quanto pesi una valigia: la si "prova a sollevare" da terra di un'altezza impercettibile, lo sforzo esercitato dal nostro braccio dovrà equilibrare quello esercitato dalla forza peso.

lativo alla generica configurazione \mathbf{x} , un qualsiasi vettore $\delta\tilde{\mathbf{x}} \in \mathbb{R}^{3N}$

$$\begin{aligned}\delta\tilde{\mathbf{x}} &= (\delta\mathbf{x}_1, \dots, \delta\mathbf{x}_N) \\ &= \underbrace{(\delta x_1, \delta y_1, \delta z_1, \dots)}_{\delta\mathbf{x}_1}, \underbrace{(\delta x_N, \delta y_N, \delta z_N)}_{\delta\mathbf{x}_N},\end{aligned}$$

tale che⁹

$$\delta\tilde{\mathbf{x}} \bullet \tilde{\nabla} f_k = \sum_{s=1}^N \delta\mathbf{x}_s \cdot \frac{\partial f_k}{\partial \mathbf{x}_s} = 0, \quad \forall k = 1, 2, \dots, m, \quad (4.24)$$

essendo $\tilde{\nabla} f_k$, $k = 1, \dots, m$, il gradiente calcolato nella configurazione \mathbf{x} . Il vettore di \mathbb{R}^3 , $\delta\mathbf{x}_i$, $i = 1, \dots, N$, viene detto **spostamento virtuale del i -esimo punto materiale** P_i .

Indichiamo con¹⁰ \mathcal{N} , il sottospazio di \mathbb{R}^{3N} generato dagli m gradienti $\tilde{\nabla} f_1, \dots, \tilde{\nabla} f_m$. Diremo dunque che il vettore $\delta\mathbf{x} \in \mathbb{R}^{3N}$ è **spostamento virtuale se $\delta\mathbf{x}$ appartiene al sottospazio ortogonale ad \mathcal{N}** , cioè se $\delta\mathbf{x} \in \mathcal{N}_\perp$. Abbiamo quindi il seguente

Teorema 4.1.2 *Ogni spostamento virtuale può esprimersi come combinazione lineare degli l vettori $\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l$, cioè*

$$\delta\tilde{\mathbf{x}} = \sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s, \quad \text{con } \delta q_s \in \mathbb{R}. \quad (4.25)$$

In particolare, lo spostamento virtuale dell' i -esimo punto materiale è

$$\delta\mathbf{x}_i = \sum_{s=1}^l \frac{\partial \mathbf{x}_i}{\partial q_s} \delta q_s, \quad i = 1, 2, \dots, N. \quad (4.26)$$

Dim. Dalla definizione 4.1.4 sappiamo che se $\delta\tilde{\mathbf{x}}$ è spostamento virtuale allora soddisfa la (4.24) per ogni $k = 1, 2, \dots, m$. Ora, siccome i gradienti $\tilde{\nabla} f_k$, $k = 1, \dots, m$, sono linearmente indipendenti, ne deduciamo che il vettore $\delta\tilde{\mathbf{x}} = (\delta\mathbf{x}_1, \dots, \delta\mathbf{x}_N)$ deve appartenere al sottospazio \mathcal{N}_\perp , la cui dimensione è $l = 3N - m$ di \mathbb{R}^{3N} . Siccome una base di \mathcal{N}_\perp è $\{\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l\}$ (vedi teorema 4.1.1), avremo la (4.25).

□

Volendo far riferimento alla sezione 1.4, la l -upla $(\delta q_1, \dots, \delta q_l)$, rappresenta la l -upla delle **componenti controvarianti** del vettore $\delta\tilde{\mathbf{x}}$ rispetto alla base locale.

⁹Attenzione ancora una volta alla notazione: nella formula che segue, e nel resto del capitolo, abbiamo usato la (4.7) per indicare il gradiente (tridimensionale) della funzione f_k rispetto alle coordinate del punto materiale P_s . Quindi $\frac{\partial f_k}{\partial \mathbf{x}_s} = \left(\frac{\partial f_k}{\partial x_s}, \frac{\partial f_k}{\partial y_s}, \frac{\partial f_k}{\partial z_s} \right)$ è un vettore di \mathbb{R}^3 e possiamo farne il prodotto scalare con $\delta\mathbf{x}_s \in \mathbb{R}^3$. Ricordiamo poi che \bullet denota il prodotto scalare in \mathbb{R}^{3N} , ovvero la somma degli N prodotti scalari di \mathbb{R}^3 .

¹⁰Il significato geometrico di \mathcal{N} verrà chiarito nella definizione 4.1.6.

Nota 4.1.4 Osserviamo che la definizione 4.1.4 scaturisce da semplici considerazioni fisiche. Un cambiamento di configurazione $\mathbf{x}_i \rightarrow \hat{\mathbf{x}}_i$, $i = 1, 2, \dots, N$, sarà compatibile con i vincoli se anche le nuove posizioni $\hat{\mathbf{x}}_i$, $i = 1, 2, \dots, N$, soddisfano le equazioni vincolari, ovvero

$$f_k(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_N) = 0, \quad \forall k = 1, \dots, m. \quad (4.27)$$

Qui però siamo interessati a variazioni **infinitesime** delle configurazioni, ovvero a “tendenze al moto” più che al moto effettivo: useremo la notazione $\delta \mathbf{x}_i$ per indicare queste variazioni. Esse possono essere caratterizzate matematicamente chiedendo che le configurazioni “variate” $\mathbf{x}_i + \delta \mathbf{x}_i$ **soddisfino le equazioni vincolari al prim’ordine** nel loro sviluppo di Taylor attorno alla configurazione $(\mathbf{x}_1, \dots, \mathbf{x}_N)$. Avremo quindi una caratterizzazione per i $\delta \mathbf{x}_i$, cioè

$$\begin{aligned} 0 &= f_k(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_N) = f_k(\mathbf{x}_1 + \delta \mathbf{x}_1, \dots, \mathbf{x}_N + \delta \mathbf{x}_N) \\ &= \underbrace{f_k(\mathbf{x}_1, \dots, \mathbf{x}_N)}_{=0} + \sum_{i=1}^N \frac{\partial f_k}{\partial \mathbf{x}_i} \cdot \delta \mathbf{x}_i, \end{aligned}$$

e quindi otteniamo la (4.24).

Nota 4.1.5 Riprendendo l’osservazione 4.1.3, vogliamo puntualizzare meglio la tesi del teorema 4.1.2. **La (4.25) è vera soltanto se le variabili lagrangiane** q_1, \dots, q_l **sono indipendenti**, cioè non soggette ad alcun ulteriore vincolo (come, ad esempio, quello dato dalla (4.11)). Se le variabili (q_1, \dots, q_l) sono vincolate (e non si vuole “eliminare” tale vincolo passando ad $(l - 1)$ coordinate lagrangiane) la (4.25) non rappresenta più il generico spostamento virtuale del sistema soggetto ad $(m + 1)$ vincoli. Infatti, applicando la definizione 4.1.4, i vettori $\delta \tilde{\mathbf{x}} \in \mathbb{R}^{3N}$ sono spostamenti virtuali soltanto se, oltre alla condizione (4.24), soddisfano anche la condizione $\delta \tilde{\mathbf{x}} \bullet \tilde{\nabla} f_{m+1} = 0$. Detto in altri termini, se \mathcal{N} denota il sottospazio di \mathbb{R}^{3N} generato dagli $(m + 1)$ gradienti $\tilde{\nabla} f_1, \dots, \tilde{\nabla} f_m, \tilde{\nabla} f_{m+1}$, $\delta \tilde{\mathbf{x}} \in \mathbb{R}^{3N}$ è spostamento virtuale se $\delta \tilde{\mathbf{x}} \in \mathcal{N}_\perp$. Il punto è che adesso il sottospazio generato da $\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l$, **non coincide** più con \mathcal{N}_\perp . Infatti $\dim \{\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l\} = 3N - m$, mentre $\dim \mathcal{N}_\perp = 3N - (m + 1)$. Quindi la generica combinazione lineare $\sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s$ non individua più uno spostamento virtuale.

La procedura fin qui analizzata si estende anche al caso di vincoli mobili. Per ogni $t \in I$, i vettori $\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l$, definiti adesso come

$$\tilde{\mathbf{u}}_j = \frac{\partial \tilde{\mathbf{x}}}{\partial q_i} = \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_j}(q_1, q_2, \dots, q_l, t) \\ \frac{\partial \mathbf{x}_2}{\partial q_j}(q_1, q_2, \dots, q_l, t) \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_j}(q_1, q_2, \dots, q_l, t) \end{pmatrix}, \quad j = 1, 2, \dots, l,$$

formano ancora la base locale relativa alle variabili (q_1, q_2, \dots, q_l) . Adesso però, fissata una qualsiasi configurazione, la base locale cambia da un istante all’altro. In ogni istante però vale sempre la (4.23), cioè

$$\tilde{\mathbf{u}}_j \bullet \tilde{\nabla} f_k = 0, \quad \forall j = 1, \dots, l; \quad \forall k = 1, \dots, m; \quad \text{e } \forall t \in I. \quad (4.28)$$

Quindi nel caso dei vincoli mobili, lo spazio degli spostamenti virtuali varierà non solo da posizione a posizione ma anche al variare del tempo.

Esempio 4.1.4 *E' dato un solo punto materiale P di massa m vincolato su una circonferenza verticale centrata in O , di raggio R , che ruota attorno all'asse z . Si vogliono determinare le equazioni vincolari, assumendo un sistema di riferimento fisso in cui l'asse delle z è la verticale ascendente passante per O .*

*Il sistema è costituito da un solo punto materiale P soggetto a due vincoli: (1) P deve stare a distanza fissata R da un punto fisso O ; (2) P appartiene a un piano verticale che ruota attorno alla verticale condotta per il punto O . La circonferenza ruotante può quindi essere vista come l'intersezione di due vincoli indipendenti (v. figura 5.1): (i) una sfera fissa di raggio R centrata nell'origine, la cui equazione implicita è $f_1(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0$; (ii) un piano ortogonale al piano x, y che ruota attorno all'asse z , la cui equazione è $f_2(x, y, z) = -x \sin \varphi(t) + y \cos \varphi(t) = 0$, essendo $\varphi(t)$ l'angolo di rotazione che è una **funzione assegnata del tempo**. In questo caso lo spazio delle configurazioni $\mathcal{C}(t)$ è la circonferenza ruotante*

$$\mathcal{C}(t) = \{f_1(x, y, z) = 0\} \cap \{f_2(x, y, z) = 0\}.$$

E' facile verificare che ∇f_1 , e ∇f_2 , sono indipendenti e quindi il teorema di funzione implicita consente di individuare un solo parametro per descrivere (almeno localmente) $\mathcal{C}(t)$. Ovviamente varie scelte sono possibile, ma quella forse più conveniente consiste nel prendere come coordinata lagrangiana l'angolo $\theta \in (-\pi, \pi]$ formato dal vettore $(P - O)$ con l'asse z , (v. ancora figura 5.1). Operando così abbiamo

$$\mathbf{x}(\theta, t) = x(\theta, t) \mathbf{e}_x + y(\theta, t) \mathbf{e}_y + z(\theta, t) \mathbf{e}_z,$$

dove

$$\begin{cases} x(\theta, t) = R \sin \theta \cos \varphi(t), \\ y(\theta, t) = R \sin \theta \sin \varphi(t), \\ z(\theta, t) = -R \cos \theta. \end{cases} \quad (4.29)$$

E' banale verificare che $f_1(\mathbf{x}(\theta, t)) = 0$, come del resto $f_2(\mathbf{x}(\theta, t)) = 0$. Il vettore che costituisce la base locale di $\mathcal{C}(t)$ è

$$\mathbf{u}_\theta = \frac{\partial \mathbf{x}}{\partial \theta} = R \cos \theta \cos \varphi(t) \mathbf{e}_x + R \cos \theta \sin \varphi(t) \mathbf{e}_y + R \sin \theta \mathbf{e}_z,$$

che, chiaramente, fissato θ , varia nel tempo. Lo spostamento virtuale è $\delta \mathbf{x} = \mathbf{u}_\theta \delta \theta$. Viene infine lasciata come esercizio la verifica che $\mathbf{u}_\theta \cdot \nabla f_1 = \mathbf{u}_\theta \cdot \nabla f_2 = 0$.

Esempio 4.1.5 *Riprendendo l'esempio 4.1.1, vogliamo adesso determinare $\tilde{\mathbf{u}}_1$, e $\tilde{\mathbf{u}}_2$ e gli spostamenti virtuali $\delta \mathbf{x}_1$ e $\delta \mathbf{x}_2$ dei due punti materiali nell'ipotesi che si faccia uso della parametrizzazione (iii). Dalle (4.22) e (4.12), abbiamo*

$$\begin{aligned} \tilde{\mathbf{u}}_1 &= \frac{\partial \tilde{\mathbf{x}}}{\partial q_1} = (0, 0, 1, 0, 0, 0), \\ \tilde{\mathbf{u}}_2 &= \frac{\partial \tilde{\mathbf{x}}}{\partial q_2} = (0, 0, 0, -\alpha t \sin q_2, \alpha t \cos q_2, 0). \end{aligned}$$

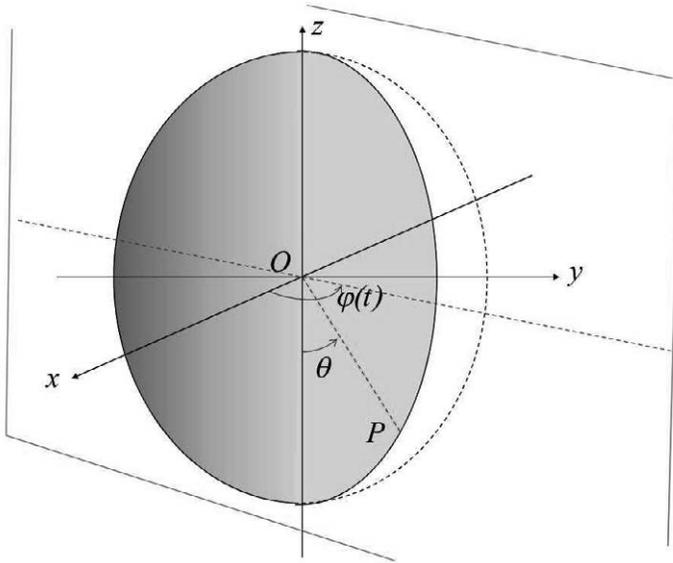


Figura 4.4: Sfera di raggio R e piano ruotante. L'angolo di rotazione $\varphi(t)$ è assegnato.

E' facile rendersi conto che $\tilde{\mathbf{u}}_1$ e $\tilde{\mathbf{u}}_2$ sono linearmente indipendenti dal momento che la matrice 2×6

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\alpha t \sin q_2 & \alpha t \cos q_2 & 0 \end{pmatrix}$$

ha sempre un minore 2×2 con determinante non nullo. Inoltre è facile provare che $\tilde{\mathbf{u}}_i \bullet \tilde{\nabla} f_k = 0, \forall i = 1, 2, \text{ e } \forall k = 1, \dots, 4$. Gli spostamenti virtuali dei punti materiali P_1 e P_2 sono dati dalla (4.26), cioè

$$\begin{aligned} \delta \mathbf{x}_1 &= \frac{\partial \mathbf{x}_1}{\partial q_1} \delta q_1 + \frac{\partial \mathbf{x}_1}{\partial q_2} \delta q_2 \\ &= \frac{\partial}{\partial q_1} [x_1(q_1, q_2, t) \mathbf{e}_x + y_1(q_1, q_2, t) \mathbf{e}_y + z_1(q_1, q_2, t) \mathbf{e}_z] \delta q_1 + \\ &\quad + \frac{\partial}{\partial q_2} [x_1(q_1, q_2, t) \mathbf{e}_x + y_1(q_1, q_2, t) \mathbf{e}_y + z_1(q_1, q_2, t) \mathbf{e}_z] \delta q_2 \\ &= \mathbf{e}_z \delta q_1 \end{aligned}$$

$$\begin{aligned} \delta \mathbf{x}_2 &= \frac{\partial}{\partial q_1} [x_2(q_1, q_2, t) \mathbf{e}_x + y_2(q_1, q_2, t) \mathbf{e}_y + z_2(q_1, q_2, t) \mathbf{e}_z] \delta q_1 + \\ &\quad + \frac{\partial}{\partial q_2} [x_2(q_1, q_2, t) \mathbf{e}_x + y_2(q_1, q_2, t) \mathbf{e}_y + z_2(q_1, q_2, t) \mathbf{e}_z] \delta q_2 \\ &= (-\alpha t \sin q_2 \mathbf{e}_x + \alpha t \cos q_2 \mathbf{e}_y) \delta q_2. \end{aligned}$$

Se invece vogliamo lo spostamento virtuale come vettore di \mathbb{R}^6 , ricordando la (4.25), abbiamo

$$\begin{aligned}\delta\tilde{\mathbf{x}} &= \tilde{\mathbf{u}}_1\delta q_1 + \tilde{\mathbf{u}}_2\delta q_2 \\ &= \left(\underbrace{0, 0, \delta q_1}_{\delta\mathbf{x}_1}, \underbrace{-\delta q_2\alpha t \sin q_2, \delta q_2\alpha t \cos q_2, 0}_{\delta\mathbf{x}_2} \right).\end{aligned}$$

4.1.2 Velocità ed energia cinetica

Un moto qualsiasi del sistema degli N punti materiali, compatibile con i vincoli, è determinato dall'assegnazione delle funzioni del tempo $q_1(t), \dots, q_l(t)$, e quindi dalle relazioni

$$\mathbf{x}_i(t) = \mathbf{x}_i(q_1(t), \dots, q_l(t), t), \quad i = 1, 2, \dots, N. \quad (4.30)$$

Derivando la (4.30) rispetto al tempo, determineremo quindi il legame tra le velocità dei punti del sistema e le coordinate lagrangiane¹¹

$$\dot{\mathbf{x}}_i(t) = \sum_{s=1}^l \frac{\partial \mathbf{x}_i}{\partial q_s}(q_1(t), \dots, q_l(t), t) \dot{q}_s + \frac{\partial \mathbf{x}_i}{\partial t}(q_1(t), \dots, q_l(t), t), \quad i = 1, \dots, N. \quad (4.31)$$

In particolare, possiamo introdurre la velocità come vettore di \mathbb{R}^{3N}

$$\tilde{\dot{\mathbf{x}}} = \begin{pmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \\ \vdots \\ \dot{\mathbf{x}}_N \end{pmatrix},$$

e, ricordando le definizioni (4.22), scrivere

$$\begin{pmatrix} \dot{\mathbf{x}}_1 \\ \vdots \\ \dot{\mathbf{x}}_N \end{pmatrix} = \sum_{s=1}^l \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_s} \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_s} \end{pmatrix} \dot{q}_s + \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial t} \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial t} \end{pmatrix},$$

cioè

$$\tilde{\dot{\mathbf{x}}} = \sum_{s=1}^l \tilde{\mathbf{u}}_s \dot{q}_s + \frac{\partial \tilde{\mathbf{x}}}{\partial t}. \quad (4.32)$$

La velocità ha dunque due componenti:

- $\sum_{s=1}^l \tilde{\mathbf{u}}_s \dot{q}_s$, che rappresenta **la velocità dei punti materiali rispetto ai vincoli**.
- $\frac{\partial \tilde{\mathbf{x}}}{\partial t}$, detta **velocità di trascinamento**. Tale componente è dovuta al moto dei vincoli.

¹¹Fare attenzione alla somiglianza formale tra la (4.31) e la (4.26): c'è però una sostanziale differenza tra queste formule e non si dovrebbe procedere oltre senza averla capita!

Possiamo ora utilizzare la (4.31) per esprimere l'energia cinetica T in funzione delle coordinate lagrangiane. Avremo

$$T = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2 = T_2 + T_1 + T_0 = \underbrace{\frac{1}{2} \sum_{h,k=1}^l a_{hk}(q_1(t), \dots, q_l(t), t) \dot{q}_h \dot{q}_k}_{T_2} + \underbrace{\sum_{h=1}^l b_h(q_1(t), \dots, q_l(t), t) \dot{q}_h}_{T_1} + \underbrace{\frac{1}{2} c(q_1(t), \dots, q_l(t), t)}_{T_0}, \quad (4.33)$$

dove i termini a_{hk} , b_h e c sono dati da¹²

$$a_{hk} = \sum_{i=1}^N m_i \frac{\partial \mathbf{x}_i}{\partial q_h} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k}, \quad (4.34)$$

$$b_h = \sum_{i=1}^N m_i \frac{\partial \mathbf{x}_i}{\partial q_h} \cdot \frac{\partial \mathbf{x}_i}{\partial t}, \quad (4.35)$$

$$c = \sum_{i=1}^N m_i \frac{\partial \mathbf{x}_i}{\partial t} \cdot \frac{\partial \mathbf{x}_i}{\partial t}. \quad (4.36)$$

Per quanto riguarda T_1 , introducendo

$$\tilde{\mathbf{b}}(\tilde{\mathbf{q}}, t) = \begin{pmatrix} b_1(q_1, \dots, q_l, t) \\ b_2(q_1, \dots, q_l, t) \\ \vdots \\ b_l(q_1, \dots, q_l, t) \end{pmatrix},$$

possiamo anche scrivere

$$T_1 = \tilde{\mathbf{b}} \bullet \dot{\tilde{\mathbf{q}}}. \quad (4.37)$$

La parte T_2 può anche essere scritta come

$$T_2 = \frac{1}{2} \tilde{\mathbf{q}}^T \mathbb{A} \tilde{\mathbf{q}}, \quad (4.38)$$

dove

$$\tilde{\mathbf{q}} = \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_l \end{pmatrix}, \quad (4.39)$$

e $\mathbb{A} = \mathbb{A}(q_1(t), \dots, q_l(t), t)$, è una matrice simmetrica $l \times l$, le cui componenti

$$a_{hk} = a_{hk}(q_1, \dots, q_l, t), \quad h, k = 1, \dots, l$$

sono date dalla (4.34).

¹²Si raccomanda di fare il calcolo delle (4.34)-(4.36), e non procedere oltre se non lo si sa fare!

Teorema 4.1.3 *La forma bilineare (4.38) è definita positiva, cioè $T_2 > 0$, e $T_2 = 0$ soltanto se $\tilde{\mathbf{q}} = 0$.*

Dim. Facciamo vedere che $\det \mathbb{A} \neq 0$. Infatti se per assurdo $\det \mathbb{A} = 0$, deve esistere almeno un autovalore nullo, per cui, denotando con $\tilde{\mathbf{w}} \in \mathbb{R}^l$, $\tilde{\mathbf{w}} \neq 0$ il relativo autovettore (non nullo), abbiamo

$$\mathbb{A}\tilde{\mathbf{w}} = 0 \quad \Rightarrow \quad \sum_{h=1}^l a_{kh} w_h = 0, \quad k = 1, 2, \dots, l.$$

Tuttavia, ricordando la (4.34), possiamo anche scrivere

$$\begin{aligned} 0 &= \tilde{\mathbf{w}}^T \mathbb{A} \tilde{\mathbf{w}} = \sum_{h=1}^l \sum_{k=1}^l a_{kh} w_k w_h \\ &= \sum_{h=1}^l \sum_{k=1}^l \left(\sum_{i=1}^N m_i \frac{\partial \mathbf{x}_i}{\partial q_h} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \right) w_k w_h \\ &= \sum_{i=1}^N m_i \left[\left(\sum_{h=1}^l w_h \frac{\partial \mathbf{x}_i}{\partial q_h} \right) \cdot \left(\sum_{k=1}^l w_k \frac{\partial \mathbf{x}_i}{\partial q_k} \right) \right] = \sum_{i=1}^N m_i \left(\sum_{h=1}^l w_h \frac{\partial \mathbf{x}_i}{\partial q_h} \right)^2. \end{aligned}$$

Ma allora, per ogni $i = 1, 2, \dots, N$, il vettore $\sum_{h=1}^l w_h \frac{\partial \mathbf{x}_i}{\partial q_h}$ è nullo, cioè

$$\begin{aligned} 0 &= \left(\sum_{h=1}^l w_h \frac{\partial \mathbf{x}_i}{\partial q_h} \right)^2 \\ &= \left(\sum_{h=1}^l w_h \frac{\partial x_i}{\partial q_h} \right)^2 + \left(\sum_{h=1}^l w_h \frac{\partial y_i}{\partial q_h} \right)^2 + \left(\sum_{h=1}^l w_h \frac{\partial z_i}{\partial q_h} \right)^2. \end{aligned}$$

Dunque tutti gli N vettori, $i = 1, \dots, N$,

$$\begin{aligned} \sum_{h=1}^l \frac{\partial \mathbf{x}_i}{\partial q_h} w_h &= \left(\sum_{h=1}^l w_h \frac{\partial x_i}{\partial q_h} \right) \mathbf{e}_x \\ &\quad + \left(\sum_{h=1}^l w_h \frac{\partial y_i}{\partial q_h} \right) \mathbf{e}_y + \left(\sum_{h=1}^l w_h \frac{\partial z_i}{\partial q_h} \right) \mathbf{e}_z, \end{aligned}$$

devono essere nulli. Esiste quindi una l -upla **non nulla**

$$\tilde{\mathbf{w}} = \begin{pmatrix} w_1 \\ \vdots \\ w_l \end{pmatrix} \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

per cui

$$\begin{aligned}
 w_1 \tilde{\mathbf{u}}_1 + \dots + w_l \tilde{\mathbf{u}}_l &= w_1 \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_1} \\ \frac{\partial \mathbf{x}_2}{\partial q_1} \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_1} \end{pmatrix} + \dots + w_l \begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_l} \\ \frac{\partial \mathbf{x}_2}{\partial q_l} \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_l} \end{pmatrix} \\
 &= \begin{pmatrix} \sum_{h=1}^l \frac{\partial \mathbf{x}_1}{\partial q_h} w_h \\ \sum_{h=1}^l \frac{\partial \mathbf{x}_2}{\partial q_h} w_h \\ \vdots \\ \sum_{h=1}^l \frac{\partial \mathbf{x}_N}{\partial q_h} w_h \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.
 \end{aligned}$$

In altri termini, gli l vettori di \mathbb{R}^{3N} , $\tilde{\mathbf{u}}_j$, $j = 1, 2, \dots, l$, **non sono linearmente indipendenti** e ciò contraddice il teorema 4.1.1. Abbiamo quindi provato che la forma bilineare (4.38) non può essere degenere. Il fatto che sia definita positiva discende banalmente dalla definizione di energia cinetica. Infatti, supponendo di “congelare” i vincoli¹³, si deduce, dall’espressione di T (4.33), che la matrice $\tilde{\mathbf{q}}^T \mathbb{A} \tilde{\mathbf{q}} > 0$, se per $\tilde{\mathbf{q}}$ non nullo. La forma quadratica (4.38) è dunque definita positiva.

□

Osserviamo infine che, sfruttando la (4.37) e la (4.38) possiamo riscrivere l’energia cinetica nella seguente forma compatta

$$T = \frac{1}{2} \tilde{\mathbf{q}}^T \mathbb{A} \tilde{\mathbf{q}} + \tilde{\mathbf{b}} \bullet \tilde{\mathbf{q}} + T_0. \quad (4.40)$$

Esempio 4.1.6 Riferendoci all’esempio 4.1.4, vogliamo adesso determinare la velocità e l’energia cinetica del punto materiale. La velocità del punto materiale P discende dalla (4.29)

$$\dot{\mathbf{x}} = \dot{\theta} \mathbf{u}_\theta + \frac{\partial \mathbf{x}}{\partial t},$$

dove

$$\frac{\partial \mathbf{x}}{\partial t} = -R \frac{d\varphi(t)}{dt} \sin \theta \sin \varphi(t) \mathbf{e}_x + R \frac{d\varphi(t)}{dt} \sin \theta \cos \varphi(t) \mathbf{e}_y,$$

rappresenta la velocità di trascinamento, mentre $\dot{\theta} \mathbf{u}_\theta$ è la velocità rispetto al vincolo.

Lo ricordiamo ancora: $\varphi(t)$ è una funzione nota, e dunque $\frac{d\varphi}{dt}$ è la velocità assegnata di rotazione del vincolo.

L’energia cinetica del punto materiale è

$$T(\theta, \dot{\theta}, t) = \frac{m}{2} |\dot{\mathbf{x}}|^2 = \frac{mR^2}{2} \left(\dot{\theta}^2 + \left(\frac{d\varphi}{dt} \right)^2 \sin^2 \theta \right). \quad (4.41)$$

¹³Per “congelare”, o meglio, immobilizzare i vincoli è sufficiente considerare nulla la velocità di trascinamento, cioè porre $\frac{\partial \mathbf{x}}{\partial t} = 0$.

dove $\frac{mR^2}{2} \left(\left(\frac{d\varphi}{dt} \right)^2 \sin^2 \theta \right)$ è la parte di energia cinetica dovuta al moto del vincolo.

In particolare, confrontando la (4.41) con la (4.33), abbiamo

$$T(\theta, \dot{\theta}, t) = \underbrace{\frac{mR^2}{2} \dot{\theta}^2}_{T_2} + \underbrace{\frac{mR^2}{2} \left(\frac{d\varphi}{dt} \right)^2 \sin^2 \theta}_{T_0}. \quad (4.42)$$

4.1.3 Punto vincolato sulla superficie

Come primo esempio di cinematica di un sistema vincolato vediamo, in dettaglio, il caso di un solo punto materiale vincolato a una superficie fissa (quindi non dipendente dal tempo) che viene assegnata come insieme di livello di una funzione. Pertanto, data $f: W \rightarrow \mathbb{R}$, funzione (o campo scalare) definita su un sottoinsieme aperto W di A , spazio tridimensionale affine euclideo, la superficie è l'insieme dei punti $P \in W$, tali che $f(P) = 0$. In simboli $\{P \in U : f(P) = 0\}$. Supponiamo che in A sia dato un riferimento cartesiano ortonormale $\{O, e_x, e_y, e_z\}$, sì che ogni punto P è identificato dal vettore

$$(P - O) = \mathbf{x}, \quad \text{con} \quad \mathbf{x} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z.$$

La superficie allora si identifica come¹⁴

$$\mathcal{C} = \{ \mathbf{x} \in \mathbb{R}^3 : f(\mathbf{x}) = 0 \}. \quad (4.43)$$

L'insieme $\mathcal{C} \subset \mathbb{R}^3$, la superficie appunto, è dunque lo spazio delle configurazioni, che ovviamente assumiamo non vuoto¹⁵. In particolare, se f è di classe¹⁶ C^2 la superficie si dice **regolare** (o **non singolare**) in $\mathbf{x}_0 \in \mathcal{C}$, se¹⁷ $\nabla f(\mathbf{x}_0) \neq 0$, ovvero se

$$|\nabla f(\mathbf{x}_0)|^2 = \left(\frac{\partial f(\mathbf{x}_0)}{\partial x} \right)^2 + \left(\frac{\partial f(\mathbf{x}_0)}{\partial y} \right)^2 + \left(\frac{\partial f(\mathbf{x}_0)}{\partial z} \right)^2 \neq 0. \quad (4.44)$$

La superficie si dirà poi **globalmente regolare** (o **globalmente non singolare**) se è regolare in ogni punto di \mathcal{C} . Supponiamo quindi di aver a che fare con una superficie globalmente regolare. Tale ipotesi consente l'applicazione del teorema della funzione implicita per ogni $\mathbf{x} \in \mathcal{C}$. In base a tale teorema $f(x, y, z) = 0$ è **localmente** il grafico di una funzione. Se, ad esempio, $\frac{\partial f}{\partial z} \neq 0$, in \mathbf{x}_0 , allora, sul piano (x, y) esiste un intorno aperto U di (x_0, y_0) , per cui la superficie può localmente scriversi come grafico, $z = z(x, y)$. In altri termini, il teorema del Dini consente di individuare un'applicazione che ai parametri $(q_1, q_2) \in U$, con $U \subset \mathbb{R}^2$, intorno aperto di (x_0, y_0) ,

¹⁴Anche adesso con \mathbf{x} intendiamo la terna (x, y, z) . Quindi $f(\mathbf{x}) = f(x, y, z)$.

¹⁵Richiesta 1 del paragrafo 4.1. Infatti, se per esempio avessimo

$$f(x, y, z) = x^2 + y^2 + z^2 + 16,$$

è chiaro che l'insieme $\mathcal{C} \subset \mathbb{R}^3$, costituito dei punti (x, y, z) tali che $f(x, y, z) = 0$, sarebbe vuoto.

¹⁶Proprietà di regolarità del vincolo (richiesta 2 del paragrafo 4.1).

¹⁷Richiesta 3 del paragrafo 4.1.

associa il vettore $\mathbf{x}(q_1, q_2)$

$$(q_1, q_2) \rightarrow \mathbf{x}(q_1, q_2) = q_1 \mathbf{e}_x + q_2 \mathbf{e}_y + z(q_1, q_2) \mathbf{e}_z \Leftrightarrow \begin{cases} x = q_1, \\ y = q_2, \\ z = z(q_1, q_2). \end{cases}$$

Questo significa che, localmente¹⁸, la superficie può essere espressa in forma parametrica mediante i parametri¹⁹ (q_1, q_2) .

Si introducono i due vettori

$$\mathbf{u}_i = \frac{\partial \mathbf{x}}{\partial q_i} = \frac{\partial x}{\partial q_i} \mathbf{e}_x + \frac{\partial y}{\partial q_i} \mathbf{e}_y + \frac{\partial z(q_1, q_2)}{\partial q_i} \mathbf{e}_z, \quad i = 1, 2, \quad (4.45)$$

detti **vettori tangenti alla superficie**. Nel caso specifico essi sono

$$\mathbf{u}_1 = \mathbf{e}_x + \frac{\partial z(q_1, q_2)}{\partial q_1} \mathbf{e}_z, \quad \mathbf{u}_2 = \mathbf{e}_y + \frac{\partial z(q_1, q_2)}{\partial q_2} \mathbf{e}_z,$$

dove, sfruttando ancora il teorema della funzione implicita,

$$\frac{\partial z(q_1, q_2)}{\partial q_1} = -\frac{\left(\frac{\partial f}{\partial x}\right)}{\left(\frac{\partial f}{\partial z}\right)}, \quad \frac{\partial z(q_1, q_2)}{\partial q_2} = -\frac{\left(\frac{\partial f}{\partial y}\right)}{\left(\frac{\partial f}{\partial z}\right)}.$$

E' facile provare che $\mathbf{u}_1, \mathbf{u}_2$ sono linearmente indipendenti. Infatti, la matrice 2×3 formata dalle componenti di $\mathbf{u}_1, \mathbf{u}_2$,

$$\begin{pmatrix} 1 & 0 & \frac{\partial z}{\partial q_1} \\ 0 & 1 & \frac{\partial z}{\partial q_2} \end{pmatrix},$$

ha rango massimo

Nota 4.1.6 *Allo stesso risultato saremmo potuti giungere anche mostrando che $\mathbf{u}_1 \wedge \mathbf{u}_2 \neq 0$. Infatti,*

$$\begin{aligned} \mathbf{u}_1 \wedge \mathbf{u}_2 &= -\frac{\partial z}{\partial q_1} \mathbf{e}_x - \frac{\partial z}{\partial q_2} \mathbf{e}_y + \mathbf{e}_z \\ &= \frac{1}{\left(\frac{\partial f}{\partial z}\right)} \left[\frac{\partial f}{\partial x} \mathbf{e}_x + \frac{\partial f}{\partial y} \mathbf{e}_y + \frac{\partial f}{\partial z} \mathbf{e}_z \right] = \frac{1}{\left(\frac{\partial f}{\partial z}\right)} \nabla f. \end{aligned}$$

Pertanto $\mathbf{u}_1 \wedge \mathbf{u}_2$ è parallelo a ∇f , che è non nullo dal momento che la superficie è regolare.

¹⁸E' importante osservare che, in generale, la parametrizzazione non è globale.

¹⁹Ovviamente, a questo livello, q_1 e q_2 altro non sono che le variabili x, y .

Inoltre abbiamo

$$\mathbf{u}_1 \cdot \nabla f = \left[\mathbf{e}_x - \frac{\left(\frac{\partial f}{\partial x} \right)}{\left(\frac{\partial f}{\partial z} \right)} \mathbf{e}_z \right] \cdot \left[\frac{\partial f}{\partial x} \mathbf{e}_x + \frac{\partial f}{\partial y} \mathbf{e}_y + \frac{\partial f}{\partial z} \mathbf{e}_z \right] = 0,$$

ed analogamente $\mathbf{u}_2 \cdot \nabla f = 0$.

Quindi, volendo generalizzare, il teorema della funzione implicita garantisce che, per ogni punto della superficie regolare, esiste un intorno in cui la superficie è rappresentabile come l'immagine di un aperto U di \mathbb{R}^2 . Ovvero esistono tre funzioni $x(q_1, q_2)$, $y(q_1, q_2)$ e $z(q_1, q_2)$ di due variabili (le variabili lagrangiane o parametri lagrangiani) $(q_1, q_2) \in U \subset \mathbb{R}^2$, tali che

$$f(x(q_1, q_2), y(q_1, q_2), z(q_1, q_2)) = 0, \quad \text{per ogni } (q_1, q_2) \in U.$$

Le curve

$$q_1 \rightarrow (x(q_1, q_2), y(q_1, q_2), z(q_1, q_2)), \quad q_2 = \text{costante},$$

e

$$q_2 \rightarrow (x(q_1, q_2), y(q_1, q_2), z(q_1, q_2)), \quad q_1 = \text{costante}.$$

formano il reticolo di linee coordinate sulla superficie $f = 0$. Introducendo poi i vettori tangenti alla superficie

$$\mathbf{u}_i = \frac{\partial x(q_1, q_2)}{\partial q_i} \mathbf{e}_x + \frac{\partial y(q_1, q_2)}{\partial q_i} \mathbf{e}_y + \frac{\partial z(q_1, q_2)}{\partial q_i} \mathbf{e}_z, \quad i = 1, 2, \quad (4.46)$$

è facile provare che sono linearmente indipendenti e che $\mathbf{u}_i \cdot \nabla f = 0$, $i = 1, 2$.

Nota 4.1.7 *E' opportuno rimarcare ancora che la parametrizzazione della superficie non è univoca. Per intendersi, anziché usare i parametri $(q_1, q_2) \in U \subset \mathbb{R}^2$, potremmo descrivere la superficie utilizzando i parametri $(s_1, s_2) \in W \subset \mathbb{R}^2$, ammesso che esista un'applicazione biunivoca che ci consenta di scrivere*

$$q_1 = q_1(s_1, s_2), \quad q_2 = q_2(s_1, s_2), \quad (4.47)$$

e le loro inverse: $s_1 = s_1(q_1, q_2)$, $s_2 = s_2(q_1, q_2)$. Utilizzando le variabili (s_1, s_2) , la superficie viene vista come l'immagine dell'aperto W , attraverso le tre funzioni $x(s_1, s_2)$, $y(s_1, s_2)$ e $z(s_1, s_2)$. L'invertibilità della trasformazione (4.47) è garantita dalla condizione $\det \mathbb{J} \neq 0$, dove, in questo caso, \mathbb{J} è la matrice jacobiana

$$\mathbb{J} = \begin{pmatrix} \frac{\partial q_1}{\partial s_1} & \frac{\partial q_1}{\partial s_2} \\ \frac{\partial q_2}{\partial s_1} & \frac{\partial q_2}{\partial s_2} \end{pmatrix}. \quad (4.48)$$

Definizione 4.1.5 Si definisce **piano tangente**, $\Pi(P_0)$, alla superficie nel punto²⁰ P_0 , il piano per P_0 i cui punti sono ortogonali a $\nabla f(P_0)$. In simboli

$$\Pi(P_0) = \{Q \in A : (Q - P_0) \cdot \nabla f(P_0) = 0\}. \quad (4.49)$$

²⁰La superficie è regolare in P .

Definizione 4.1.6 Si definisce *spazio normale* alla superficie \mathcal{S} nel punto P_0 , e lo si denota con $\mathcal{N}(P_0)$, il complemento ortogonale dello spazio tangente, ovvero il sottospazio di²¹ V formato da tutti i vettori ortogonali a $\Pi(P_0)$,

$$\mathcal{N}(P_0) = \{v \in V, : v = \lambda \nabla f, \text{ con } \lambda \in \mathbb{R}\}. \quad (4.50)$$

In particolare se il punto P_0 è identificato dal vettore $(P_0 - O) = x_0$, avremo che il punto Q , identificato dal vettore $(Q - O) = x$, giace sul piano tangente se $(x - x_0) \cdot \nabla f(x_0) = 0$. Evidentemente, una definizione equivalente di piano tangente può essere anche data in termini dei due vettori u_1 e u_2 ,

$$\Pi(P_0) = \{Q \in A : (Q - P_0) = \alpha_1 u_1 + \alpha_2 u_2, \text{ con } \alpha_i \in \mathbb{R}, i = 1, 2\},$$

mentre una definizione equivalente di spazio normale è

$$\mathcal{N}(P_0) = \{v \in V, : v = \lambda (u_1(P_0) \wedge u_2(P_0)), \text{ con } \lambda \in \mathbb{R}\}.$$

Possiamo quindi concludere che ogni spostamento virtuale relativo alla generica configurazione x è un vettore del piano tangente, cioè

$$\delta x = u_1 \delta q_1 + u_2 \delta q_2, \quad (4.51)$$

dove $\delta q_1, \delta q_2$ sono numeri reali qualsiasi.

Definiamo poi, in ogni punto della superficie, la matrice simmetrica i cui elementi sono i prodotti scalari dei vettori tangenti (4.46)

$$\mathbb{A} = \begin{pmatrix} u_1 \cdot u_1 & u_2 \cdot u_1 \\ u_1 \cdot u_2 & u_2 \cdot u_2 \end{pmatrix}, \quad (4.52)$$

usualmente denotata²² $\begin{pmatrix} E & F \\ F & G \end{pmatrix}$, che viene detta **prima forma quadratica fondamentale della superficie**. Osserviamo che la matrice \mathbb{A} non è altro che la matrice metrica (1.8) relativa alla base $\{u_1, u_2\}$. Infatti, la restrizione del prodotto scalare al piano tangente dà luogo ad una forma bilineare simmetrica definita positiva. In altre parole, se $\pi_1 = \alpha_1 u_1 + \alpha_2 u_2$, e $\pi_2 = \beta_1 u_1 + \beta_2 u_2$, sono due vettori del piano tangente, allora

$$\begin{aligned} \pi_1 \cdot \pi_2 &= (\alpha_1 u_1 + \alpha_2 u_2) \cdot (\beta_1 u_1 + \beta_2 u_2) = \\ &= \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\ &= \begin{pmatrix} \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}. \end{aligned} \quad (4.53)$$

²¹Ricordiamo che V è lo spazio vettoriale che soggiace allo spazio affine euclideo A .

²²Si faccia attenzione al fatto che le componenti della matrice E, F e G sono, in generale, funzioni di (q_1, q_2) , ovvero $E = E(q_1, q_2)$, $F = F(q_1, q_2)$ e $G = G(q_1, q_2)$.

La forma è definita positiva proprio perché il prodotto scalare è definito positivo. Infatti, per ogni vettore $\boldsymbol{\pi} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 \neq 0$ del piano tangente potremo scrivere (assumendo, per esempio, $\alpha_2 \neq 0$)

$$\boldsymbol{\pi} \cdot \boldsymbol{\pi} = \alpha_2^2 (E\xi^2 + 2F\xi + G), \quad \text{con } \xi = \frac{\alpha_1}{\alpha_2}.$$

Ora il trinomio in ξ è sempre positivo dal momento che il discriminante è

$$(F^2 - GE) = -|\mathbf{u}_1 \wedge \mathbf{u}_2|^2 < 0,$$

(v. formula (4.57)).

Nota 4.1.8 Osserviamo che la forma bilineare (4.53) è indipendente dalla parametrizzazione della superficie. Infatti, rifacendoci alla nota 4.1.7, se al posto di (q_1, q_2) facciamo uso della coppia (s_1, s_2) , i cui vettori tangenti sono $\mathbf{h}_1 = \frac{\partial \mathbf{x}}{\partial s_1}$, $\mathbf{h}_2 = \frac{\partial \mathbf{x}}{\partial s_2}$, la matrice della forma bilineare, rispetto alla nuova base $\mathbf{h}_1, \mathbf{h}_2$, è data da

$$\hat{\mathbb{M}} = \begin{pmatrix} \hat{E} & \hat{F} \\ \hat{F} & \hat{G} \end{pmatrix} = \begin{pmatrix} \mathbf{h}_1 \cdot \mathbf{h}_1 & \mathbf{h}_1 \cdot \mathbf{h}_2 \\ \mathbf{h}_1 \cdot \mathbf{h}_2 & \mathbf{h}_1 \cdot \mathbf{h}_2 \end{pmatrix}.$$

Si può facilmente dimostrare che

$$\begin{pmatrix} \hat{E} & \hat{F} \\ \hat{F} & \hat{G} \end{pmatrix} = \mathbb{J}^T \begin{pmatrix} E & F \\ F & G \end{pmatrix} \mathbb{J}, \quad (4.54)$$

dove \mathbb{J} è data dalla (4.48). In particolare, dati due vettori $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$ sullo spazio tangente, tali che

$$\boldsymbol{\pi}_1 \rightarrow \begin{cases} \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 \\ \delta_1 \mathbf{h}_1 + \delta_2 \mathbf{h}_2 \end{cases} \quad \text{con } \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \mathbb{J} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix},$$

$$\boldsymbol{\pi}_2 \rightarrow \begin{cases} \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 \\ \gamma_1 \mathbf{h}_1 + \gamma_2 \mathbf{h}_2 \end{cases} \quad \text{con } \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \mathbb{J} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix},$$

mostriamo che $\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2$ è indipendente dalla base prescelta. Infatti, indicando con $(\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2)_{(\mathbf{u}_1, \mathbf{u}_2)}$ la forma bilineare espressa nella base $\mathbf{u}_1, \mathbf{u}_2$, e con $(\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2)_{(\mathbf{h}_1, \mathbf{h}_2)}$ l'espressione della forma rispetto a $\mathbf{h}_1, \mathbf{h}_2$, proviamo che

$$(\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2)_{(\mathbf{u}_1, \mathbf{u}_2)} = (\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2)_{(\mathbf{h}_1, \mathbf{h}_2)}.$$

Dalla²³ (4.53)

$$\begin{aligned}
 (\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2)_{(\mathbf{u}_1, \mathbf{u}_2)} &= \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\
 &= \begin{pmatrix} \delta_1 & \delta_2 \end{pmatrix} \mathbb{J}^T \begin{pmatrix} E & F \\ F & G \end{pmatrix} \mathbb{J} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \\
 &\stackrel{(4.54)}{=} \begin{pmatrix} \delta_1 & \delta_2 \end{pmatrix} \begin{pmatrix} \hat{E} & \hat{F} \\ \hat{F} & \hat{G} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = (\boldsymbol{\pi}_1 \cdot \boldsymbol{\pi}_2)_{(\mathbf{h}_1, \mathbf{h}_2)}.
 \end{aligned}$$

Ora, i parametri lagrangiani (q_1, q_2) saranno, in generale, funzioni del tempo. Scriveremo pertanto

$$\mathbf{x}(t) = x(q_1(t), q_2(t)) \mathbf{e}_x + y(q_1(t), q_2(t)) \mathbf{e}_y + z(q_1(t), q_2(t)) \mathbf{e}_z.$$

Il vettore velocità del punto materiale è

$$\begin{aligned}
 \dot{\mathbf{x}}(t) &= \sum_{i=1,2} \underbrace{\left(\frac{\partial x}{\partial q_i} \mathbf{e}_x + \frac{\partial y}{\partial q_i} \mathbf{e}_y + \frac{\partial z}{\partial q_i} \mathbf{e}_z \right)}_{\mathbf{u}_i} \dot{q}_i(t) \\
 &= \dot{q}_1(t) \mathbf{u}_1 + \dot{q}_2(t) \mathbf{u}_2.
 \end{aligned} \tag{4.55}$$

Quindi $\dot{\mathbf{x}}(t)$ è un vettore dello spazio tangente e le sue componenti, rispetto alla base $\mathbf{u}_1, \mathbf{u}_2$, sono date dalla coppia $\begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \end{pmatrix}$.

L'energia cinetica del punto è $T = \frac{m}{2} |\dot{\mathbf{x}}|^2$. Siccome $\dot{\mathbf{x}}$ è un vettore del piano tangente facciamo uso della forma bilineare (4.53) per calcolarne la norma, cioè

$$\begin{aligned}
 T &= \frac{m}{2} |\dot{\mathbf{x}}|^2 = \frac{m}{2} \begin{pmatrix} \dot{q}_1 & \dot{q}_2 \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} \\
 &= \frac{m}{2} (E\dot{q}_1^2 + 2F\dot{q}_1\dot{q}_2 + G\dot{q}_2^2).
 \end{aligned} \tag{4.56}$$

Concludiamo mostrando come, nel caso di un solo punto materiale vincolato su una superficie regolare, il teorema 4.1.3 sia facilmente dimostrabile. Infatti

$$\begin{aligned}
 \det \mathbb{A} &= EG - F^2 = |\mathbf{u}_1|^2 |\mathbf{u}_2|^2 - (\mathbf{u}_2 \cdot \mathbf{u}_1)^2 \\
 &= |\mathbf{u}_1|^2 |\mathbf{u}_2|^2 (1 - \cos^2 \theta) = |\mathbf{u}_1|^2 |\mathbf{u}_2|^2 \sin^2 \theta = |\mathbf{u}_1 \wedge \mathbf{u}_2|^2 > 0,
 \end{aligned} \tag{4.57}$$

dal momento che \mathbf{u}_1 e \mathbf{u}_2 non sono paralleli.

²³Ricordiamo che

$$\underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}}_{\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix}}^T = \underbrace{\begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}}_{\begin{pmatrix} \delta_1 & \delta_2 \end{pmatrix}}^T \mathbb{J}^T$$

Capitolo 5

Le equazioni di moto per sistemi vincolati

Sia $\{P_i, m_i, i = 1, \dots, N\}$, un sistema di punti materiali, che supponiamo soggetto a un sistema di m vincoli olonomi

$$f_k(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = 0, \quad k = 1, \dots, m. \quad (5.1)$$

Supponiamo poi che il sistema di punti materiali sia soggetto a delle forze, indipendentemente dal fatto che il suo moto debba risultare a priori vincolato. Indicheremo con \mathbf{F}_i la risultante delle forze applicate al punto P_i . Chiameremo queste **forze direttamente applicate**.

Sotto l'azione di tali forze il moto del sistema retto dalle equazioni $m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i$, $i = 1, \dots, N$, non risulterà in genere incompatibile con le equazioni vincolari (5.1). Per esempio si pensi a un punto appoggiato su un piano orizzontale: in questo caso la forza direttamente applicata è la forza peso, e il punto abbandonerà il piano orizzontale durante il moto sotto l'azione di questa forza, violando così il vincolo che gli impone a rimanere sul piano orizzontale. Se vogliamo salvaguardare i vincoli e le equazioni di Newton, dobbiamo supporre che i vincoli "esercitino" delle forze \mathbf{R}_i sui punti del sistema in modo che le soluzioni delle equazioni

$$m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i + \mathbf{R}_i, \quad i = 1, \dots, N, \quad (5.2)$$

risultino compatibili con le equazioni vincolari (5.1). Chiameremo le forze \mathbf{R}_i , dovute ai vincoli, **reazioni vincolari**.

Le (5.1) assieme alle (5.2) formano un sistema di $3N + m$ equazioni da cui dovremmo ricavare sia le incognite di moto (le $3N$ coordinate dei punti P_i) sia le $3N$ componenti delle reazioni vincolari \mathbf{R}_i . Poiché $m < 3N$, il sistema risulta indeterminato, ovvero non possiamo da queste sole "informazioni" determinare la soluzione completa del problema del moto vincolato (a meno di non caratterizzare maggiormente le reazioni \mathbf{R}_i).

5.1 Dinamica di un punto vincolato sulla superficie.

Iniziamo esaminando il caso più semplice di un solo punto materiale (P, m) vincolato a muoversi su una superficie fissa regolare (vincolo olonomo fisso) data dalla (4.43). Come abbiamo già detto le equazioni di moto

$$m \ddot{\mathbf{x}} = \mathbf{F} + \mathbf{R}, \quad (5.3)$$

dovranno contenere la forza \mathbf{R} incognita, il cui scopo è quello di far sì che la soluzione di (5.3) soddisfi, ad ogni istante t , la condizione vincolare $f(x, y, z) = 0$. Ciò ci dà un sistema indeterminato, avendo solo quattro equazioni per determinare sei incognite, le coordinate di P e le componenti di \mathbf{R} .

L'indeterminatezza del problema così formulato è dovuta al fatto che non abbiamo specificato "come" il vincolo, nella sua realizzazione fisica, intervenga per assicurarci il soddisfacimento della condizione $f(x, y, z) = 0$. Pensiamo per esempio, ad un oggetto appoggiato su di un tavolo: il piano del tavolo impedisce all'oggetto di cadere, in termini più formali, genera quella forza necessaria a contrastare l'effetto della forza peso. Quindi una componente della forza vincolare \mathbf{R} è tale da impedire al corpo di abbandonare il vincolo. Ma cosa succede se tentiamo di muovere l'oggetto sul piano stesso? Ogni vincolo "reale" si opporrà a tale moto: è l'effetto che va sotto il nome di **attrito**. Questo effetto viene schematizzato tramite una **componente di \mathbf{R} che ha sempre direzione e verso opposti alla velocità del corpo**. Ora, siccome la velocità del punto materiale giace sul piano tangente, la componente di \mathbf{R} dovuta all'attrito starà sul piano tangente. Potremo quindi effettuare questa scomposizione

$$\mathbf{R} = \mathbf{R}_{\parallel} + \mathbf{R}_{\perp}, \quad (5.4)$$

dove \mathbf{R}_{\parallel} **giace nel piano tangente** (e dunque sarà esprimibile come combinazione lineare di \mathbf{u}_1 e \mathbf{u}_2) mentre \mathbf{R}_{\perp} **giace nello spazio tangente**, per cui, ricordando la (4.50),

$$\mathbf{R}_{\perp} = \lambda \nabla f, \quad (5.5)$$

dove $\lambda \in \mathbb{R}$, è un parametro incognito, detto **moltiplicatore di Lagrange**, il cui valore dipenderà dalla posizione e dalla velocità del punto.

La forza \mathbf{R}_{\parallel} è causata dall'attrito. Essa viene generalmente espressa in termini di altre quantità dinamiche (es. velocità, \mathbf{R}_{\perp} , etc.) e di parametri fisici (coefficienti di attrito, etc.). Ovviamente, per \mathbf{R}_{\parallel} esistono vari modelli di attrito¹ o, come si usa dire in meccanica, varie equazioni costitutive. Una volta selezionato il modello il vettore \mathbf{R}_{\parallel} viene espressa (in maniera più o meno complicata) in funzione di altre quantità.

La forza \mathbf{R}_{\perp} rappresenta la componente della forza vincolare che la superficie (il vincolo) esercita sulla particella per mantenerlo su di essa. Notiamo che, in generale, \mathbf{R}_{\perp} è incognita (cioè, data la (5.5), λ è incognito). Infatti, a differenza di \mathbf{R}_{\parallel} , non esiste alcuna espressione costitutiva per \mathbf{R}_{\perp} , dal momento che essa deve poter assumere qualunque valore necessario per non far allontanare il punto materiale dalla superficie.

In molti casi possiamo ridurre l'intensità di \mathbf{R}_{\parallel} (componente di \mathbf{R} dovuta all'attrito), e pensare ad un modello ideale in cui questa resistenza al moto viene a mancare

¹Ad esempio, il modello di attrito viscoso prevede che $\mathbf{R}_{\parallel} = -\eta \dot{\mathbf{x}}$, dove η viene detto coefficiente di attrito viscoso.

del tutto. Chiameremo **ideale** o **liscio** un vincolo per cui

$$\mathbf{R}_{\parallel} = 0. \tag{5.6}$$

La caratterizzazione del vincolo liscio è immediata. Poiché, come detto, il vettore \mathbf{R}_{\parallel} giace, istante per istante, sul piano tangente, avremo che il vincolo è liscio se e solo se

$$\mathbf{R} \cdot \delta \mathbf{x} = 0, \text{ per ogni } \delta \mathbf{x} \in \Pi(P), \tag{5.7}$$

e di conseguenza², ricordando la definizione 4.1.6,

$$\mathbf{R} \in \mathcal{N}(P) \Leftrightarrow \mathbf{R} = \lambda \nabla f. \tag{5.8}$$

Nota 5.1.1 *E' opportuno chiarire il significato fisico della (5.7). Se consideriamo un intervallo di tempo δt "molto piccolo", il lavoro δL infinitesimo, detto anche **lavoro virtuale**, fatto dalla forza \mathbf{R} durante tale intervallo temporale è (si ricordi la definizione 3.2.6)*

$$\delta L = W \delta t = \mathbf{R} \cdot \dot{\mathbf{x}} \delta t.$$

Se adesso applichiamo la (4.55) si ottiene

$$\begin{aligned} \delta L &= \mathbf{R} \cdot \left[\mathbf{u}_1 \underbrace{\dot{q}_1(t) \delta t}_{\delta q_1} + \mathbf{u}_2 \underbrace{\dot{q}_2(t) \delta t}_{\delta q_2} \right] \\ &= \mathbf{R} \cdot [\mathbf{u}_1 \delta q_1 + \mathbf{u}_2 \delta q_2] = \mathbf{R} \cdot \delta \mathbf{x}. \end{aligned}$$

Dunque, in termini fisici la (5.7) può esprimersi così: **il lavoro della forza vincolare \mathbf{R} è nullo per ogni spostamento virtuale.**

Quindi, nel caso in cui la superficie sia liscia, l'equazione di moto (5.3) e l'equazione del vincolo danno luogo ad un sistema, dette equazioni di Lagrange di prima specie,

$$\begin{cases} m \ddot{\mathbf{x}} = \mathbf{F} + \lambda \nabla f, \\ f(\mathbf{x}) = 0, \end{cases} \tag{5.9}$$

di quattro equazioni nelle quattro incognite $\mathbf{x} = (x, y, z)$ e λ .

Vediamo adesso come si determina λ nel caso del singolo punto materiale vincolato su una superficie regolare. Partendo dalla (5.9)₁, si ha

$$\lambda |\nabla f|^2 = (m \ddot{\mathbf{x}} - \mathbf{F}) \cdot \nabla f, \tag{5.10}$$

da cui

$$\lambda = \frac{(m \ddot{\mathbf{x}} - \mathbf{F}) \cdot \nabla f}{|\nabla f|^2},$$

dove il termine $\ddot{\mathbf{x}} \cdot \nabla f$ può essere ulteriormente elaborato. Considerando il vettore posizione \mathbf{x} come funzione del tempo³ $\mathbf{x} = \mathbf{x}(t)$, avremo

$$f(\mathbf{x}(t)) = 0, \Rightarrow \frac{df(\mathbf{x}(t))}{dt} = 0, \Leftrightarrow \frac{d}{dt} f(x(t), y(t), z(t)) = 0,$$

²Se vogliamo anche la (5.8) è un'assunzione costitutiva: si dà a priori la forma della forza vincolare.

³In altri termini, pur sapendo che $\mathbf{x} = \mathbf{x}(q_1(t), q_2(t))$, considerare $\mathbf{x} = \mathbf{x}(t)$ significa "dimenticarci" della "mediazione" di q_1 e q_2 e leggere \mathbf{x} direttamente come funzione del tempo.

da cui, ricordando la regola di derivazione di funzione composta,

$$\frac{\partial f}{\partial x} \dot{x} + \frac{\partial f}{\partial y} \dot{y} + \frac{\partial f}{\partial z} \dot{z} = 0, \Leftrightarrow \nabla f \cdot \dot{\mathbf{x}} = 0 \Rightarrow \frac{d}{dt} (\nabla f \cdot \dot{\mathbf{x}}) = 0.$$

Svolgendo la derivata rispetto al tempo si ottiene

$$\nabla f \cdot \ddot{\mathbf{x}} + \left(\frac{d \nabla f}{dt} \right) \cdot \dot{\mathbf{x}} = 0, \Rightarrow \nabla f \cdot \ddot{\mathbf{x}} = - \left(\frac{d \nabla f}{dt} \right) \cdot \dot{\mathbf{x}}. \quad (5.11)$$

Dobbiamo quindi calcolare $\frac{d \nabla f}{dt}$

$$\begin{aligned} \frac{d \nabla f}{dt} &= \frac{d}{dt} \left[\frac{\partial f(x(t), y(t), z(t))}{\partial x} \mathbf{e}_x \right. \\ &\quad \left. + \frac{\partial f(x(t), y(t), z(t))}{\partial y} \mathbf{e}_y + \frac{\partial f(x(t), y(t), z(t))}{\partial z} \mathbf{e}_z \right] \\ &= \left[\frac{\partial^2 f}{\partial x^2} \dot{x} + \frac{\partial^2 f}{\partial y \partial x} \dot{y} + \frac{\partial^2 f}{\partial z \partial x} \dot{z} \right] \mathbf{e}_x + \\ &\quad \left[\frac{\partial^2 f}{\partial x \partial y} \dot{x} + \frac{\partial^2 f}{\partial y^2} \dot{y} + \frac{\partial^2 f}{\partial z \partial y} \dot{z} \right] \mathbf{e}_y + \\ &\quad \left[\frac{\partial^2 f}{\partial x \partial z} \dot{x} + \frac{\partial^2 f}{\partial y \partial z} \dot{y} + \frac{\partial^2 f}{\partial z^2} \dot{z} \right] \mathbf{e}_z. \end{aligned}$$

Per cui, introducendo la matrice $\mathbb{H}(f)$ delle derivate seconde della f , detta **matrice hessiana**,

$$\mathbb{H}(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}, \quad (5.12)$$

si ottiene

$$\frac{d \nabla f}{dt} = \underbrace{\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}}_{\mathbb{H}(f)} \underbrace{\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix}}_{\dot{\mathbf{x}}}.$$

Di conseguenza

$$\left(\frac{d \nabla f}{dt} \right) \cdot \dot{\mathbf{x}} = (\mathbb{H}(f) \dot{\mathbf{x}}) \cdot \dot{\mathbf{x}}, \stackrel{(5.11)}{\Rightarrow} \nabla f \cdot \ddot{\mathbf{x}} = -\dot{\mathbf{x}} \cdot \mathbb{H}(f) \dot{\mathbf{x}}.$$

Abbiamo quindi determinato λ

$$\lambda = - \frac{m \dot{\mathbf{x}} \cdot \mathbb{H}(f) \dot{\mathbf{x}} + \mathbf{F} \cdot \nabla f}{|\nabla f|^2}, \quad (5.13)$$

in termini della forza \mathbf{F} , della velocità $\dot{\mathbf{x}}$, e delle caratteristiche geometriche del vincolo.

L'estensione al caso di una superficie mobile (e quindi dipendente dal tempo) è immediata. Anche in questo caso la definizione di vincolo liscio è data dalla (5.6) o, equivalentemente, dalla (5.7). Ma, anche se il "lavoro virtuale", ovvero il prodotto scalare della reazione con un qualsiasi $\delta\mathbf{x}$ del piano tangente è nullo, la forza vincolare \mathbf{R} può effettuare un lavoro effettivo non nullo.

Per chiarire questo punto si pensi a un corpo che si muova sul piano di un ascensore in movimento. Ovviamente la condizione di "vincolo liscio" deve solo tener conto che tra il piano dell'ascensore e il corpo in moto non c'è attrito, indipendentemente dal fatto che l'ascensore stia salendo, scendendo o sia fermo. Quindi la reazione vincolare deve risultare ortogonale al piano dell'ascensore (condizione garantita dalla (5.7)). Tuttavia, quando l'ascensore sta salendo (o scendendo) la reazione vincolare compierà un lavoro "effettivo" non nullo (un ascensore che sale fa aumentare l'energia potenziale del punto, anche se questo sta fermo sul piano dell'ascensore).

Questo, come vedremo, ha per conseguenza che, nel caso dei vincoli mobili, non si ha, in generale, la conservazione dell'energia meccanica anche in presenza di forze di tipo conservativo.

5.2 L'equazione simbolica della dinamica

La generalizzazione della definizione di vincolo liscio si ottiene in analogia con la (5.7). Sia, al solito, $\{P_i, m_i\}$, $i = 1, \dots, N$, un sistema di punti materiali soggetti agli m vincoli olonomi (5.1). Indichiamo con $\tilde{\mathbf{R}} \in \mathbb{R}^{3N}$, il vettore delle N reazioni vincolari

$$\tilde{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_1 \\ \vdots \\ \mathbf{R}_N \end{pmatrix}, \quad (5.14)$$

intendendo con \mathbf{R}_i , $i = 1, 2, \dots, N$, la forza vincolare (o meglio reazione vincolare) che agisce sull' i -esimo punto generata al complesso degli m vincoli. Anche in questo caso, il sistema formato dalle (5.2) e (5.1) è indeterminato, contando $6N$ incognite contro le $3N + m$ equazioni, a meno che non si facciano ipotesi costitutive sulle reazioni vincolari. L'ipotesi più semplice è quella di vincoli lisci, che così definiamo in generale:

Definizione 5.2.1 Diremo che i vincoli sono *lisci* quando $\tilde{\mathbf{R}} \bullet \delta\tilde{\mathbf{x}} = 0$, per ogni spostamento virtuale $\delta\tilde{\mathbf{x}}$, cioè quando le reazioni vincolari \mathbf{R}_i soddisfano la condizione

$$\sum_{i=1}^N \mathbf{R}_i \cdot \delta\mathbf{x}_i = 0, \quad \text{per ogni sistema di } (\delta\mathbf{x}_1, \dots, \delta\mathbf{x}_N). \quad (5.15)$$

Riferendoci alla nota 5.1.1, esprimeremo la (5.15) dicendo che **il lavoro virtuale delle reazioni vincolari è nullo per ogni sistema di spostamenti virtuali**.

Vediamo come l'assunzione di vincoli lisci ci permetta di ridurre ad m il numero delle incognite dovute alle reazioni vincolari. Infatti, la (5.15) implica che $\tilde{\mathbf{R}} \in \mathcal{N}$, dove,

lo ricordiamo, \mathcal{N} è il sottospazio la cui base è $\{\tilde{\nabla}f_1, \tilde{\nabla}f_2, \dots, \tilde{\nabla}f_m\}$, cioè

$$\tilde{\mathbf{R}} = \sum_{k=1}^m \lambda_k \tilde{\nabla}f_k, \quad (5.16)$$

che, tradotta nelle singole “componenti” di $\tilde{\mathbf{R}}$ significa

$$\mathbf{R}_i = \sum_{k=1}^m \lambda_k \frac{\partial f_k}{\partial \mathbf{x}_i}, \quad i = 1, \dots, N, \quad (5.17)$$

dove $\frac{\partial f_k}{\partial \mathbf{x}_i}$ è definito dalla (4.7) e dove λ_k , $k = 1, \dots, m$, sono parametri arbitrari (moltiplicatori di Lagrange).

Sostituendo le espressioni (5.17) nelle equazioni di moto (5.2) e considerando anche le (5.1), otteniamo il seguente sistema, dette ancora **equazioni di Lagrange di prima specie**,

$$\begin{cases} m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i + \sum_{k=1}^m \lambda_k \frac{\partial f_k}{\partial \mathbf{x}_i}, & i = 1, \dots, N, \\ f_k(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = 0, & k = 1, \dots, m. \end{cases} \quad (5.18)$$

Lo ribadiamo ancora: le N equazioni (5.18)₁ sono valide solo nell’ipotesi di vincoli lisci.

Il processo di eliminazione delle λ_k in (5.18), possibile in linea di principio, è, nel caso di $N > 1$ punti materiali, algebricamente molto complesso. Torneremo sulla questione delle equazioni di Lagrange di prima specie nella sezione 5.3.2.

Vediamo adesso come sia possibile percorrere un’altra strada che conduce direttamente ad un sistema di tante equazioni quanti sono i gradi di libertà. Prima di sviluppare tale procedura è opportuno analizzare più in dettaglio \mathbf{F}_i , $i = 1, \dots, N$, risultante di tutte le forze, non vincolari, che agiscono sull’ i -esimo punto materiale. Dividiamo quindi le forze che agiscono sull’ i -esimo punto materiale in due categorie:

1. Forze conservative (o di tipo gradiente), la cui risultante indicheremo con $\mathbf{F}_{i, cons}$. In tal caso sappiamo che esiste una funzione, detta **energia potenziale**,

$$V = V(\mathbf{x}_1, \dots, \mathbf{x}_N),$$

che, per brevità di notazione indicheremo anche con $V = V(\tilde{\mathbf{x}})$, per cui

$$\mathbf{F}_{i, cons} = -\frac{\partial V(\mathbf{x}_1, \dots, \mathbf{x}_N)}{\partial \mathbf{x}_i}, \quad (5.19)$$

dove $\frac{\partial}{\partial \mathbf{x}_i}$, rappresenta il gradiente rispetto alle coordinate (x_i, y_i, z_i) , come espresso dalla (4.7).

2. Tutte le altre forze che non sono né conservative né dovute all’interazione coi vincoli. La loro risultante viene denotata con \mathbf{S}_i .

Ritornando alle (5.18)₁ ed applicando la suddivisione appena illustrata

$$\mathbf{F}_i = \mathbf{F}_{i, cons} + \mathbf{S}_i = -\frac{\partial V}{\partial \mathbf{x}_i} + \mathbf{S}_i, \quad (5.20)$$

scriveremo

$$m_i \ddot{\mathbf{x}}_i = -\frac{\partial V(\tilde{\mathbf{x}})}{\partial \mathbf{x}_i} + \mathbf{S}_i + \sum_{k=1}^m \lambda_k \frac{\partial f_k}{\partial \mathbf{x}_i}, \quad i = 1, \dots, N. \quad (5.21)$$

Moltiplichiamo ora scalarmente, in \mathbb{R}^3 , ognuna delle equazioni per lo spostamento virtuale $\delta \mathbf{x}_i$ dell' i -esimo punto materiale e sommiamo le N equazioni scalari così ottenute. In virtù della assunzione di vincoli lisci, ovvero della (5.15), le reazioni vincolari "scompaiono" dalla somma, e abbiamo

$$\sum_{i=1}^N \left(m_i \ddot{\mathbf{x}}_i + \frac{\partial V(\tilde{\mathbf{x}})}{\partial \mathbf{x}_i} - \mathbf{S}_i \right) \cdot \delta \mathbf{x}_i = 0, \quad \forall (\delta \mathbf{x}_1, \dots, \delta \mathbf{x}_N), \quad (5.22)$$

che è nota con il nome di **equazione simbolica della dinamica**.

E' interessante scrivere la (5.22) nel formalismo compatto di \mathbb{R}^{3N} . Infatti introducendo

$$\tilde{\mathbf{S}} = \begin{pmatrix} \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_N \end{pmatrix},$$

e, ricordando la (5.19) e la (4.8),

$$\tilde{\mathbf{F}}_{cons} = \begin{pmatrix} \mathbf{F}_{1, cons} \\ \vdots \\ \mathbf{F}_{N, cons} \end{pmatrix} = \begin{pmatrix} -\frac{\partial V(\tilde{\mathbf{x}})}{\partial \mathbf{x}_1} \\ \vdots \\ -\frac{\partial V(\tilde{\mathbf{x}})}{\partial \mathbf{x}_N} \end{pmatrix} = -\tilde{\nabla} V(\tilde{\mathbf{x}}),$$

il sistema (5.21) di N equazioni vettoriali (e quindi $3N$ equazioni scalari) può esser scritto come

$$\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} + \tilde{\nabla} V(\tilde{\mathbf{x}}) - \tilde{\mathbf{S}} - \sum_{k=1}^m \lambda_k \tilde{\nabla} f_k = 0, \quad (5.23)$$

dove abbiamo sfruttato la (5.17) per esprimere le reazioni vincolari $\tilde{\mathbf{R}}$. Se adesso moltiplichiamo scalarmente in \mathbb{R}^{3N} la (5.23) per il generico spostamento virtuale $\delta \tilde{\mathbf{x}} \in \mathbb{R}^{3N}$, cioè

$$\left[\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} + \tilde{\nabla} V(\tilde{\mathbf{x}}) - \tilde{\mathbf{S}} \right] \bullet \delta \tilde{\mathbf{x}} - \underbrace{\sum_{k=1}^m \lambda_k \tilde{\nabla} f_k \bullet \delta \tilde{\mathbf{x}}}_{=0} = 0, \quad (5.24)$$

otteniamo, sviluppando il prodotto scalare \bullet di \mathbb{R}^{3N} in somma di N prodotti scalari \cdot di \mathbb{R}^3 , l'equazione simbolica della dinamica (5.22).

Nota 5.2.1 Osserviamo che la (5.24) si presta ad un'interessante lettura "energetica". Infatti, considerando $\delta\tilde{\mathbf{x}} = \sum_{i=1}^l \delta q_i \tilde{\mathbf{u}}_i$, con δq_i , $i = 1, \dots, l$, "infinitesimi", sappiamo che $\delta\tilde{\mathbf{x}}$ rappresenta di fatto uno "spostamento infinitesimo" del sistema degli N punti materiali compatibile con i vincoli. Ricordando la nota 5.1.1, il lavoro virtuale δL compiuto da **tutte le forze** (comprese le forze d'inerzia, ovvero $-m_i \ddot{\mathbf{x}}_i$, $i = 1, \dots, N$) durante lo spostamento virtuale $\delta\tilde{\mathbf{x}}$, è

$$\delta L = \left[\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} + \tilde{\nabla} V - \tilde{\mathbf{R}} - \tilde{\mathbf{S}} \right] \bullet \delta\tilde{\mathbf{x}}.$$

La (5.24) comporta dunque $\delta L = 0$. Ciò significa che il **lavoro virtuale compiuto da tutte le forze su un qualsiasi spostamento virtuale del sistema degli N punti materiali è nullo**, ovvero i lavori compiuti dalle singole forze si bilanciano perfettamente sì che δL si annulla. L'equazione $\delta L = 0$, prende il nome di **principio dei lavori virtuali**.

Osserviamo adesso che la (5.24) vale per ogni $\delta\tilde{\mathbf{x}}$. Avremo pertanto

$$\left[\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} + \tilde{\nabla} V(\tilde{\mathbf{x}}) - \tilde{\mathbf{S}} \right] \bullet \tilde{\mathbf{u}}_j = 0, \quad \forall j = 1, 2, \dots, l, \quad (5.25)$$

dal momento che, in virtù del teorema 4.1.2, o meglio della (4.25), ogni spostamento virtuale può essere espresso come combinazione degli l vettori indipendenti $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \dots, \tilde{\mathbf{u}}_l$. Quindi, in definitiva, **l'equazione simbolica della dinamica (5.24) si traduce nel sistema di l equazioni indipendenti (5.25), che sono le proiezioni di (5.23) lungo i vettori della base locale $\{\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_l\}$.**

Nella sezione 5.3 vedremo come esprimere il prodotto scalare (in \mathbb{R}^{3N})

$$\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} \bullet \tilde{\mathbf{u}}_j,$$

in termini dell'energia cinetica del sistema di punti materiali. In altri termini dimostreremo una generalizzazione del teorema 3.1.1 illustrato nel capitolo 3. Per far ciò utilizzeremo l'energia cinetica del sistema espressa tramite le coordinate lagrangiane.

Esempio 5.2.1 Mostriamo che il vincolo di rigidità che lega due punti è un vincolo liscio nel senso della definizione 5.2.1. Consideriamo due punti materiali P_1 e P_2 , le cui coordinate sono rispettivamente date da \mathbf{x}_1 e \mathbf{x}_2 , vincolati a mantenere costante la loro distanza d , cioè

$$f(\mathbf{x}_1, \mathbf{x}_2) = |\mathbf{x}_1 - \mathbf{x}_2|^2 - d^2 = 0,$$

che scriveremo esplicitamente come

$$f(\mathbf{x}_1, \mathbf{x}_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - d^2.$$

Indichiamo poi con \mathbf{d} , il vettore $(P_1 - P_2)$, cioè

$$\mathbf{d} = (x_1 - x_2) \mathbf{e}_x + (y_1 - y_2) \mathbf{e}_y + (z_1 - z_2) \mathbf{e}_z,$$

mentre il vettore $(P_2 - P_1)$ sarà $-\mathbf{d}$. Abbiamo poi

$$\tilde{\nabla} f = \begin{pmatrix} \frac{\partial f}{\partial \mathbf{x}_1} \\ \frac{\partial f}{\partial \mathbf{x}_2} \end{pmatrix} = \begin{pmatrix} 2\mathbf{d} \\ -2\mathbf{d} \end{pmatrix},$$

dal momento che $\frac{\partial f}{\partial \mathbf{x}_1} = 2(x_1 - x_2) \mathbf{e}_x + 2(y_1 - y_2) \mathbf{e}_y + 2(z_1 - z_2) \mathbf{e}_z$, e analogamente per $\frac{\partial f}{\partial \mathbf{x}_2}$. Denotiamo adesso con \mathbf{R}_1 la forza vincolare che agisce su P_1 , mentre quella che agisce su P_2 è \mathbf{R}_2 . Ora la forza \mathbf{R}_1 è dovuta al punto P_2 , mentre \mathbf{R}_2 , forza che agisce su P_2 , è originata da P_1 . Di conseguenza \mathbf{R}_1 e \mathbf{R}_2 sono forze interne, l'una l'opposta dell'altra per il principio di azione e reazione, e, sempre per la terza legge di Newton, sono parallele al vettore $(P_1 - P_2)$. Scriveremo dunque

$$\mathbf{R}_1 = \alpha \mathbf{d}, \quad e \quad \mathbf{R}_2 = -\mathbf{R}_1 = -\alpha \mathbf{d},$$

con $\alpha \in \mathbb{R}$, per cui

$$\tilde{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix} = \begin{pmatrix} \alpha \mathbf{d} \\ -\alpha \mathbf{d} \end{pmatrix}.$$

Consideriamo adesso un qualsiasi spostamento virtuale

$$\delta \tilde{\mathbf{x}} = \begin{pmatrix} \delta \mathbf{x}_1 \\ \delta \mathbf{x}_2 \end{pmatrix}.$$

In virtù della definizione 4.1.4 abbiamo

$$0 = \underbrace{\begin{pmatrix} \delta \mathbf{x}_1 \\ \delta \mathbf{x}_2 \end{pmatrix}}_{\delta \tilde{\mathbf{x}}} \cdot \underbrace{\begin{pmatrix} 2\mathbf{d} \\ -2\mathbf{d} \end{pmatrix}}_{\tilde{\nabla} f} = 2(\delta \mathbf{x}_1 - \delta \mathbf{x}_2) \cdot \mathbf{d}.$$

Ma allora

$$\tilde{\mathbf{R}} \cdot \delta \tilde{\mathbf{x}} = \begin{pmatrix} \alpha \mathbf{d} \\ -\alpha \mathbf{d} \end{pmatrix} \cdot \begin{pmatrix} \delta \mathbf{x}_1 \\ \delta \mathbf{x}_2 \end{pmatrix} = \alpha(\delta \mathbf{x}_1 - \delta \mathbf{x}_2) \cdot \mathbf{d} = 0,$$

cioè $\tilde{\mathbf{R}}$ soddisfa la (5.15): il vincolo è liscio. Tale risultato è estendibile anche a sistemi di $N > 2$ punti, a due a due vincolati dal vincolo di rigidità.

5.3 Le equazioni di Lagrange

Consideriamo il solito sistema di N punti materiali vincolato⁴, le cui configurazioni siano descritte tramite un insieme di coordinate lagrangiane (q_1, \dots, q_l) . Come abbiamo già visto, il legame tra le coordinate cartesiane dei punti e le coordinate lagrangiane sarà espresso dalle relazioni

$$\mathbf{x}_i = \mathbf{x}_i(q_1, \dots, q_l, t), \quad i = 1, 2, \dots, N, \quad (5.26)$$

nelle quali, come sappiamo, il tempo potrà comparire esplicitamente quando i vincoli sono mobili.

Teorema 5.3.1 Per ogni $j = 1, 2, \dots, l$ vale la seguente relazione

$$\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} \cdot \tilde{\mathbf{u}}_j = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j}, \quad (5.27)$$

essendo T l'energia cinetica del sistema di N punti materiali data dalla (4.33).

Dim. Partendo dalla (4.33) abbiamo

$$\frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial \dot{q}_j}.$$

Dalla (4.31) ricaviamo $\frac{\partial \dot{\mathbf{x}}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{x}_i}{\partial q_j}$, e quindi riscriviamo la precedente uguaglianza come

$$\frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j}. \quad (5.28)$$

Se deriviamo adesso la (5.28) rispetto al tempo si ottiene

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) = \sum_{i=1}^N m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} + \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{x}_i}{\partial q_j} \right).$$

Chiaramente il primo addendo è il termine di sinistra nella (5.27)

$$\sum_{i=1}^N m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} = \begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} \cdot \underbrace{\begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_j} \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_j} \end{pmatrix}}_{\tilde{\mathbf{u}}_j},$$

mentre, per quel che riguarda il secondo addendo, possiamo ricorrere al fatto che nell'espressione di \mathbf{x}_i , in funzione delle coordinate lagrangiane, le variabili q_j hanno un ruolo

⁴Le considerazioni ovviamente valgono anche per un sistema libero, ovvero non soggetto a vincoli

di variabili indipendenti rispetto al tempo. E' dunque possibile scambiare la derivata rispetto a tempo con la derivata parziale rispetto a q_j e ottenere⁵ $\frac{d}{dt} \left(\frac{\partial \mathbf{x}_i}{\partial q_j} \right) = \frac{\partial \dot{\mathbf{x}}_i}{\partial q_j}$.

Abbiamo quindi

$$\begin{aligned} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{x}_i}{\partial q_j} \right) &= \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial q_j} \\ &= \sum_{i=1}^N m_i \frac{\partial \dot{\mathbf{x}}_i^2}{\partial q_j} = \frac{\partial}{\partial q_j} \left(\sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2 \right) = \frac{\partial T}{\partial q_j}. \end{aligned} \quad (5.29)$$

Quindi, riunendo tutti i termini otteniamo

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) = \begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} \bullet \tilde{\mathbf{u}}_j + \frac{\partial T}{\partial q_j},$$

da cui la (5.27) discende banalmente.

□

La (5.27) viene usualmente detta **identità di Lagrange**, mentre il secondo membro è generalmente noto con il nome di **binomio di Lagrange**.

Torniamo adesso al sistema di l equazioni (5.25) ed applichiamo la (5.27). Otteniamo

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} + \tilde{\nabla} V(\tilde{\mathbf{x}}) \bullet \tilde{\mathbf{u}}_j - \tilde{\mathbf{S}} \bullet \tilde{\mathbf{u}}_j = 0, \quad j = 1, 2, \dots, l. \quad (5.30)$$

Ora, consideriamo l'energia potenziale come funzione delle variabili lagrangiane q_1, \dots, q_l , cioè

$$\hat{V}(\tilde{\mathbf{q}}, t) = V(\tilde{\mathbf{x}}(\tilde{\mathbf{q}}, t)) = V(\mathbf{x}_1(q_1, \dots, q_l, t), \dots, \mathbf{x}_N(q_1, \dots, q_l, t)),$$

e deriviamo rispetto a q_j . Troviamo

$$\frac{\partial \hat{V}}{\partial q_j} = \sum_{i=1}^N \frac{\partial V}{\partial \mathbf{x}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_j} = \underbrace{\begin{pmatrix} \frac{\partial V}{\partial \mathbf{x}_1} \\ \vdots \\ \frac{\partial V}{\partial \mathbf{x}_N} \end{pmatrix}}_{\tilde{\nabla} V(\tilde{\mathbf{x}})} \bullet \underbrace{\begin{pmatrix} \frac{\partial \mathbf{x}_1}{\partial q_j} \\ \vdots \\ \frac{\partial \mathbf{x}_N}{\partial q_j} \end{pmatrix}}_{\tilde{\mathbf{u}}_j}. \quad (5.31)$$

Introduciamo inoltre

$$\Xi_j = \tilde{\mathbf{S}} \bullet \tilde{\mathbf{u}}_j = \sum_{i=1}^N \mathbf{S}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_j}, \quad (5.32)$$

⁵Questo passaggio è stato ampiamente illustrato nella dimostrazione del teorema 3.1.1.

detta **componente Lagrangiana della forza \tilde{S}** .

Possiamo dunque riscrivere il sistema (5.30) nella forma nota sotto il nome di **equazioni di Lagrange di seconda specie** o, più semplicemente, **equazioni di Lagrange**

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial (T - \hat{V})}{\partial q_j} = \Xi_j, \quad j = 1, 2, \dots, l. \quad (5.33)$$

La funzione energia potenziale V non dipende da \dot{q}_i , $i = 1, \dots, N$. Si introduce quindi, come nella sezione 3.3, la **funzione di Lagrange** o **Lagrangiana** del sistema

$$\mathcal{L}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l, t) = T(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l, t) - \hat{V}(q_1, \dots, q_l, t), \quad (5.34)$$

e si riscrivono le equazioni di Lagrange nella forma compatta

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j} = \Xi_j, \quad j = 1, 2, \dots, l, \quad (5.35)$$

che, lo ricordiamo ancora, valgono nel caso in cui tutti i vincoli siano olonomi e lisci.

Se poi le forze cui è sottoposto il sistema sono soltanto quelle vincolari, del tipo (5.16), e conservative, cioè $\tilde{S} = 0$, oppure se $\tilde{S} \bullet \tilde{u}_j = 0$, $\forall j = 1, \dots, l$, allora il sistema (5.35) si semplifica in

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \quad j = 1, 2, \dots, l. \quad (5.36)$$

Nota 5.3.1 Rimarchiamo ancora che le equazioni di Lagrange (5.35) discendono dalla (5.25). Ma allora, riferendoci alla sezione 1.4, le l equazioni (5.35) possono essere lette come le **componenti covarianti** di $\mathbf{F} = m\mathbf{a}$, rispetto alla base locale.

Nota 5.3.2 Ricordando la scomposizione (4.33) dell'energia cinetica T come $T = T_2 + T_1 + T_0$, osserviamo che la componente $T_1 = \sum_{i=1}^l b_i \dot{q}_i$, è completamente ininfluente ai fini delle equazioni di moto se: (i) se tutti i $b_i = b_i(q_1, \dots, q_l)$, $i = 1, \dots, l$, non dipendono esplicitamente dal tempo; (ii) $\frac{\partial b_i}{\partial q_i} = \frac{\partial b_i}{\partial q_j}$, $\forall i, j = 1, \dots, l$. Nelle equazioni di moto il termine T_1 dà luogo all'espressione $\frac{d}{dt} \left(\frac{\partial T_1}{\partial \dot{q}_j} \right) - \frac{\partial T_1}{\partial q_j}$, che si annulla. Infatti

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial T_1}{\partial \dot{q}_j} \right) - \frac{\partial T_1}{\partial q_j} &= \frac{db_j(q_1(t), \dots, q_l(t))}{dt} - \sum_{i=1}^l \frac{\partial b_i}{\partial q_j} \dot{q}_i \\ &= \sum_{i=1}^l \frac{\partial b_j}{\partial q_i} \dot{q}_i - \sum_{i=1}^l \underbrace{\left(\frac{\partial b_i}{\partial q_j} \right)} \dot{q}_i = 0. \end{aligned}$$

5.3.1 Equazioni di Lagrange e statica del punto materiale sulla superficie

Continuiamo con l'esempio di un solo punto materiale vincolato a muoversi una superficie fissa, riscrivendo la (5.3), nella quale teniamo conto della scomposizione (5.20), cioè

$$m\ddot{\mathbf{x}} = \underbrace{\mathbf{F}_{cons}}_{-\nabla V} + \mathbf{R} + \mathbf{S}. \quad (5.37)$$

Se adesso proiettiamo la (5.37) nello spazio tangente, otteniamo due equazioni

$$m \ddot{\mathbf{x}} \cdot \mathbf{u}_i = -\nabla V \cdot \mathbf{u}_i + \mathbf{R} \cdot \mathbf{u}_i + \mathbf{S} \cdot \mathbf{u}_i, \quad i = 1, 2. \quad (5.38)$$

Se adesso la superficie è liscia, cioè vale la (5.8), sfruttando

$$m \ddot{\mathbf{x}} \cdot \mathbf{u}_i = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i}, \quad i = 1, 2,$$

e

$$\nabla V \cdot \mathbf{u}_i = \frac{\partial \hat{V}}{\partial q_i}, \quad i = 1, 2, \quad \text{con} \quad \hat{V}(q_1, q_2) = V(\mathbf{x}(q_1, q_2)),$$

abbiamo

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = \Xi_i, \quad i = 1, 2. \quad (5.39)$$

dove $\Xi_i = \mathbf{S} \cdot \mathbf{u}_i$, $i = 1, 2$, e dove

$$\mathcal{L} = \mathcal{L}(q_1, q_2, \dot{q}_1, \dot{q}_2) = T(q_1, q_2, \dot{q}_1, \dot{q}_2) - \hat{V}(q_1, q_2).$$

Le due equazioni (5.39) rappresentano quindi le equazioni di moto (5.35) per un punto materiale vincolato su una superficie regolare liscia.

Nota 5.3.3 *Notiamo che nelle due equazioni di moto (5.39) intervengono le componenti di \mathbf{R} e di \mathbf{S} rispetto alla base $\{\mathbf{u}_1, \mathbf{u}_2\}$ dello spazio tangente. L'equazione (5.37) può essere infatti scomposta nella parte ortogonale al piano tangente ed in quella che giace sul piano tangente*

$$\left[-m\ddot{\mathbf{x}}_{\parallel} + (\mathbf{F}_{cons})_{\parallel} + \mathbf{R}_{\parallel} + \mathbf{S}_{\parallel} \right] + \left[-m\ddot{\mathbf{x}}_{\perp} + (\mathbf{F}_{cons})_{\perp} + \mathbf{R}_{\perp} + \mathbf{S}_{\perp} \right] = 0,$$

dove, $\ddot{\mathbf{x}} = \ddot{\mathbf{x}}_{\parallel} + \ddot{\mathbf{x}}_{\perp}$, con $\ddot{\mathbf{x}}_{\parallel} \in \Pi(P)$ e $\ddot{\mathbf{x}}_{\perp} \in \mathcal{N}(P)$. Perciò moltiplicando scalarmente la (5.37) per \mathbf{u}_i , $i = 1, 2$, si isola la parte $[\dots]_{\parallel}$ e si ottengono, sotto l'ipotesi (5.8), le equazioni di moto (5.39). Questa è una notevole semplificazione dal momento che si "elimina" tutta la parte $[\dots]_{\perp}$. Rimane quindi da analizzare la parte $[\dots]_{\perp}$. Per far ciò, assumendo $\mathbf{R}_{\perp} = \lambda \nabla f$, è sufficiente moltiplicare scalarmente la (5.37) per ∇f

$$(-m\ddot{\mathbf{x}} + \mathbf{F}_{cons} + \lambda \nabla f + \mathbf{S}) \cdot \nabla f = 0, \quad (5.40)$$

da cui discende la (5.10). La (5.40) è, come illustrato nella sezione 5.1, il punto di partenza per la determinazione di λ , ossia di \mathbf{R} .

Le equazioni (5.38) possono essere anche raggruppate in un'unica equazione

$$\sum_{i=1}^2 (-m \ddot{\mathbf{x}} - \nabla V + \mathbf{R} + \mathbf{S}) \cdot \mathbf{u}_i = 0,$$

da cui discende l'equazione simbolica della dinamica

$$\left(\begin{array}{cccc} -m \ddot{\mathbf{x}} & - & \nabla V & + \\ \text{forza inerzia} & & \text{forza conservativa} & \\ & & \mathbf{R} & + \\ & & \text{reazione vincolare} & \\ & & \mathbf{S} & \\ & & \text{altre forze} & \end{array} \right) \cdot \delta \mathbf{x} = 0, \quad (5.41)$$

nel momento in cui si considera il generico spostamento virtuale (4.51).

Analizziamo adesso la **statica** del punto materiale. Se, $\forall t > 0$, $\ddot{\mathbf{x}} \equiv \dot{\mathbf{x}} \equiv 0$, la (5.37) diventa

$$\underbrace{\mathbf{F}_{cons}}_{-\nabla V} + \mathbf{R} + \mathbf{S} = 0, \quad (5.42)$$

e le (5.38) si riducono a

$$\begin{cases} (-\nabla V + \mathbf{R} + \mathbf{S}) \cdot \mathbf{u}_1 = 0, \\ (-\nabla V + \mathbf{R} + \mathbf{S}) \cdot \mathbf{u}_2 = 0, \end{cases} \quad (5.43)$$

che, assumendo al solito la superficie priva di attrito, si traducono in

$$\begin{cases} -\frac{\partial \hat{V}}{\partial q_1} + \mathbf{S} \cdot \mathbf{u}_1 = 0, \\ -\frac{\partial \hat{V}}{\partial q_2} + \mathbf{S} \cdot \mathbf{u}_2 = 0. \end{cases} \quad (5.44)$$

Quindi, data \mathbf{S} , le soluzioni (\bar{q}_1, \bar{q}_2) , ammesso che esistano⁶, del sistema (5.44) rappresentano le configurazioni di equilibrio del punto materiale sulla superficie. La (5.44) può anche essere utilizzata come formula per determinare la forza \mathbf{S} che deve essere applicata al punto materiale affinché esso sia in equilibrio in (\bar{q}_1, \bar{q}_2) . In tal caso le incognite sono le due componenti di \mathbf{S} nel piano tangente, cioè

$$\mathbf{S} = S_1 \boldsymbol{\kappa}_1 + S_2 \boldsymbol{\kappa}_2, \quad \text{dove, al solito, } \boldsymbol{\kappa}_1 = \frac{\mathbf{u}_1}{|\mathbf{u}_1|}, \quad \boldsymbol{\kappa}_2 = \frac{\mathbf{u}_2}{|\mathbf{u}_2|}.$$

Siccome $\boldsymbol{\kappa}_i \cdot \mathbf{u}_i = \frac{\mathbf{u}_i \cdot \mathbf{u}_i}{|\mathbf{u}_i|} = |\mathbf{u}_i|$, $i = 1, 2$, la (5.44) equivale al seguente sistema

$$\begin{cases} -\frac{\partial \hat{V}}{\partial q_1} + S_1 |\mathbf{u}_1| = 0, \\ -\frac{\partial \hat{V}}{\partial q_2} + S_2 |\mathbf{u}_2| = 0. \end{cases} \quad (5.45)$$

Si noti che in presenza di sole forze conservative, ovvero nel caso in cui $\mathbf{S} = 0$, le posizioni di equilibrio sono date dalle soluzioni di

$$\frac{\partial \hat{V}}{\partial q_1} = 0, \quad \text{e} \quad \frac{\partial \hat{V}}{\partial q_2} = 0,$$

che quindi corrispondono al punto di estremo locale dell'energia potenziale. Le configurazioni stabili corrisponderanno ai minimi isolati dell'energia potenziale (criterio di Dirichlet), come proveremo nella sezione 5.9.1.

Sempre assumendo la superficie liscia, dalla (5.42), si ricava

$$\underbrace{\mathbf{R}}_{\lambda \nabla f} = -\mathbf{S} - \mathbf{F}_{cons}. \quad (5.46)$$

⁶In generale, il sistema (5.44) può avere più di una soluzione, ma anche non averne.

Moltiplicando scalarmente la (5.46) per ∇f otteniamo

$$\lambda = -\frac{(\mathbf{F}_{cons} + \mathbf{S}) \cdot \nabla f}{|\nabla f|^2},$$

che, nel caso in cui $\mathbf{S} = 0$, si semplifica in $\lambda = -\frac{\mathbf{F}_{cons} \cdot \nabla f}{|\nabla f|^2}$.

Esempio 5.3.1 Consideriamo il moto di un punto materiale di massa m su una sfera liscia di raggio R , soggetto alla sola forza peso.

La sfera è definita implicitamente da $f(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0$. Escludendo i poli⁷, può essere così parametrizzata

$$\begin{cases} x = R \sin \theta \cos \phi, \\ y = R \sin \theta \sin \phi, \\ z = R \cos \theta, \end{cases}$$

dove i parametri sono la colatitudine $\theta \in (0, \pi)$, e la longitudine $\phi \in (0, 2\pi]$, come mostrato nella figura 5.1. I vettori tangenti sono

$$\begin{aligned} \mathbf{u}_\theta &= R \cos \theta \cos \phi \mathbf{e}_x + R \cos \theta \sin \phi \mathbf{e}_y - R \sin \theta \mathbf{e}_z, \\ \mathbf{u}_\phi &= -R \sin \theta \sin \phi \mathbf{e}_x + R \sin \theta \cos \phi \mathbf{e}_y. \end{aligned}$$

Si verifica facilmente che $|\mathbf{u}_\theta \wedge \mathbf{u}_\phi| = R^4 \sin^2 \theta$. Quindi la parametrizzazione è regolare ovunque eccetto ai poli. Calcolando la matrice \mathbb{A} data dalla (4.52) si ha

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{pmatrix}$$

e pertanto, sfruttando la (4.56), l'energia cinetica del punto materiale è

$$T = \frac{m}{2} \begin{pmatrix} \dot{\theta} & \dot{\phi} \end{pmatrix} \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{pmatrix} \begin{pmatrix} \dot{\theta} \\ \dot{\phi} \end{pmatrix} = \frac{mR^2}{2} (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta).$$

La forza peso è $\mathbf{F}_{peso} = -mg\mathbf{e}_z$, la cui energia potenziale è $V(z) = mgz$, e quindi $\hat{V}(\theta, \phi) = mgR \cos \theta$. La funzione di Lagrange è

$$\mathcal{L} = T - \hat{V} = \frac{m}{2} R^2 (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) - mgR \cos \theta$$

e le equazioni del moto sono

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) - \frac{\partial \mathcal{L}}{\partial \theta} = 0, \Rightarrow \ddot{\theta} - \left(\dot{\phi}^2 \sin \theta \cos \theta + \frac{g}{R} \sin \theta \right) = 0, \quad (5.47)$$

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \Rightarrow \frac{d}{dt} (mR^2 \dot{\phi} \sin^2 \theta) = 0. \quad (5.48)$$

⁷Una rappresentazione regolare della sfera necessita di due parametrizzazioni come quelle fornite dalla cosiddetta proiezione stereografica.

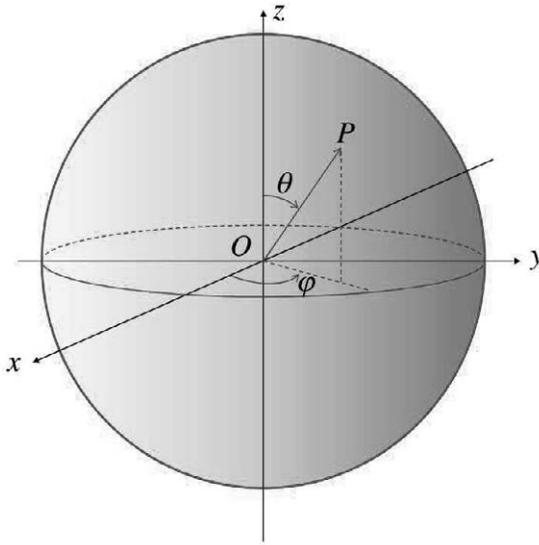


Figura 5.1: Sfera di raggio R .

Se vogliamo, la (5.48) è ovvia dal momento che la ϕ è coordinata ciclica. Abbiamo quindi

$$\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = mR^2 \dot{\phi} \sin^2 \theta = A_o, \quad (5.49)$$

dove A_o è una costante indipendente dal tempo. Dalla (5.49) ricaviamo subito che se $\dot{\phi} = 0$, allora $A_o = 0$, e la (5.47) si riduce all'equazione del pendolo di lunghezza R , $\ddot{\theta} - \frac{g}{R} \sin \theta = 0$.

Utilizzando la (5.49) per esprimere $\dot{\phi} = \frac{A_o}{mR^2 \sin^2 \theta}$, e sostituirlo nella (5.47) otteniamo

$$\ddot{\theta} = \left(\frac{A_o}{mR^2} \right)^2 \frac{\cos \theta}{\sin^3 \theta} + \frac{g}{R} \sin \theta,$$

che possiamo anche scrivere come

$$\ddot{\theta} = -\frac{dV_{\text{eff}}(\theta)}{d\theta}, \quad \text{con } V_{\text{eff}}(\theta) = \frac{1}{2} \left(\frac{A_o}{mR^2} \right)^2 \frac{1}{\sin^2 \theta} + \frac{g}{R} \cos \theta, \quad (5.50)$$

energia potenziale efficace. Il grafico di $V_{\text{eff}}(\theta)$ per $\theta \in (0, \pi)$ è riportato in figura 5.2, e mostra due asintoti verticali ed un minimo in corrispondenza di $\bar{\theta}$, soluzione di $\left(\frac{A_o}{mR^2} \right)^2 \frac{\cos \bar{\theta}}{\sin^3 \bar{\theta}} + \frac{g}{R} \sin \bar{\theta} = 0$. Se $\theta(t) = \bar{\theta}$, la particella percorre un'orbita circolare che risulta stabile (in quanto $\bar{\theta}$ è minimo di V_{eff}) In particolare, moltiplicando

la (5.50) per $\dot{\theta}$, è facile ottenere

$$\frac{d}{dt} \left(\frac{\dot{\theta}^2}{2} + V_{\text{eff}}(\theta) \right) = 0, \quad \Rightarrow \quad \dot{\theta} = \pm \sqrt{2(E - V_{\text{eff}}(\theta))}, \quad (5.51)$$

con E costante corrispondente all'energia meccanica.

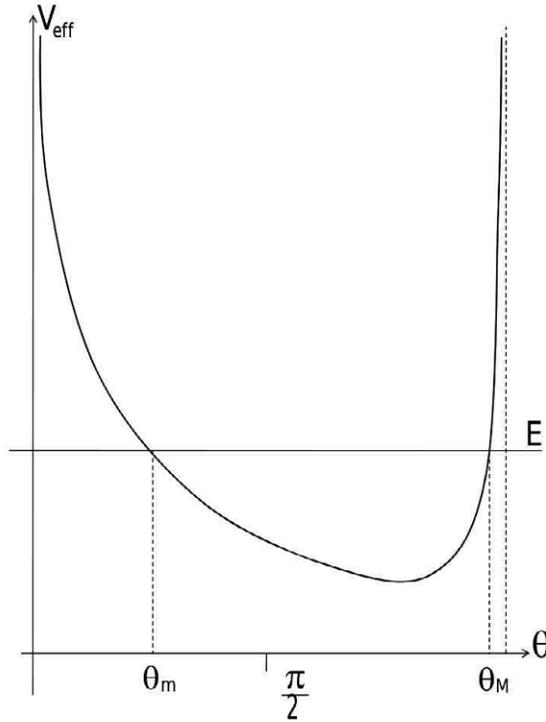


Figura 5.2: Energia potenziale $V_{\text{eff}}(\theta)$.

Dall'analisi qualitativa si deduce che $\theta(t)$ è periodica e, fissato il livello energetico E , il moto si svolge fra θ_m e θ_M . Il periodo di oscillazione è

$$T = 2 \int_{\theta_m}^{\theta_M} \frac{du}{\sqrt{2(E - V_{\text{eff}}(u))}}.$$

Il moto complessivo non è necessariamente periodico. Ponendo $\phi = \phi(\theta)$, si ha

$$\dot{\phi} = \frac{d\phi}{d\theta} \dot{\theta}, \text{ da cui}$$

$$\frac{d\phi}{d\theta} = \frac{\dot{\phi}}{\dot{\theta}} \stackrel{(5.51)}{=} \frac{A_o}{mR^2 \sin^2 \theta} \frac{1}{\sqrt{2(E - V_{\text{eff}}(\theta))}}.$$

Il moto sarà effettivamente periodico se, durante l'oscillazione compiuta fra θ_m e θ_M (andata e ritorno) la coordinata angolare ϕ ha subito una variazione pari ad un frazione di π , ovvero

$$\begin{aligned} \underbrace{2(\phi(\theta_M) - \phi(\theta_m))}_{\Delta\phi} &= 2 \int_{\theta_m}^{\theta_M} \frac{d\phi}{d\theta} d\theta \\ &= \frac{2A_o}{mR^2} \int_{\theta_m}^{\theta_M} \frac{1}{\sin^2 \theta} \frac{d\theta}{\sqrt{2(E - V_{\text{eff}}(\theta))}} = \frac{r}{s} 2\pi, \end{aligned}$$

con $r, s \in \mathbb{N}$. Se è così allora, dopo s oscillazioni la traiettoria si chiude dal momento che l'angolo $\Delta\phi$ "spazzato" è $2\pi r$.

Lasciamo infine come esercizio la dimostrazione che A_o è, a parte il segno, la componente lungo l'asse z del momento angolare della particella calcolato rispetto ad O .

Esempio 5.3.2 Consideriamo il moto di un punto materiale su un cono liscio soggetto alla sola forza peso. L'equazione implicita del cono è $x^2 + y^2 = \frac{z^2}{c^2}$. Il cono quindi è regolare ovunque, eccetto che nell'origine. Possiamo esprimere la superficie nella seguente forma parametrica (si considera solo $z > 0$)

$$\begin{cases} x = r \cos \phi, \\ y = r \sin \phi, \\ z = cr, \end{cases} \quad \phi \in (0, 2\pi], \quad r \in (0, +\infty).$$

I vettori che costituiscono la base del piano tangente sono

$$\begin{aligned} \mathbf{u}_r &= \cos \phi \mathbf{e}_x + \sin \phi \mathbf{e}_y + c\mathbf{e}_z, \\ \mathbf{u}_\phi &= -r \sin \phi \mathbf{e}_x + r \cos \phi \mathbf{e}_y. \end{aligned}$$

L'energia cinetica del punto è

$$T = \frac{m}{2} \begin{pmatrix} \dot{r} & \dot{\phi} \end{pmatrix} \begin{pmatrix} 1+c^2 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \dot{r} \\ \dot{\phi} \end{pmatrix} = \frac{m}{2} \left[(1+c^2) \dot{r}^2 + r^2 \dot{\phi}^2 \right].$$

L'unica forza "attiva" - cioè non dovuta al vincolo - che agisce sulla particella è il peso la cui energia potenziale è $V(x, y, z) = mgz$, per cui $\hat{V}(r, \phi) = mgr$. La funzione Lagrangiana è dunque

$$\mathcal{L} = T - \hat{V} = \frac{m}{2} \left[(1+c^2) \dot{r}^2 + r^2 \dot{\phi}^2 \right] - mgr.$$

Scriviamo adesso le equazioni del moto tenendo conto del fatto che ϕ è coordinata ciclica, cioè $\frac{\partial \mathcal{L}}{\partial \phi} = 0$,

$$\underbrace{\frac{\partial \mathcal{L}}{\partial \dot{\phi}}}_{mr^2 \dot{\phi}} = A_o = \text{costante},$$

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{r}} \right) - \frac{\partial \mathcal{L}}{\partial r} = 0, \quad \Rightarrow \quad (1 + c^2) \ddot{r} - r \dot{\phi}^2 + gc = 0,$$

che possiamo anche scrivere così

$$\ddot{r} = \frac{2a}{r^3} - b = -\frac{dV_{\text{eff}}(r)}{dr}$$

dove

$$V_{\text{eff}}(r) = \frac{a}{r^2} + br.$$

con $2a = \left(\frac{A_o}{m} \right)^2 \frac{1}{1 + c^2}$, $b = \frac{gc}{1 + c^2}$. Il grafico di $V_{\text{eff}}(r)$, riportato in figura 5.3, evidenzia la presenza di un minimo che corrisponde ad un'orbita circolare stabile. In generale il moto è periodico e si può trovare un'espressione del periodo T corrispondente al livello energetico E .

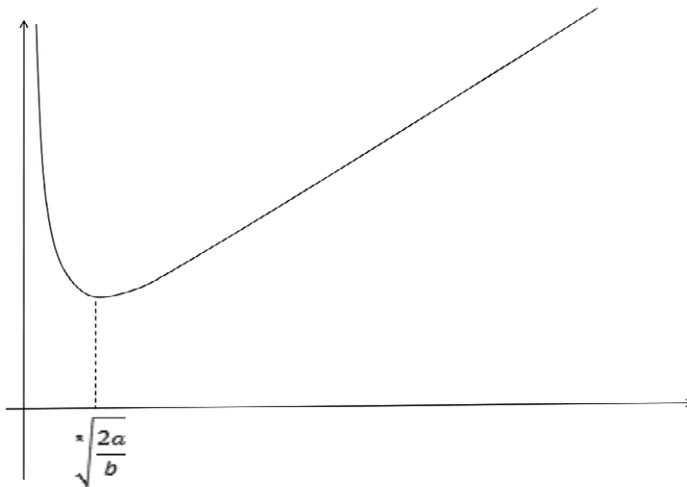


Figura 5.3: Energia potenziale $V_{\text{eff}}(r)$ nel caso di una particella, soggetta soltanto alla forza peso, vincolata su un cono.

Esempio 5.3.3 *E' dato un punto materiale, soggetto alla sola forza peso, posto su un cilindro liscio il cui asse è orizzontale. Si vuol determinare:*

1. *le equazioni di moto;*
2. *la reazione vincolare in condizioni dinamiche;*
3. *la forza \mathbf{S} che deve essere applicata al punto materiale affinché la configurazione (x_o, θ_o) sia posizione di equilibrio. Si vuole calcolare anche la corrispondente reazione vincolare.*

1. *La forma implicita dell'equazione del cilindro è $f(x, y, z) = y^2 + z^2 - R^2 = 0$, dove R è il raggio. Una parametrizzazione della superficie è dunque la seguente*

$$\begin{cases} x \in \mathbb{R}, \\ y = R \cos \theta, \\ z = R \sin \theta, \end{cases} \quad \text{con } -\pi < \theta \leq \pi.$$

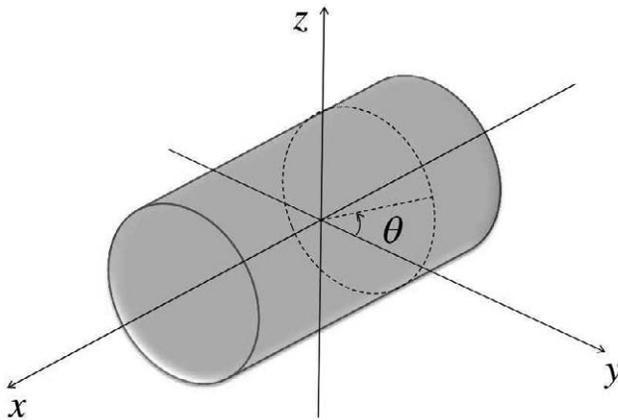


Figura 5.4: Cilindro orizzontale di raggio R .

Abbiamo poi

$$\mathbf{u}_x = \mathbf{e}_1, \quad \mathbf{u}_\theta = -R \sin \theta \mathbf{e}_2 + R \cos \theta \mathbf{e}_3. \quad (5.52)$$

L'energia cinetica della particella è

$$T = \frac{m}{2} (\dot{x}^2 + R^2 \dot{\theta}^2),$$

mentre l'energia potenziale relativa alla forza peso è $\hat{V} = mgR \sin \theta$. La Lagrangiana è

$$\mathcal{L} = \frac{m}{2} (\dot{x}^2 + R^2 \dot{\theta}^2) - mgR \sin \theta,$$

da cui si nota immediatamente che la x è ciclica, cioè

$$\frac{\partial \mathcal{L}}{\partial x} = 0, \Rightarrow \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = 0, \Rightarrow \dot{x} = \text{cost.}$$

Per quanto riguarda l'equazione per θ , abbiamo

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) - \frac{\partial \mathcal{L}}{\partial \theta} = 0, \Rightarrow \ddot{\theta} = -\frac{g}{R} \cos \theta,$$

la cui interpretazione fisica è immediata: la particella trasla con velocità longitudinale costante e contemporaneamente oscilla come un pendolo di lunghezza R . In particolare, se $\dot{x} = 0$, la posizione $\theta = -\frac{\pi}{2}$, che corrisponde al punto materiale nella parte bassa del cilindro, è posizione di equilibrio stabile.

2. Si parte dalla formula (5.13). Siccome

$$\nabla f = 2ye_y + 2ze_z, \quad |\nabla f|^2 = 4(y^2 + z^2) = 4R^2,$$

abbiamo $\nabla f \cdot \mathbf{F} = (2ye_y + 2ze_z) \cdot (-mge_z) = -2zmg = -2mgR \sin \theta$, mentre la matrice (5.12) è data da

$$\mathbb{H}(f) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Dobbiamo quindi esprimere $\dot{\mathbf{x}}$ rispetto alla base ortonormale $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$,

$$\begin{aligned} \dot{\mathbf{x}} &= \dot{x}\mathbf{u}_x + \dot{\theta}\mathbf{u}_\theta \stackrel{(5.52)}{=} \\ &= \dot{x}\mathbf{e}_x + \dot{\theta}(-R \sin \theta \mathbf{e}_y + R \cos \theta \mathbf{e}_z). \end{aligned}$$

Quindi

$$m\dot{\mathbf{x}} \cdot \mathbb{H}(f) \dot{\mathbf{x}} = 2mR^2\dot{\theta}^2,$$

da cui $\lambda = -\frac{mR\dot{\theta}^2 - mg \sin \theta}{2R}$, e

$$\mathbf{R} = -\lambda \nabla f = -\left(mR\dot{\theta}^2 - mg \sin \theta\right) (\cos \theta \mathbf{e}_y + \sin \theta \mathbf{e}_z). \quad (5.53)$$

3. L'incognita adesso è \mathbf{S} , o meglio, le componenti di \mathbf{S} sul piano tangente

$$\mathbf{S} = S_x \boldsymbol{\kappa}_x + S_\theta \boldsymbol{\kappa}_\theta, \quad \text{con } \boldsymbol{\kappa}_x = \frac{\mathbf{u}_x}{|\mathbf{u}_x|} = \mathbf{e}_x, \quad \boldsymbol{\kappa}_\theta = \frac{\mathbf{u}_\theta}{|\mathbf{u}_\theta|} = \frac{\mathbf{u}_\theta}{R}.$$

Dalla (5.45) abbiamo

$$\begin{cases} -\frac{\partial \hat{V}}{\partial x} + S_x = 0 \\ -\frac{\partial \hat{V}}{\partial \theta} + S_\theta R = 0 \end{cases} \Rightarrow \begin{cases} S_x = 0 \\ S_\theta = mg \cos \theta_o \end{cases}$$

Volendo esprimere \mathbf{S} rispetto alla base cartesiana $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ si ha

$$\mathbf{S} = mg \cos \theta_o \frac{\mathbf{u}_o}{R} \stackrel{(5.52)}{=} mg \cos \theta_o (-\sin \theta_o \mathbf{e}_y + \cos \theta_o \mathbf{e}_z).$$

Notiamo che \mathbf{S} si annulla se $\theta_o = \pm \frac{\pi}{2}$. Per quanto riguarda la reazione vincolare \mathbf{R} , sfruttando $\lambda = -\frac{\mathbf{F}_{cons} \cdot \nabla f}{|\nabla f|^2}$, abbiamo

$$\lambda = \frac{mg \sin \theta}{2R}, \Rightarrow \mathbf{R} = mg \sin \theta (\cos \theta \mathbf{e}_y + \sin \theta \mathbf{e}_z)$$

che coincide con la (5.53) in caso $\dot{\theta} = 0$.

5.3.2 Equazioni di Lagrange di prima specie

Riprendiamo quanto visto nell'osservazione 4.1.3 della sezione 4.1, supponendo che il sistema meccanico con l gradi di libertà sia soggetto all'ulteriore vincolo olonomo (4.11), che per brevità indicheremo anche con $\Phi(\tilde{\mathbf{q}}) = 0$. Evidentemente, come già osservato, si può procedere come nella sezione 4.1 e sfruttare l'equazione vincolare per ridurre i gradi di libertà passando così da l ad $(l - 1)$. Tuttavia si può evitare tale strada (che, in taluni casi, può essere estremamente complicata), tenendo conto del vincolo in un altro modo.

Assumendo che il nuovo vincolo sia liscio, consideriamo le forze vincolari agenti sui punti materiali. Queste, riprendendo la notazione della (5.14), verranno così indicate

$$\tilde{\mathbf{R}}_T = \tilde{\mathbf{R}} + \tilde{\mathbf{R}}_{nv},$$

dove $\tilde{\mathbf{R}} \in \mathbb{R}^{3N}$ è dato dalla (5.14) e rappresenta le forze (o meglio le reazioni) vincolari dovute agli m vincoli olonomi (4.4), mentre

$$\tilde{\mathbf{R}}_{nv} = \begin{pmatrix} \mathbf{R}_{1, nv} \\ \vdots \\ \mathbf{R}_{N, nv} \end{pmatrix},$$

rappresenta le forze vincolari dovute al nuovo vincolo $\Phi(\tilde{\mathbf{q}}) = 0$. Siccome tutti i vincoli sono lisci, applicando la definizione 5.2.1, deduciamo che $\tilde{\mathbf{R}}_T$ deve appartenere al sottospazio di \mathbb{R}^{3N} la cui base è $\{\tilde{\nabla} f_1, \tilde{\nabla} f_2, \dots, \tilde{\nabla} f_m, \tilde{\nabla} f_{m+1}\}$. Pertanto, ricordando la (5.16), scriveremo

$$\tilde{\mathbf{R}}_T = \sum_{k=1}^m \lambda_k \tilde{\nabla} f_k + \lambda \tilde{\nabla} f_{m+1}.$$

La (5.23) diventa quindi

$$\begin{pmatrix} m_1 \ddot{\mathbf{x}}_1 \\ \vdots \\ m_N \ddot{\mathbf{x}}_N \end{pmatrix} + \tilde{\nabla} V(\tilde{\mathbf{x}}) - \tilde{\mathbf{S}} - \underbrace{\left[\sum_{k=1}^m \lambda_k \tilde{\nabla} f_k + \lambda \tilde{\nabla} f_{m+1} \right]}_{\tilde{\mathbf{R}}_T} = 0. \quad (5.54)$$

Se adesso moltiplichiamo scalarmente la (5.54) per $\tilde{\mathbf{u}}_k$, $k = 1, \dots, l$, e applichiamo la (5.27), la (5.31), la (5.32) e la (4.23) otteniamo l equazioni del tipo

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} + \frac{\partial \hat{V}}{\partial q_k} - \Xi_k - \lambda \tilde{\nabla} f_{m+1} \bullet \tilde{\mathbf{u}}_k = 0. \quad (5.55)$$

In particolare, ricordando sia la definizione (5.34) sia l'espressione (4.11) del nuovo vincolo, che consente di scrivere

$$\frac{\partial \Phi}{\partial q_k} = \sum_{i=1}^N \frac{\partial f_{m+1}}{\partial \mathbf{x}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} = \tilde{\nabla} f_{m+1} \bullet \tilde{\mathbf{u}}_k,$$

abbiamo che il sistema (5.55) può essere posto nella seguente forma

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} = \Xi_k + \lambda \frac{\partial \Phi}{\partial q_k}, \quad k = 1, \dots, l. \quad (5.56)$$

anche dette **equazioni di Lagrange di prima specie**. La (5.56) è un sistema di l equazioni differenziali a cui va aggiunta l'equazione del vincolo $\Phi(\tilde{\mathbf{q}}) = 0$, nelle $(l + 1)$ incognite $q_1(t), \dots, q_l(t)$ e $\lambda(t)$.

Osserviamo che anziché moltiplicare scalarmente la (5.54) per $\tilde{\mathbf{u}}_k$, avremmo potuto moltiplicarla per la generica combinazione lineare $\sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s$, e poi selezionare, di volta in volta, i coefficienti δq_s in modo da ridurre $\sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s$ ai singoli vettori $\tilde{\mathbf{u}}_k$, $k = 1, \dots, l$. La peculiarità della combinazione lineare $\sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s$ è, come rimarcato nella nota 4.1.5, che questa **non rappresenta uno spostamento virtuale del sistema soggetto agli $(m + 1)$ vincoli**. Pertanto $\left(\sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s \right) \bullet \tilde{\mathbf{R}}_T$, non si annulla, ma genera il contributo

$$\left(\sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s \right) \bullet \tilde{\mathbf{R}}_{nv} = \lambda \sum_{s=1}^l \delta q_s \tilde{\mathbf{u}}_s \bullet \tilde{\nabla} f_{m+1} = \lambda \sum_{s=1}^l \delta q_s \frac{\partial \Phi}{\partial q_s}.$$

La procedura illustrata si estende facilmente anche al caso in cui si abbiano $n > 1$ vincoli olonomi lisci supplementari del tipo (4.11), ed anche al caso in questi dipendano dal tempo, cioè

$$\Phi_j(q_1, \dots, q_l, t) = 0, \quad j = 1, \dots, n. \quad (5.57)$$

In tal caso il sistema di equazioni di Lagrange di prima specie diventa

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} = \Xi_k + \sum_{j=1}^n \lambda_j \frac{\partial \Phi_j}{\partial q_k}, \quad k = 1, \dots, l, \quad (5.58)$$

ovvero un sistema di l equazioni differenziali accoppiato con le n equazioni vincolari (5.57). Abbiamo così un sistema di $(l + n)$ equazioni nelle $(l + n)$ incognite $q_1(t), \dots, q_l(t)$ e $\lambda_1(t), \dots, \lambda_n(t)$.

Rimandiamo alla sezione 8.4.1 per l'illustrazione di un esempio di applicazione delle equazioni di Lagrange di prima specie.

Si conclude con un'osservazione sul significato dei termini $\sum_{j=1}^n \lambda_j \frac{\partial \Phi_j}{\partial q_k}$, che appaiono nella (5.58). In tal senso proprio la forma della (5.58) è illuminante: i termini $\sum_{j=1}^n \lambda_j \frac{\partial \Phi_j}{\partial q_k}$ si assimilano, seguendo la definizione (5.32), alle componenti Lagrangiane della forza $\tilde{\mathbf{R}}_{nv}$ dovuta all'ulteriore vincolo.

5.4 Risolubilità delle equazioni di Lagrange

Abbiamo annunciato che le equazioni di Lagrange, che sono equivalenti all'equazione simbolica della dinamica, sono sufficienti per determinare la dinamica di un sistema meccanico⁸. Questo significa che possiamo determinare, almeno per un tempo "piccolo" in accordo con la teoria delle equazioni differenziali, le quantità $q_j(t)$, $j = 1, \dots, l$, una volta assegnate delle condizioni iniziali $q_j(t_0)$ e $\dot{q}_j(t_0)$, $j = 1, \dots, l$.

Osserviamo per prima cosa che le (5.35) sono un sistema di equazioni differenziali ordinarie del secondo ordine, in quanto le derivate di ordine massimo che compaiono quando si sviluppa la derivata totale rispetto al tempo nel termine $\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right)$ sono derivate seconde. Per la forma di T (vedi la (4.33)), abbiamo⁹

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) \\ &= \frac{d}{dt} \left[\sum_{h=1}^l a_{hk}(q_1(t), \dots, q_l(t), t) \dot{q}_k + b_k(q_1(t), \dots, q_l(t), t) \right]. \end{aligned}$$

Eseguendo quindi la derivazione rispetto al tempo compaiono le derivate seconde

$$\sum_{h=1}^l a_{hk}(q_1(t), \dots, q_l(t), t) \ddot{q}_h,$$

e il sistema di equazioni di Lagrange ha dunque la forma

$$\sum_{h=1}^l a_{hk} \ddot{q}_h = G_k(q_1(t), \dots, q_l(t), \dot{q}_1(t), \dots, \dot{q}_l(t), t), \quad k = 1, \dots, l, \quad (5.59)$$

che, in forma matriciale compatta, si scriverà

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1l} \\ \vdots & & \vdots \\ a_{l1} & \cdots & a_{ll} \end{pmatrix}}_{\mathbb{A}} \underbrace{\begin{pmatrix} \ddot{q}_1 \\ \vdots \\ \ddot{q}_l \end{pmatrix}}_{\ddot{\mathbf{q}}} = \underbrace{\begin{pmatrix} G_1(\tilde{\mathbf{q}}(t), \tilde{\dot{\mathbf{q}}}(t), t) \\ \vdots \\ G_l(\tilde{\mathbf{q}}(t), \tilde{\dot{\mathbf{q}}}(t), t) \end{pmatrix}}_{\tilde{\mathbf{G}}(\tilde{\mathbf{q}}(t), \tilde{\dot{\mathbf{q}}}(t), t)}.$$

Ricordando il teorema 4.1.3, che la matrice \mathbb{A} è definita positiva e quindi invertibile. Possiamo pertanto risolvere algebricamente le (5.59) rispetto alle derivate seconde \ddot{q}_h , $h = 1, \dots, l$, ottenendo un sistema di equazioni differenziali ordinarie in forma normale corredate con dati iniziali

$$\begin{cases} \ddot{\mathbf{q}} = \mathbb{A}^{-1} \tilde{\mathbf{G}}(\tilde{\mathbf{q}}(t), \tilde{\dot{\mathbf{q}}}(t), t), \\ \tilde{\mathbf{q}}(t_0) = \tilde{\mathbf{q}}_0, \\ \tilde{\dot{\mathbf{q}}}(t_0) = \tilde{\dot{\mathbf{q}}}_0. \end{cases} \quad (5.60)$$

⁸Purché soggetto a vincoli lisci.

⁹Si consiglia di non procedere oltre se non si ha chiara questa prima derivazione rispetto a \dot{q}_k . Si noti in particolare che il fattore $\frac{1}{2}$ scompare per la simmetria della matrice a_{hk} .

Il problema (5.60) è un tipico problema di Cauchy al quale possiamo applicare la teoria sviluppata nella sezione 2.3, e così affermare che, se le funzioni \tilde{G} godono delle opportune regolarità, per ogni insieme di dati iniziali $q_1(t_0), \dots, q_l(t_0), \dot{q}_1(t_0), \dots, \dot{q}_l(t_0)$, esiste un'unica soluzione (locale) del sistema.

5.5 Invarianza delle equazioni di Lagrange

Supponiamo che (q_1, \dots, q_l) siano le l coordinate lagrangiane utilizzate per descrivere il sistema di N punti materiali soggetto a m vincoli lisci. Sia $\mathcal{L}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l, t)$ la relativa funzione di Lagrange. Supponiamo adesso di individuare altri l parametri lagrangiani (η_1, \dots, η_l) , legati alle (q_1, \dots, q_l) tramite la trasformazione

$$q_i = q_i(\eta_1, \dots, \eta_l), \quad i = 1, 2, \dots, l, \tag{5.61}$$

e

$$\eta_j = \eta_j(q_1, \dots, q_l), \quad j = 1, 2, \dots, l, \tag{5.62}$$

tale che $\det \mathbb{J} \neq 0$, dove

$$\mathbb{J} = \begin{pmatrix} \frac{\partial q_1}{\partial \eta_1} & \frac{\partial q_2}{\partial \eta_1} & \dots & \frac{\partial q_l}{\partial \eta_1} \\ \frac{\partial q_1}{\partial \eta_2} & \frac{\partial q_2}{\partial \eta_2} & \dots & \frac{\partial q_l}{\partial \eta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial q_1}{\partial \eta_l} & \frac{\partial q_2}{\partial \eta_l} & \dots & \frac{\partial q_l}{\partial \eta_l} \end{pmatrix}.$$

Le coordinate x_i dell' i -esimo punto sono date da¹⁰ $x_i = x_i(\tilde{q}, t)$, oppure $x_i = x_i(\tilde{\eta}, t)$, dove

$$x_i(\tilde{\eta}, t) = x_i(\tilde{q}(\tilde{\eta}), t),$$

e viceversa. Avremo poi

$$\dot{q}_i = \sum_{k=1}^l \frac{\partial q_i}{\partial \eta_k} \dot{\eta}_k, \quad i = 1, \dots, l, \tag{5.63}$$

$$\dot{\eta}_i = \sum_{k=1}^l \frac{\partial \eta_i}{\partial q_k} \dot{q}_k, \quad i = 1, \dots, l, \tag{5.64}$$

¹⁰Ricordando che $\tilde{q} = (q_1, q_2, \dots, q_l)$ e $\dot{\tilde{q}} = (\dot{q}_1, \dots, \dot{q}_l)$, poniamo adesso $\tilde{\eta} = (\eta_1, \eta_2, \dots, \eta_l)$, e $\dot{\tilde{\eta}} = (\dot{\eta}_1, \dots, \dot{\eta}_l)$. Inoltre con $\tilde{q}(\tilde{\eta})$ intendiamo $\tilde{q}(\tilde{\eta}) = (q_1(\eta_1, \dots, \eta_l), \dots, q_l(\eta_1, \dots, \eta_l))$.

e

$$\begin{aligned}
 \sum_{s=1}^l \frac{\partial \mathbf{x}_i \cdot}{\partial q_s} \dot{q}_s &= \sum_{s=1}^l \frac{\partial \mathbf{x}_i}{\partial q_s} \sum_{k=1}^l \frac{\partial q_s \cdot}{\partial \eta_k} \dot{\eta}_k \\
 &= \sum_{k=1}^l \left(\sum_{s=1}^l \frac{\partial \mathbf{x}_i}{\partial q_s} \frac{\partial q_s}{\partial \eta_k} \right) \dot{\eta}_k \\
 &= \sum_{k=1}^l \frac{\partial \mathbf{x}_i \cdot}{\partial \eta_k} \dot{\eta}_k,
 \end{aligned}$$

per cui $\dot{\mathbf{x}}_i$ potrà essere espressa sia in termini delle coordinate q che in termini delle η ,

$$\dot{\mathbf{x}}_i = \begin{cases} \sum_{s=1}^l \frac{\partial \mathbf{x}_i}{\partial q_s} \dot{q}_s + \frac{\partial \mathbf{x}_i}{\partial t}, \\ \sum_{k=1}^l \frac{\partial \mathbf{x}_i}{\partial \eta_k} \dot{\eta}_k + \frac{\partial \mathbf{x}_i}{\partial t}, \end{cases}$$

Notiamo poi che dalla (5.63) discende¹¹

$$\frac{\partial \dot{q}_i}{\partial \dot{\eta}_j} = \sum_{k=1}^l \frac{\partial q_i}{\partial \eta_k} \underbrace{\frac{\partial \dot{\eta}_k}{\partial \dot{\eta}_j}}_{\delta_{kj}} = \frac{\partial q_i}{\partial \eta_j}. \quad (5.65)$$

La funzione di Lagrange nelle variabili η si ottiene a partire dalla \mathcal{L} , ovvero¹²

$$L(\tilde{\boldsymbol{\eta}}, \tilde{\dot{\boldsymbol{\eta}}}, t) = \mathcal{L} \left(\tilde{\mathbf{q}}(\tilde{\boldsymbol{\eta}}), \underbrace{\sum_{k=1}^l \frac{\partial q_1}{\partial \eta_k} \dot{\eta}_k}_{\dot{q}_1}, \dots, \underbrace{\sum_{k=1}^l \frac{\partial q_l}{\partial \eta_k} \dot{\eta}_k}_{\dot{q}_l}, t \right). \quad (5.66)$$

Dimostriamo adesso che L , data dalla (5.66), soddisfa le equazioni di Lagrange

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_i} \right) - \frac{\partial L}{\partial \eta_i} = 0, \quad i = 1, 2, \dots, l, \quad (5.67)$$

se \mathcal{L} soddisfa le (5.36). Si comincia col valutare

$$\frac{\partial L}{\partial \eta_k} = \sum_{i=1}^l \left(\frac{\partial \mathcal{L}}{\partial q_i} \frac{\partial q_i}{\partial \eta_k} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \eta_k} \right),$$

¹¹Si ricorda ancora che $\delta_{kj} = \begin{cases} 1 & \text{se } k = j \\ 0 & \text{se } k \neq j \end{cases}$

¹²Si noti che questo ci dice che le funzioni \mathcal{L} e L sono *funzionalmente* diverse: l'uguaglianza $\mathcal{L} = L$ deve essere intesa come uguaglianza dei valori delle due funzioni quando sono calcolate rispettivamente nelle coordinate che rappresentano lo stesso stato cinematico.

e poi

$$\frac{\partial L}{\partial \dot{\eta}_k} = \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \dot{\eta}_k} \stackrel{(5.65)}{=} \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial q_i}{\partial \eta_k}.$$

Abbiamo quindi

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_k} \right) &= \sum_{i=1}^l \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \frac{\partial q_i}{\partial \eta_k} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d}{dt} \left(\frac{\partial q_i}{\partial \eta_k} \right) \right] \\ &= \sum_{i=1}^l \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \frac{\partial q_i}{\partial \eta_k} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \eta_k} \right], \end{aligned}$$

da cui si ricava facilmente la tesi. Infatti

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_k} \right) - \frac{\partial L}{\partial \eta_k} = \sum_{i=1}^l \frac{\partial q_i}{\partial \eta_k} \underbrace{\left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} \right]}_{=0} = 0, \quad k = 1, 2, \dots, l.$$

Non solo, ma abbiamo ottenuto anche

$$\begin{pmatrix} \frac{\partial q_1}{\partial \eta_1} & \frac{\partial q_2}{\partial \eta_1} & \dots & \frac{\partial q_l}{\partial \eta_1} \\ \frac{\partial q_1}{\partial \eta_2} & \frac{\partial q_2}{\partial \eta_2} & \dots & \frac{\partial q_l}{\partial \eta_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial q_1}{\partial \eta_l} & \frac{\partial q_2}{\partial \eta_l} & \dots & \frac{\partial q_l}{\partial \eta_l} \end{pmatrix} \begin{pmatrix} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_1} \right) - \frac{\partial \mathcal{L}}{\partial q_1} \\ \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_2} \right) - \frac{\partial \mathcal{L}}{\partial q_2} \\ \vdots \\ \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_l} \right) - \frac{\partial \mathcal{L}}{\partial q_l} \end{pmatrix} = \begin{pmatrix} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_1} \right) - \frac{\partial L}{\partial \eta_1} \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_2} \right) - \frac{\partial L}{\partial \eta_2} \\ \vdots \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_l} \right) - \frac{\partial L}{\partial \eta_l} \end{pmatrix}.$$

Per cui se uno dei due vettori colonna è nullo lo è anche l'altro dal momento che $\det \mathbb{J} \neq 0$. Vale dunque il seguente

Teorema 5.5.1 *Le soluzioni di (5.67) si ottengono dalle soluzioni di (5.36) tramite le trasformazioni (5.62), (5.64).*

Questo risultato è, in un certo qual senso, ovvio: il moto è lo stesso, indipendentemente dal sistema di coordinate scelto. Di più, tutto il formalismo lagrangiano è stato, fin dall'inizio, concepito per scrivere le equazioni di moto *con la stessa ricetta* in qualsiasi sistema di coordinate. Quindi tale teorema è, se vogliamo, una verifica a posteriori della correttezza del formalismo stesso. Tuttavia questa invarianza "analitica" è importante quando si voglia introdurre un nuovo sistema di coordinate a partire da un preesistente sistema, senza "ritornare" al sistema meccanico, cioè senza tornare a ricalcolare la funzione di Lagrange dalla sua definizione meccanica.

Questo risultato si enuncia anche dicendo che **le equazioni di Lagrange sono invarianti per cambiamento di coordinate.**

5.6 Coordinate cicliche

In questa sezione vogliamo generalizzare al caso di N punti materiali i concetti già introdotti nel paragrafo 3.3.1. Come già visto, sappiamo che in certi casi una delle coordinate lagrangiane, diciamo q_i , non compaia esplicitamente nella funzione di Lagrange \mathcal{L} (ma la \dot{q}_i deve comparire), cioè

$$\frac{\partial \mathcal{L}}{\partial q_i} = 0. \quad (5.68)$$

Se ciò accade, ricordando la (5.35), si ha

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) = 0,$$

se $\Xi_j \equiv 0$, per ogni $j = 1, 2, \dots, l$ (il che avviene, per esempio, se $\tilde{\mathcal{S}} \equiv 0$). Quindi, riprendendo la definizione 3.3.1 del capitolo 3, diremo che la coordinata q_i è **ciclica**, se (5.68) è verificata. Ricordando la proposizione 3.3.2 ed **assumendo** $\tilde{\mathcal{S}} \equiv 0$, abbiamo che la funzione $\frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ è una costante del moto¹³ (o integrale primo),

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \text{costante},$$

dove “costante” significa costante rispetto al tempo. In accordo con la definizione 3.3.2 (ovvero con la definizione 9.0.1), $\frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ è detta **momento coniugato** alla variabile q_i .

Il punto che adesso vogliamo approfondire è una procedura, definita nell’ambito generale di sistemi con l gradi di libertà, che consenta di sfruttare efficacemente le variabili cicliche al fine di semplificare le equazioni di Lagrange (5.36). Ovviamente la strada è già stata tracciata nelle sezione 3.3.1. Adesso vogliamo generalizzarla supponendo che esistano $0 < \nu \leq l$, variabili cicliche. Assumiamo quindi che:

- q_α , $\alpha = 1, \dots, \nu$, **siano variabili cicliche**. Ciò implica che \mathcal{L} **non dipende esplicitamente dalle** q_α , cioè

$$\mathcal{L} = \mathcal{L}(\dot{q}_1, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t).$$

- q_i , $i = \nu + 1, \dots, l$, **siano non cicliche**.

Abbiamo quindi ν momenti coniugati

$$p_\alpha = \frac{\partial \mathcal{L}(\dot{q}_1, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t)}{\partial \dot{q}_\alpha}, \quad \alpha = 1, \dots, \nu, \quad (5.69)$$

che sono costanti. Vale questa importante

Proposizione 5.6.1 *Le ν relazioni (5.69) possono essere invertite e da esse si possono ricavare le \dot{q}_α , $\alpha = 1, \dots, \nu$, in funzione:*

¹³Stiamo assumendo di aver a che fare con un sistema di N punti materiali soggetto a vincoli olonomi lisci e sottoposto all’azione di forze puramente conservative (per intenderci, $\tilde{\mathcal{S}} \equiv 0$).

- dei ν momenti coniugati p_α , $\alpha = 1, \dots, \nu$;
- delle rimanenti $l - \nu$, \dot{q}_i , cioè delle \dot{q}_i , $i = \nu + 1, \dots, l$;
- di tutte le q_i non cicliche, cioè di q_i , $i = \nu + 1, \dots, l$;
- del tempo t .

Dim. Scriviamo le (5.69) in forma estesa

$$\begin{cases} p_1 = \Phi_1(\dot{q}_1, \dots, \dot{q}_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t), \\ p_2 = \Phi_2(\dot{q}_1, \dots, \dot{q}_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t), \\ \vdots \\ p_\nu = \Phi_\nu(\dot{q}_1, \dots, \dot{q}_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t), \end{cases} \quad (5.70)$$

dove $\Phi_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}$, $\alpha = 1, \dots, \nu$. Adesso guardiamo al sistema (5.70) come ad una relazione fra la ν -upla (p_1, \dots, p_ν) e la ν -upla $(\dot{q}_1, \dots, \dot{q}_\nu)$. Evidentemente tale relazione sarà biunivoca, e quindi invertibile, se la relativa matrice jacobiana

$$\mathbb{J} = \begin{pmatrix} \frac{\partial \Phi_1}{\partial \dot{q}_1} & \cdots & \frac{\partial \Phi_\nu}{\partial \dot{q}_1} \\ \vdots & & \vdots \\ \frac{\partial \Phi_1}{\partial \dot{q}_\nu} & \cdots & \frac{\partial \Phi_\nu}{\partial \dot{q}_\nu} \end{pmatrix},$$

è invertibile, ovvero se $\det \mathbb{J} \neq 0$. Ora, ricordando l'espressione delle Φ_α , si ha

$$\mathbb{J} = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial \dot{q}_1^2} & \cdots & \frac{\partial \mathcal{L}}{\partial \dot{q}_1 \partial \dot{q}_\nu} \\ \vdots & & \vdots \\ \frac{\partial^2 \mathcal{L}}{\partial \dot{q}_\nu \partial \dot{q}_1} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial \dot{q}_\nu^2} \end{pmatrix},$$

da cui, data la (5.34) con T espressa dalla (4.33), ricaviamo

$$\mathbb{J} = \begin{pmatrix} a_{11} & \cdots & a_{1\nu} \\ \vdots & & \vdots \\ a_{\nu 1} & \cdots & a_{\nu\nu} \end{pmatrix},$$

cioè \mathbb{J} coincide con un minore di ordine ν della matrice dell'energia cinetica \mathbb{A} , le cui componenti sono definite dalla (4.34). Ora il teorema 4.1.3 garantisce che $\det \mathbb{A} \neq 0$, e di conseguenza anche¹⁴ $\det \mathbb{J} \neq 0$.

□

¹⁴Ricordiamo che se con \mathbb{A}_j denotiamo il minore di ordine j di una matrice \mathbb{A} , vale il seguente:
Teorema. Sia \mathbb{A} una matrice simmetrica di ordine n . La forma quadratica relativa alla matrice \mathbb{A} è definita positiva se e solo se $\det \mathbb{A}_j > 0$, per ogni $j \leq n$.

Applicando quindi tale proposizione, potremo scrivere

$$\dot{q}_\alpha = \widehat{q}_\alpha(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t), \quad \alpha = 1, \dots, \nu, \quad (5.71)$$

e, ricordando la definizione (3.34) della sezione 3.3.1, introdurre la funzione di Routh¹⁵

$$\mathcal{R} = \widehat{\mathcal{L}} - \sum_{\alpha=1}^{\nu} p_\alpha \widehat{q}_\alpha, \quad (5.72)$$

dove, in analogia con la sezione 3.3.1, $\widehat{\mathcal{L}}$ denota la funzione di Lagrange \mathcal{L} , in cui, al posto delle ν variabili \dot{q}_α , $\alpha = 1, \dots, \nu$, sono state sostituite le espressioni (5.71)

$$\widehat{\mathcal{L}} \left(\underbrace{\widehat{q}_1(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t), \dots, \widehat{q}_\nu(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t)}_{\dot{q}_\alpha \text{ viste come funzioni delle rimanenti } \dot{q}, \text{ delle } p \text{ e di tutte le } q \text{ non cicliche}} \right) \\ \underbrace{\dot{q}_{\nu+1}, \dots, \dot{q}_l}_{\text{rimanenti } \dot{q}}, \underbrace{q_{\nu+1}, \dots, q_l, t}_{\text{tutte le } q \text{ non cicliche}}. \quad (5.73)$$

Di conseguenza

$$\mathcal{R} = \mathcal{R}(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t). \quad (5.74)$$

Lo ricordiamo ancora, p_1, \dots, p_ν , sono costanti il cui valore è stabilito dalle condizioni iniziali.

In particolare abbiamo che le equazioni per le $l - \nu$, coordinate non cicliche sono

$$\frac{d}{dt} \left(\frac{\partial \mathcal{R}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{R}}{\partial q_i} = 0, \quad i = \nu + 1, \dots, l. \quad (5.75)$$

Infatti, come già provato nella proposizione 3.3.3, considerando $i > \nu$, abbiamo

$$\frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_i} = \sum_{\alpha=1}^{\nu} \underbrace{\frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_\alpha}}_{p_\alpha} \frac{\partial \widehat{q}_\alpha}{\partial \dot{q}_i} + \frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_i} = \sum_{\alpha=1}^{\nu} p_\alpha \frac{\partial \widehat{q}_\alpha}{\partial \dot{q}_i} + \frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_i},$$

$$\frac{\partial \widehat{\mathcal{L}}}{\partial q_i} = \sum_{\alpha=1}^{\nu} \sum_{\alpha=1}^{\nu} \underbrace{\frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_\alpha}}_{p_\alpha} \frac{\partial \widehat{q}_\alpha}{\partial q_i} + \frac{\partial \widehat{\mathcal{L}}}{\partial q_i} = \sum_{\alpha=1}^{\nu} p_\alpha \frac{\partial \widehat{q}_\alpha}{\partial q_i} + \frac{\partial \widehat{\mathcal{L}}}{\partial q_i},$$

¹⁵Osserviamo che se q_α è ciclica per la Lagrangiana, cioè $\frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$, è ciclica anche per la funzione di Routh, ovvero $\frac{\partial \mathcal{R}}{\partial q_\alpha} = 0$, cioè \mathcal{R} non dipende esplicitamente da q_α .

e quindi¹⁶

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial \hat{\mathcal{L}}}{\partial \dot{q}_i} \right) - \frac{\partial \hat{\mathcal{L}}}{\partial q_i} &= \sum_{\alpha=1}^{\nu} p_{\alpha} \left[\frac{d}{dt} \left(\frac{\partial \hat{q}_{\alpha}}{\partial \dot{q}_i} \right) - \frac{\partial \hat{q}_{\alpha}}{\partial q_i} \right] + \\ &\quad \underbrace{\frac{d}{dt} \left(\frac{\partial \hat{\mathcal{L}}}{\partial \dot{q}_i} \right) - \frac{\partial \hat{\mathcal{L}}}{\partial q_i}}_{=0} \\ &= \frac{d}{dt} \left(\sum_{\alpha=1}^{\nu} p_{\alpha} \frac{\partial \hat{q}_{\alpha}}{\partial \dot{q}_i} \right) - \sum_{\alpha=1}^{\nu} p_{\alpha} \frac{\partial \hat{q}_{\alpha}}{\partial q_i} \\ &= \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_i} \left(\sum_{\alpha=1}^{\nu} p_{\alpha} \hat{q}_{\alpha} \right) \right) - \frac{\partial}{\partial q_i} \left(\sum_{\alpha=1}^{\nu} p_{\alpha} \hat{q}_{\alpha} \right). \end{aligned}$$

Pertanto, per ogni $i = \nu + 1, \dots, l$, cioè per ogni **coordinata non ciclica**, possiamo scrivere

$$\frac{d}{dt} \left[\frac{\partial}{\partial \dot{q}_i} \left(\hat{\mathcal{L}} - \sum_{\alpha=1}^{\nu} p_{\alpha} \hat{q}_{\alpha} \right) \right] - \frac{\partial}{\partial q_i} \left(\hat{\mathcal{L}} - \sum_{\alpha=1}^{\nu} p_{\alpha} \hat{q}_{\alpha} \right) = 0,$$

cioè le (5.75). Si può provare, seguendo essenzialmente lo stesso procedimento illustrato nella sezione 5.4, che il sistema di $l - \nu$ equazioni differenziali (5.75) ammette (almeno localmente) un'unica soluzione. Possiamo quindi, almeno in linea di principio, risolverlo e determinare le q_i non cicliche, cioè $q_i = q_i(t)$, $i = \nu + 1, \dots, l$.

Rimane ancora da determinare l'evoluzione delle ν coordinate cicliche. Possiamo utilizzare le ν equazioni (5.71) che costituiscono un sistema del prim'ordine in forma normale per le ν variabili q_{α} . Non solo, ma siccome la parte di destra della (5.71) non dipende dalle coordinate cicliche q_{α} , $\alpha = 1, \dots, \nu$, possiamo integrarla e così, per ogni $\alpha = 1, \dots, l$, ottenere

$$q_{\alpha}(t) = q_{\alpha}(0) + \int_0^t \hat{q}_{\alpha}(p_1, \dots, p_{\nu}, \dot{q}_{\nu+1}(\tau), \dots, \dot{q}_l(\tau), q_{\nu+1}(\tau), \dots, q_l(\tau), \tau) d\tau. \tag{5.76}$$

E' infine interessante notare che le funzioni \hat{q}_{α} , $\alpha = 1, \dots, \nu$, si possono derivare direttamente dalla funzione \mathcal{R} . Infatti, tenendo conto delle (5.71) e (5.73), derivando la (5.72) rispetto a p_{α} , abbiamo

$$\begin{aligned} -\frac{\partial \mathcal{R}}{\partial p_{\alpha}} &= \hat{q}_{\alpha} + \sum_{\beta=1}^{\nu} p_{\beta} \frac{\partial \hat{q}_{\beta}}{\partial p_{\alpha}} - \frac{\partial \hat{\mathcal{L}}}{\partial p_{\alpha}} \\ &= \hat{q}_{\alpha} + \sum_{\beta=1}^{\nu} p_{\beta} \frac{\partial \hat{q}_{\beta}}{\partial p_{\alpha}} - \underbrace{\sum_{\beta=1}^{\nu} \frac{\partial \hat{\mathcal{L}}}{\partial \hat{q}_{\beta}} \frac{\partial \hat{q}_{\beta}}{\partial p_{\alpha}}}_{p_{\beta}} = \hat{q}_{\alpha}. \end{aligned}$$

¹⁶Si ricordi che p_{α} è costante, per cui $\frac{dp_{\alpha}}{dt} = 0$, e $\frac{\partial p_{\alpha}}{\partial q_i} = 0$.

Quindi

$$\dot{q}_\alpha = - \underbrace{\frac{\partial \mathcal{R}}{\partial p_\alpha}}_{\hat{q}_\alpha}, \quad \alpha = 1, \dots, \nu. \quad (5.77)$$

In definitiva, se il sistema presenta $\nu \leq l$, variabili cicliche, il problema della determinazione del moto diminuisce di complessità. Infatti, introducendo la funzione di Routh (5.72), abbiamo $l - \nu$ (anziché l) equazioni di tipo lagrangiano, cioè le (5.75), che governano l'evoluzione delle $l - \nu$ variabili non cicliche, mentre l'evoluzione delle ν variabili cicliche è data dalla (5.76), che, sfruttando la (5.77), può anche essere così riscritta

$$q_\alpha(t) = q_\alpha(0) - \int_0^t \frac{\partial \mathcal{R}}{\partial p_\alpha} d\tau, \quad \alpha = 1, \dots, \nu.$$

5.7 La funzione Hamiltoniana e la conservazione dell'energia

Proseguiamo nella generalizzazione dei risultati della sezione 3.3.1 considerando un sistema caratterizzato da l parametri lagrangiani. Introduciamo la funzione **Hamiltoniana** la cui definizione, nel caso di un singolo punto materiale è data dalla (3.29), mentre per un sistema con l gradi di libertà è

$$\mathcal{H}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l, t) = \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{L}. \quad (5.78)$$

Calcolando la derivata totale di \mathcal{H} si ottiene

$$\frac{d\mathcal{H}}{dt} = \sum_{i=1}^l \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \dot{q}_i + \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} - \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} - \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial q_i} \dot{q}_i - \frac{\partial \mathcal{L}}{\partial t}. \quad (5.79)$$

Il secondo e terzo addendo si elidono a vicenda mentre il primo e il quarto termine generano, in virtù delle equazioni di Lagrange (5.35),

$$\sum_{i=1}^l \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} \right] \dot{q}_i = \sum_{i=1}^l \Xi_i \dot{q}_i.$$

Ne segue che, se $q_i(t)$, $i = 1, \dots, l$, soddisfano il sistema (5.35), allora

$$\frac{d\mathcal{H}}{dt} = \sum_{i=1}^l \Xi_i \dot{q}_i - \frac{\partial \mathcal{L}}{\partial t}, \quad (5.80)$$

e di conseguenza

$$\frac{d\mathcal{H}}{dt} = 0, \quad (5.81)$$

se la funzione Lagrangiana non dipende esplicitamente dal tempo e se $\sum_{i=1}^l \Xi_i \dot{q}_i = 0$. In particolare, dalla (5.32) otteniamo

$$\sum_{i=1}^l \Xi_i \dot{q}_i = \sum_{i=1}^l \tilde{\mathbf{S}} \cdot \tilde{\mathbf{u}}_i \dot{q}_i = \tilde{\mathbf{S}} \cdot \left(\sum_{i=1}^l \tilde{\mathbf{u}}_i \dot{q}_i \right),$$

dove, al solito \bullet è il prodotto scalare in \mathbb{R}^{3N} . Ora, ricordando la formula (4.32) e le considerazioni ad essa successive, $\sum_{i=1}^l \tilde{\mathbf{u}}_i \dot{q}_i$ rappresenta la velocità del sistema degli N punti materiali rispetto ai vincoli. Quindi, denotando con

$$\tilde{\mathbf{x}}_{vcl} = \sum_{i=1}^l \tilde{\mathbf{u}}_i \dot{q}_i,$$

la velocità del sistema rispetto ai vincoli, possiamo riscrivere la (5.80) in questa forma più compatta

$$\frac{d\mathcal{H}}{dt} = \tilde{\mathbf{S}} \bullet \tilde{\mathbf{x}}_{vcl} - \frac{\partial \mathcal{L}}{\partial t}, \quad (5.82)$$

la cui interpretazione fisica sarà evidente una volta chiarita l'interpretazione fisica di \mathcal{H} . Vogliamo quindi dare, come nella sezione 3.3.1, un'interpretazione fisica della funzione \mathcal{H} . Partendo dalla definizione (5.34) e ricordando la (4.33), abbiamo

$$\mathcal{L} = T_2 + T_1 + T_0 - V.$$

Eseguendo passaggi analoghi a quelli sviluppati nella sezione 3.3.1, è facile provare che¹⁷

$$\sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i = \sum_{i=1}^l \frac{\partial (T_2 + T_1 + T_0)}{\partial \dot{q}_i} \dot{q}_i = 2T_2 + T_1, \quad (5.83)$$

per cui,

$$\mathcal{H} = 2T_2 + T_1 - (T_2 + T_1 + T_0 - V) = T_2 + V - T_0. \quad (5.84)$$

In particolare, se $T_0 \equiv 0$, la funzione \mathcal{H} corrisponde alla somma dell'energia cinetica e dell'energia potenziale. Abbiamo quindi questa semplice

Proposizione 5.7.1 *Nel caso di un sistema olonomo a vincoli fissi, la funzione \mathcal{H} , corrisponde all'energia meccanica totale, cioè*

$$\mathcal{H} = T + V. \quad (5.85)$$

Dim. Nel caso in cui la Lagrangiana provenga da un sistema meccanico soggetto a vincoli fissi, abbiamo $\mathcal{L} = T_2 - V$, cioè $T \equiv T_2$, da cui discende immediatamente la (5.85).

□

La (5.85) è stata già incontrata nel paragrafo 3.3.1, dove avevamo a che fare con un solo punto materiale libero, cioè privo di vincoli. **Tuttavia sottolineiamo ancora che, in generale, $\mathcal{H} \neq T + V$, l'uguaglianza vale soltanto se $T \equiv T_2$.**

Prima di continuare, confrontando la (5.84) con la (5.85), viene naturale introdurre la cosiddetta **energia potenziale efficace**

$$V_{eff} = V - T_0, \quad (5.86)$$

¹⁷Questo calcolo è un caso particolare di un teorema sulle forme quadratiche (detto "di Eulero") che dice che, data la forma quadratica $A(\xi, \xi) = \sum_{i,j=1}^n a_{ij} \xi_i \xi_j$ con (a_{ij}) matrice simmetrica, allora $\sum_{k=1}^n \frac{\partial A}{\partial \xi_k} \xi_k = 2A(\xi, \xi)$.

ed interpretare la (5.84) come

$$\mathcal{H} = T_2 + V_{eff}, \quad (5.87)$$

cioè come **energia meccanica totale di un sistema soggetto a forze la cui energia potenziale è** (5.86).

Prima di procedere è tuttavia necessario rimarcare la differenza fra i due termini che compongono V_{eff} . Infatti V è proprio l'energia potenziale dovuta a forze conservative prodotte da ben precisi agenti fisici (il peso, una o più molle, interazioni elettrostatiche, etc.), $-T_0$ non è invece generata da nessun agente fisico se non dal moto dei vincoli. Questi, trascinando con il loro moto il sistema di punti materiali, sono i responsabili delle forze d'inerzia (rappresentate nell'Hamiltoniana \mathcal{H} proprio da $-T_0$).

Tornando alla (5.82), la conseguenza "fisica" di tale formula è evidente: la variazione nell'unità di tempo dell'energia meccanica di un sistema è imputabile a due termini:

- (i) $\tilde{\mathcal{S}} \bullet \tilde{\mathbf{x}}_{vcl}$, lavoro fatto, nell'unità di tempo (e quindi la potenza) dalle forze che non sono conservative¹⁸ (e che quindi che non ammettono energia potenziale);
- (ii) $\frac{\partial \mathcal{L}}{\partial t}$, che, in generale, è non nullo nel caso di vincoli mobili (anche se in taluni casi particolari può annullarsi pur avendo a che fare con vincoli in movimento).

Quindi, tornando all'energia meccanica totale espressa dalla (5.87) possiamo concludere che essa si conserva se $\frac{\partial \mathcal{L}}{\partial t} = 0$, e se $\tilde{\mathcal{S}} \bullet \tilde{\mathbf{x}}_{vcl} = 0$.

Nota 5.7.1 Osserviamo che il termine T_1 , cioè il termine omogeneo di primo grado nelle \dot{q}_i , non ha alcuna influenza sull'Hamiltoniana \mathcal{H} . Nella nota 5.3.2 abbiamo osservato che T_1 non influisce nelle equazioni di Lagrange **soltanto se non dipende esplicitamente dal tempo e se** $\frac{\partial b_i}{\partial q_i} = \frac{\partial b_j}{\partial q_j}$, $\forall i, j = 1, \dots, l$. Invece, per quel che riguarda la \mathcal{H} , T_1 è sempre ininfluenza, indipendentemente dalla sua forma.

Proseguendo il confronto con la sezione 3.3.1, concludiamo il paragrafo illustrando come sfruttare le eventuali coordinate cicliche anche nell'Hamiltoniana. Supponiamo, come nella sezione 5.6, di aver a che fare con un sistema olonomo a vincoli lisci soggetto soltanto a forze conservative, cioè $\tilde{\mathcal{S}} \equiv 0$, e che q_α , $\alpha = 1, \dots, \nu$, siano cicliche (cioè $\frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$, $\forall \alpha = 1, \dots, \nu$). Deduciamo allora che le q_α , $\alpha = 1, \dots, \nu$, sono cicliche anche per la funzione Hamiltoniana¹⁹. Infatti, dalla definizione (5.78),

$$\frac{\partial \mathcal{H}}{\partial q_\alpha} = \sum_{i=1}^l \frac{\partial}{\partial q_\alpha} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \dot{q}_i - \underbrace{\frac{\partial \mathcal{L}}{\partial q_\alpha}}_{=0} = \sum_{i=1}^l \frac{\partial}{\partial \dot{q}_i} \left(\frac{\partial \mathcal{L}}{\partial q_\alpha} \right) \dot{q}_i = 0.$$

¹⁸Osserviamo che le forze di attrito possono essere inserite in $\tilde{\mathcal{S}}$. Si tratta soltanto definire come reazioni vincolari $\tilde{\mathbf{R}}$ quella componente delle forze vincolari per cui $\tilde{\mathbf{R}} \bullet \delta \tilde{\mathbf{x}} = 0$, per ogni spostamento virtuale $\delta \tilde{\mathbf{x}}$. In tale "riclassificazione" delle forze le forze vincolari dovute all'attrito si inseriscono in $\tilde{\mathcal{S}}$, insieme a tutte le altre (eventuali) forze che non sono conservative.

¹⁹Come del resto lo sono anche per la funzione di Routh \mathcal{R} (vedi nota 15) a piè di pagina.

Procediamo quindi come nella sezione 3.3.1, introducendo

$$\widehat{\mathcal{H}}(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t) = \mathcal{H}(\widehat{q}_1(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_1, \dots, q_l, t), \dots, \widehat{q}_\nu(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_1, \dots, q_l, t), \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t). \quad (5.88)$$

Dimostriamo che

$$\widehat{\mathcal{H}}(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t) = \sum_{i=\nu+1}^l \frac{\partial \mathcal{R}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{R},$$

dove \mathcal{R} è la funzione di Routh, data dalla (5.72), che dipende da p_α , $\alpha = 1, \dots, \nu$, e dalle variabili non cicliche (si veda la (5.74)). Con gli stessi passaggi illustrati nella formula (3.36), si prova che

$$\frac{\partial \mathcal{R}}{\partial \dot{q}_i} = \frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_i}, \quad i = \nu + 1, \dots, l,$$

dove $\widehat{\mathcal{L}}$ è data dalla (5.73). Quindi, osservando che

$$\begin{aligned} \mathcal{H} &= \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{L} = \left[\sum_{\alpha=1}^{\nu} \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \dot{q}_\alpha + \sum_{i=\nu+1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i \right] - \mathcal{L} \\ &= \sum_{\alpha=1}^{\nu} p_\alpha \dot{q}_\alpha + \sum_{i=\nu+1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{L}, \end{aligned}$$

abbiamo

$$\begin{aligned} \widehat{\mathcal{H}}(p_1, \dots, p_\nu, \dot{q}_{\nu+1}, \dots, \dot{q}_l, q_{\nu+1}, \dots, q_l, t) &= \sum_{\alpha=1}^{\nu} p_\alpha \widehat{q}_\alpha + \sum_{i=\nu+1}^l \frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_i} \dot{q}_i - \widehat{\mathcal{L}} \\ &= \sum_{i=\nu+1}^l \underbrace{\frac{\partial \widehat{\mathcal{L}}}{\partial \dot{q}_i}}_{\frac{\partial \mathcal{R}}{\partial \dot{q}_i}} \dot{q}_i - \underbrace{\left(\widehat{\mathcal{L}} - \sum_{\alpha=1}^{\nu} p_\alpha \widehat{q}_\alpha \right)}_{\mathcal{R}} \end{aligned}$$

La versione “semplificata” dell’energia meccanica si ottiene inserendo \widehat{q}_α , $\alpha = 1, \dots, \nu$, nella Hamiltoniana \mathcal{H} . Così facendo si ottiene la nuova Hamiltoniana $\widehat{\mathcal{H}}$, che dipende soltanto dalle $l - \nu$ variabili non cicliche e dalle ν costanti p_α , $\alpha = 1, \dots, \nu$.

Esempio 5.7.1 *Proseguiamo con lo studio del moto di un punto materiale P su una circonferenza ruotante, moto già trattato negli esempi 4.1.4 e 4.1.6, determinando la funzione Hamiltoniana nell’ipotesi che il vincolo sia liscio e che l’unica forza non vincolare agente su P sia la forza peso diretta nel verso opposto dell’asse z . L’energia potenziale della forza peso è data da*

$$V(x, y, z) = mgz, \Rightarrow \widehat{V}(\theta) = -mgR \cos \theta,$$

e quindi, ricordando la (4.41), la funzione Lagrangiana è

$$\mathcal{L}(\theta, \dot{\theta}, t) = T - \hat{V} = \frac{mR^2}{2} \left[\dot{\theta}^2 + \left(\frac{d\varphi}{dt} \right)^2 \sin^2 \theta \right] + mgR \cos \theta,$$

che può dipendere esplicitamente dal tempo tramite $\frac{d\varphi}{dt}$. Evidentemente se $\frac{d\varphi}{dt}$ è costante, il che vuol dire che la circonferenza ruota uniformemente attorno all'asse z , cioè

$$\frac{d\varphi}{dt} = \omega, \quad \Leftrightarrow \quad \varphi(t) = \omega t + \varphi_0,$$

allora la funzione \mathcal{L} **non dipende esplicitamente dal tempo** t . Sfruttando la (5.84) e la (4.42) dove sono specificate T_2 e T_0 , abbiamo

$$\begin{aligned} \mathcal{H} &= T_2 + V - T_0 \\ &= \frac{mR^2}{2} \dot{\theta}^2 - mgR \cos \theta - \frac{mR^2}{2} \left(\frac{d\varphi}{dt} \right)^2 \sin^2 \theta. \end{aligned}$$

Quindi $V_{\text{eff}} = V - T_0 = -mgR \cos \theta - \frac{mR^2}{2} \left(\frac{d\varphi}{dt} \right)^2 \sin^2 \theta$. La variazione nel tempo di \mathcal{H} è data

$$\frac{d\mathcal{H}}{dt} = -\frac{\partial \mathcal{L}}{\partial t} = mR^2 \frac{d\varphi}{dt} \frac{d^2\varphi}{dt^2} \sin^2 \theta,$$

e quindi $\frac{d\mathcal{H}}{dt} = 0$, se $\frac{d\varphi}{dt} \equiv 0$, cioè vincolo è fisso, o se $\frac{d^2\varphi}{dt^2} \equiv 0$, cioè la rotazione è uniforme. Notiamo che \mathcal{H} si conserva anche se il vincolo è mobile, a patto che la rotazione sia uniforme.

Esempio 5.7.2 *E' dato un punto materiale P vincolato a muoversi su una superficie di rotazione ottenuta ruotando attorno all'asse z la curva $z = f(x)$. La superficie è priva di attrito ed il punto è soggetto alla forza peso ed a quella esercitata da una molla di costante elastica k , ed avente lunghezza a riposo e massa trascurabili, il cui secondo estremo è fissato in O .*

La forma implicita della superficie è $z = f(x^2 + y^2)$. Una descrizione parametrica può essere data in termini di $r = x^2 + y^2$, distanza di P dall'asse z , e dell'angolo ϕ che la proiezione del vettore $(P - O)$ sul piano (x, y) , forma con l'asse x ,

$$\begin{cases} x = r \cos \phi, \\ y = r \sin \phi, \\ z = f(r), \end{cases} \quad \phi \in (-\pi, \pi], \quad r > 0.$$

I vettori base del piano tangente sono

$$\begin{aligned} \mathbf{u}_r &= \cos \phi \mathbf{e}_1 + \sin \phi \mathbf{e}_2 + f'(r) \mathbf{e}_3, \\ \mathbf{u}_\phi &= -r \sin \phi \mathbf{e}_1 + r \cos \phi \mathbf{e}_2, \end{aligned}$$

per cui $|\mathbf{u}_r \wedge \mathbf{u}_\phi|^2 = r^2 (f'^2 + 1) \neq 0$, se $r \neq 0$. L'energia cinetica del punto è

$$\begin{aligned} T &= \frac{m}{2} \begin{pmatrix} \dot{r} & \dot{\phi} \end{pmatrix} \begin{pmatrix} 1 + (f'(r))^2 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \dot{r} \\ \dot{\phi} \end{pmatrix} \\ &= \frac{m}{2} \left[(1 + (f'(r))^2) \dot{r}^2 + r^2 \dot{\phi}^2 \right], \end{aligned}$$

mentre l'energia potenziale dovuta al peso è $V_{\text{peso}}(x, y, z) = mgz$, e quindi $\hat{V}_{\text{peso}} = mgf(r)$, e quella dovuta alla molla è $V_{\text{molla}} = \frac{k}{2} (x^2 + y^2 + z^2)$, ovvero $\hat{V}_{\text{molla}} = \frac{k}{2} (r^2 + f^2(r))$. La Lagrangiana è

$$\mathcal{L} = \frac{m}{2} \left[(1 + (f'(r))^2) \dot{r}^2 + r^2 \dot{\phi}^2 \right] - mgf(r) - \frac{k}{2} (r^2 + f^2(r)).$$

La coordinata ϕ è ciclica. Infatti

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) - \underbrace{\frac{\partial \mathcal{L}}{\partial \phi}}_{=0} = 0, \Rightarrow \underbrace{\frac{\partial \mathcal{L}}{\partial \dot{\phi}}}_{mr^2 \dot{\phi}} = A_o = \text{costante}.$$

L'equazione per r è

$$\left(1 + (f'(r))^2 \right) \ddot{r} + f' f'' \dot{r}^2 - r \dot{\phi}^2 + gf' + \frac{k}{m} (r + f' f'') = 0,$$

che scriveremo così

$$\left(1 + (f'(r))^2 \right) \ddot{r} + f' f'' \dot{r}^2 - \frac{A_o^2}{m^2 r^3} + gf' + \frac{k}{m} (r + f' f'') = 0. \quad (5.89)$$

La (5.89) rientra nella classe di equazioni del secondo ordine analizzate nella sezione 2.6.3. Dobbiamo quindi scrivere la (5.89) nella corrispondente forma conservativa, individuando l'energia meccanica, ossia la funzione di Hamilton

$$\mathcal{H} = \frac{\partial T}{\partial \dot{r}} \dot{r} + \frac{\partial T}{\partial \dot{\phi}} \dot{\phi} - \mathcal{L}.$$

Infatti, essendo il vincolo liscio e fisso, $\frac{d\mathcal{H}}{dt} = 0$. Scriveremo quindi

$$\mathcal{H} = \frac{m}{2} \left[(1 + (f'(r))^2) \dot{r}^2 + r^2 \dot{\phi}^2 \right] + mgf(r) + \frac{k}{2} (r^2 + f^2(r)). \quad (5.90)$$

Lo studio qualitativo può essere svolto a partire direttamente dalla (5.90). Infatti, sostituendo $\dot{\phi}$ in termini di A_o e r , otteniamo $\hat{\mathcal{H}}$. Come caso particolare consideriamo $f(r) = cr^2$, e supponiamo $k = 0$ (assenza della molla), per cui

$$\hat{\mathcal{H}}(\dot{r}, r) = \frac{m}{2} \left[(1 + 4c^2 r^2) \dot{r}^2 + \frac{A_o^2}{m^2 r^2} \right] + mgr^2 = E,$$

dove E è costante e si calcola a partire dalle condizioni iniziali. Avremo quindi

$$(1 + 4c^2 r^2) \dot{r}^2 = \frac{E}{m} - \hat{V}_{eff}(r), \quad \text{con} \quad \hat{V}_{eff}(r) = \frac{1}{m} \left(mgr^2 + \frac{A_o^2}{2mr^2} \right). \quad (5.91)$$

In particolare, siccome il segno del primo membro della (5.91) è sempre non negativo, gli r "accessibili" sono quelli per cui su cui $\hat{V}_{eff}(r) \leq \frac{E}{m}$.

Esempio 5.7.3 Consideriamo un punto materiale P vincolato a muoversi su una superficie di rotazione ottenuta ruotando attorno all'asse z la curva $x = h(z)$, con $h(z) \geq 0$. La superficie è liscia ed il punto è soggetto alla forza peso diretta nel verso opposto a quello dell'asse z .

Una descrizione parametrica della superficie è

$$\begin{cases} z \in \mathbb{R} \\ x = h(z) \cos \varphi, \\ y = h(z) \sin \varphi, \end{cases}$$

dove φ è l'angolo che la proiezione del vettore $(P - O)$ sul piano x, y forma con l'asse x . La forma implicita della superficie è

$$f(x, y, z) = x^2 + y^2 - h^2(z) = 0.$$

I vettori base del piano tangente sono

$$\begin{aligned} \mathbf{u}_z &= h'(z) \cos \varphi \mathbf{e}_1 + h'(z) \sin \varphi \mathbf{e}_2 + \mathbf{e}_3, \\ \mathbf{u}_\varphi &= -h(z) \sin \varphi \mathbf{e}_1 + h(z) \cos \varphi \mathbf{e}_2, \end{aligned}$$

per cui l'energia cinetica è

$$\begin{aligned} T &= \frac{m}{2} \begin{pmatrix} \dot{z} & \dot{\varphi} \end{pmatrix} \begin{pmatrix} 1 + (h'(z))^2 & 0 \\ 0 & h^2(z) \end{pmatrix} \begin{pmatrix} \dot{z} \\ \dot{\varphi} \end{pmatrix} \\ &= \frac{m}{2} \left[(1 + (h'(z))^2) \dot{z}^2 + h^2(z) \dot{\varphi}^2 \right]. \end{aligned}$$

L'energia potenziale è $V_{peso}(x, y, z) = mgz$, e quindi $\hat{V}_{peso} = mgz$, che comporta la seguente funzione di Lagrange

$$\mathcal{L} = \frac{m}{2} \left[(1 + (h'(z))^2) \dot{z}^2 + h^2(z) \dot{\varphi}^2 \right] - mgz.$$

Da essa si ricava immediatamente che φ è ciclica, cioè

$$\varphi = \frac{A_o}{m} \frac{1}{h^2(z)}, \quad \text{con} \quad A_o = \text{costante}. \quad (5.92)$$

La funzione di Hamilton è dunque

$$\mathcal{H} = \frac{m}{2} \left[(1 + (h'(z))^2) \dot{z}^2 + h^2(z) \dot{\varphi}^2 \right] - mgz,$$

che, sostituendo $\dot{\varphi}$ con l'espressione data dalla (5.92), assume questa forma

$$\widehat{\mathcal{H}}(\dot{z}, z) = \frac{m}{2} \left(1 + (h'(z))^2 \right) \dot{z}^2 + \frac{A_o^2}{2m} \frac{1}{h^2(z)} + mgz.$$

L'energia potenziale efficace è dunque

$$V_{\text{eff}}(z) = \frac{A_o^2}{2m} \frac{1}{h^2(z)} + mgz.$$

Esempio 5.7.4 Come ultimo esempio consideriamo quello di un punto materiale vincolato a stare sulla cicloide (4.17) e rappresentata nella figura 5.5. In tal caso la

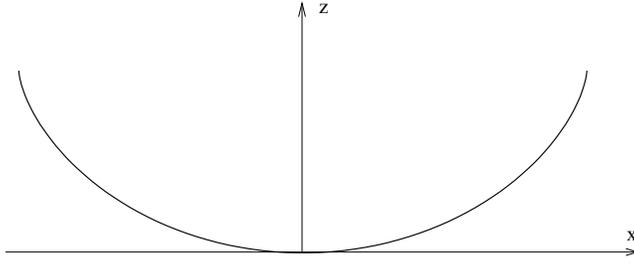


Figura 5.5: Rappresentazione della curva (4.17), detta cicloide.

particella ha un solo grado di libertà: ψ . Si suppone che il vincolo sia liscio e che il punto materiale, la cui massa è m , sia soggetto soltanto all'azione della forza peso diretta nel verso opposto dell'asse z . È facile scrivere la funzione Hamiltoniana dal momento che

$$T = mR^2\dot{\psi}^2(1 + \cos \psi), \quad V = mgR(1 - \cos \psi), \quad (5.93)$$

e quindi

$$\mathcal{H} = mR^2\dot{\psi}^2(1 + \cos \psi) + mgR(1 - \cos \psi).$$

Se poniamo

$$q^2 = R(1 - \cos \psi), \quad (5.94)$$

Derivando (5.94) si ottiene

$$R^2\dot{\psi}^2 = 4 \frac{q^2 \dot{q}^2}{\sin^2 \psi}, \quad (5.95)$$

e sostituendo nella (5.93)

$$T = 4m \frac{q^2 \dot{q}^2}{\sin^2 \psi} (1 + \cos \psi) = 4mR\dot{q}^2, \quad V = mgq^2, \quad (5.96)$$

da cui si ricava

$$\mathcal{H} = 4mR\dot{q}^2 + mgq^2,$$

che è la funzione di Hamilton dell'oscillatore armonico, $\dot{q}^2 + \omega^2 q^2$, se $\omega^2 = g/4R^2$. Si deduce quindi che le oscillazioni della particella sotto l'azione della sola forza peso non dipendono dall'energia. La variabile q compie un'oscillazione armonica di

periodo $T = 4\pi\sqrt{\frac{R}{g}}$, indipendente dalle condizioni iniziali ovvero dalla quota massima (purché minore di R) che si raggiunge durante il moto. Durante una oscillazione completa della variabile q la variabile $z = q^2$ compie due oscillazioni complete (una quando $q > 0$ e una quando $q < 0$) e quindi il periodo di oscillazione del punto sulla cicloide è dato da $T/2 = 2\pi\sqrt{\frac{R}{g}}$.

5.8 Il teorema di Noether

La conservazione dei momenti coniugati²⁰ p_α , $\alpha = 1, \dots, \nu$, nel caso delle variabili cicliche e la conservazione della funzione Hamiltoniana, sono legati alle proprietà di invarianza della funzione di Lagrange. Nel caso di una coordinata ciclica q_h , la condizione di ciclicità, $\frac{\partial \mathcal{L}}{\partial q_h} = 0$, ci dice che comunque si “trasli” la coordinata q_h la funzione Lagrangiana rimane immutata (in questo caso la cosa è “ovvia” visto che la Lagrangiana non dipende da q_h).

Cerchiamo di chiarire questo punto con qualche esempio. Come primo esempio consideriamo il moto centrale analizzato nella sezione 3.4. In questo caso la variabile φ è ciclica. Questo semplicemente traduce il fatto che, essendo tutto il sistema simmetrico per rotazioni attorno all’asse perpendicolare al piano del moto. La variabile φ non deve intervenire nella dinamica: solo la sua variazione nel tempo, cioè $\dot{\varphi}$, ha importanza. Un secondo esempio è quello del punto materiale libero (cioè una particella a cui non sono applicate forze). In tal caso la Lagrangiana è

$$\mathcal{L} \equiv T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

Qui tutte e tre le coordinate (x, y, z) sono cicliche. E questo è la “conseguenza” della omogeneità dello spazio: ogni punto dello spazio è indistinguibile dagli altri. In termini matematici apparentemente più raffinati, il problema del moto del punto libero deve essere invariante per traslazione (o meglio ancora, per l’azione del gruppo delle traslazioni dello spazio euclideo).

Anche la conservazione dell’energia, seppur meno palesemente, è legata a un’invarianza traslazionale: quella delle traslazioni temporali (cioè della variabile t).

Per generalizzare queste osservazioni occorre formalizzare meglio il concetto di **invarianza**. Per evitare complicazioni limitiamoci soltanto a funzioni Lagrangiane indipendenti dal tempo. La funzione \mathcal{L} è quindi una funzione definita su un insieme $\mathcal{S} = D \times \mathbb{R}^l$ dove $D \subset \mathbb{R}^l$ è un aperto, ed è il dominio delle coordinate lagrangiane²¹ $\tilde{q} = (q_1, \dots, q_l)$, mentre il secondo \mathbb{R}^l è il dominio delle velocità lagrangiane $\tilde{q} = (\dot{q}_1, \dots, \dot{q}_l)$.

Definizione 5.8.1 *Uno pseudo-gruppo²² ad un parametro di diffeomorfismi di D è un’applicazione da $\Phi : I \times D \rightarrow D$, dove $I \subset \mathbb{R}$ è un intervallo contenente lo zero e*

²⁰Si veda la formula (5.69).

²¹Utilizziamo, per semplicità, la notazione introdotta nella sezione 4.1.

²²Il nome “pseudo-gruppo” proviene dal fatto che in genere $I \neq \mathbb{R}$ e quindi la terza proprietà richiesta è solo “localmente” una proprietà di gruppo. Nel seguito, per non appesantire, elimineremo il prefisso “pseudo”

D è un aperto di \mathbb{R}^l

$$\underbrace{(s, \tilde{\mathbf{q}})}_{(s, (q_1, \dots, q_l))} \xrightarrow{\Phi} \Phi(s, \tilde{\mathbf{q}}) = (\Phi_1(s, \tilde{\mathbf{q}}), \Phi_2(s, \tilde{\mathbf{q}}), \dots, \Phi_l(s, \tilde{\mathbf{q}})) \in D \subset \mathbb{R}^l,$$

che gode delle seguenti proprietà:

1. per ogni $s \in I$ la mappa $\Phi(s, \cdot) : D \rightarrow D$ è un diffeomorfismo (ovvero è un'applicazione biunivoca da D in D , differenziabile con inversa differenziabile); per ogni $\tilde{\mathbf{q}} \in D$ la mappa $\Phi(\cdot, \tilde{\mathbf{q}}) : I \rightarrow D$ è differenziabile;
2. $\Phi(0, \cdot) : D \rightarrow D$ è la trasformazione identica, cioè $\Phi(0, \tilde{\mathbf{q}}) = \tilde{\mathbf{q}}$ per ogni $\tilde{\mathbf{q}} \in D$;
3. per ogni $s_1, s_2 \in I$ tali che $s_1 + s_2 \in I$, vale $\Phi(s_1, \Phi(s_2, \cdot)) = \Phi(s_2, \Phi(s_1, \cdot)) = \Phi(s_1 + s_2, \cdot)$.

Nota 5.8.1 Fissato un qualunque punto di D , ovvero una qualunque l -upla $\tilde{\mathbf{q}} \in D$, consideriamo l'applicazione $\Phi(s, \tilde{\mathbf{q}}) : I \rightarrow D$, cioè $s \rightarrow \Phi(s, \tilde{\mathbf{q}})$. Questa rappresenta una "curva" in D . Evidentemente, per ogni $\tilde{\mathbf{q}} \in D$ passa una sola curva del tipo $\Phi(s, \tilde{\mathbf{q}})$. Tuttavia la "rappresentazione" della curva non è univoca a causa della proprietà di traslazione al punto 3.

Esempio 5.8.1 Mostriamo che la **traslazione della coordinata lagrangiana** q_1 , ovvero l'applicazione $\Phi : \mathbb{R} \times \mathbb{R}^l \rightarrow \mathbb{R}^l$, così definita

$$(s, (q_1, \dots, q_l)) \xrightarrow{\Phi} (q_1 + s, q_2, \dots, q_l)$$

è un gruppo a un parametro. Le proprietà 1 e 2 sono soddisfatte. Proviamo la proprietà 3. Abbiamo

$$\begin{aligned} \Phi(s_1, \Phi(s_2, \tilde{\mathbf{q}})) &= \Phi(s_1, (q_1 + s_2, q_2, \dots, q_l)) = (q_1 + s_2 + s_1, q_2, \dots, q_l) \\ &= \Phi(s_1 + s_2, (q_1, q_2, \dots, q_l)), \end{aligned}$$

e

$$\begin{aligned} \Phi(s_2, \Phi(s_1, \tilde{\mathbf{q}})) &= \Phi(s_2, (q_1 + s_1, q_2, \dots, q_l)) = (q_1 + s_1 + s_2, q_2, \dots, q_l) \\ &= \Phi(s_1 + s_2, (q_1, q_2, \dots, q_l)), \end{aligned}$$

e quindi la proprietà 3 è verificata.

Nota 5.8.2 E' facile verificare che se $\Phi(s, \tilde{\mathbf{q}})$ è un gruppo ad un parametro, allora $\Phi(-s, \tilde{\mathbf{q}})$ è l'applicazione inversa di $\Phi(s, \tilde{\mathbf{q}})$, ovvero che $\Phi(-s, \Phi(s, \tilde{\mathbf{q}})) = \tilde{\mathbf{q}}$. Applicando infatti la proprietà 3 abbiamo $\Phi(-s, \Phi(s, \tilde{\mathbf{q}})) = \Phi(s - s, \tilde{\mathbf{q}}) = \Phi(0, \tilde{\mathbf{q}})$, che, per la proprietà 2, è proprio l'identità.

Definizione 5.8.2 Sia Φ un gruppo a un parametro nel senso della definizione 5.8.1. Chiameremo **rilevamento del gruppo allo spazio** $\mathcal{S} = D \times \mathbb{R}^l \subset \mathbb{R}^{2l}$ il gruppo a un parametro $T\Phi : I \times \mathcal{S} \rightarrow \mathcal{S}$, così definito:

$$\begin{aligned} T\Phi \left(s, \tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}} \right) &= T\Phi \left(s, (q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l) \right) \\ &= \left(\Phi_1(s, \tilde{\mathbf{q}}), \dots, \Phi_l(s, \tilde{\mathbf{q}}), \sum_{k=1}^l \frac{\partial \Phi_1}{\partial q_k} \dot{q}_k, \dots, \sum_{k=1}^l \frac{\partial \Phi_l}{\partial q_k} \dot{q}_k \right) \end{aligned} \quad (5.97)$$

Notiamo che se leggiamo la generica componente $\Phi_j(s, \tilde{\mathbf{q}})$, $j = 1, \dots, l$, della l -upla $\Phi(s, \tilde{\mathbf{q}})$ come la j -esima coordinata di una particella che si “muove” sotto l’azione del gruppo di diffeomorfismi, allora $\sum_{k=1}^l \frac{\partial \Phi_j}{\partial q_k} \dot{q}_k$, altro non è che la velocità della particella lungo la j -esima coordinata. Infatti, tornando alla sezione 4.1.2, se le coordinate dell’ i -esimo punto materiale $\mathbf{x}_i(\tilde{\mathbf{q}}) = (x_i(\tilde{\mathbf{q}}), y_i(\tilde{\mathbf{q}}), z_i(\tilde{\mathbf{q}}))$ non dipendono esplicitamente dal tempo, allora considerando, a titolo di esempio, la componente y_i del vettore posizione, la velocità del punto materiale rispetto ad y è data da $\dot{y}_i = \sum_{k=1}^l \frac{\partial y_i}{\partial q_k} \dot{q}_k$. Quindi, volendo proseguire questa analogia, se le (q_1, \dots, q_l) vengono mappate dal gruppo nelle $\Phi(s, \tilde{\mathbf{q}})$ (che leggiamo come l -upla delle coordinate) è ovvio che le $\tilde{\dot{\mathbf{q}}}$ dovranno essere mappate nelle “velocità”, ovvero in $\sum_{k=1}^l \frac{\partial \Phi_j}{\partial q_k} \dot{q}_k$, $j = 1, \dots, l$.

Proposizione 5.8.1 Il rilevamento di un gruppo a un parametro è esso stesso un gruppo a un parametro che agisce sui vettori di \mathbb{R}^{2l} .

Dim. Per prima cosa bisogna mostrare che, fissato un qualunque $s \in I$, l’applicazione $T\Phi \left(s, \tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}} \right) : \mathcal{S} \rightarrow \mathcal{S}$, è invertibile, sapendo che $\Phi(s, \tilde{\mathbf{q}})$ è un gruppo ad un parametro e che quindi soddisfa la proprietà 1 della definizione 5.8.1. Proprio in virtù di tale proprietà, abbiamo che, per ogni $s \in I$ e per ogni $\tilde{\mathbf{q}} \in D$, $\det \mathbb{J}_\Phi \neq 0$ dove \mathbb{J}_Φ è la matrice jacobiana

$$\mathbb{J}_\Phi = \begin{pmatrix} \frac{\partial \Phi_1}{\partial q_1} & \dots & \frac{\partial \Phi_l}{\partial q_1} \\ \vdots & & \vdots \\ \frac{\partial \Phi_1}{\partial q_l} & \dots & \frac{\partial \Phi_l}{\partial q_l} \end{pmatrix}.$$

Se calcoliamo la matrice jacobiana dell’applicazione $T\Phi$ abbiamo

$$\mathbb{J}_{T\Phi} = \begin{pmatrix} \mathbb{J}_\Phi & \mathbb{D} \\ 0 & \mathbb{J}_\Phi \end{pmatrix},$$

dove \mathbb{D} è la seguente matrice $l \times l$

$$\mathbb{D} = \begin{pmatrix} \sum_{k=1}^l \frac{\partial^2 \Phi_1}{\partial q_1 \partial q_k} \dot{q}_k & \dots & \sum_{k=1}^l \frac{\partial^2 \Phi_l}{\partial q_1 \partial q_k} \dot{q}_k \\ \vdots & & \vdots \\ \sum_{k=1}^l \frac{\partial^2 \Phi_1}{\partial q_l \partial q_k} \dot{q}_k & \dots & \sum_{k=1}^l \frac{\partial^2 \Phi_l}{\partial q_l \partial q_k} \dot{q}_k \end{pmatrix}.$$

Sfruttando le proprietà delle matrici triangolari a blocchi, $\det \mathbb{J}_{T\Phi} = (\det \mathbb{J}_\Phi)^2 \neq 0$, e quindi $T\Phi$ è invertibile.

Passiamo adesso alla proprietà 2. Siccome $\Phi(0, \tilde{\mathbf{q}}) = \tilde{\mathbf{q}}$, abbiamo

$$\begin{aligned} (0, \tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) \xrightarrow{T\Phi} & \left(\underbrace{(\Phi_1(0, \tilde{\mathbf{q}}), \dots, \Phi_l(0, \tilde{\mathbf{q}}))}_{(q_1, \dots, q_l)}, \sum_{k=1}^l \frac{\partial \Phi_1(0, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k, \dots \right. \\ & \left. \dots, \sum_{k=1}^l \frac{\partial \Phi_l(0, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k \right). \end{aligned}$$

Ora, per ogni $j = 1, \dots, l$, si ha $\Phi_j(0, \tilde{\mathbf{q}}) = q_j$, e quindi $\sum_{k=1}^l \frac{\partial \Phi_j(0, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k = \sum_{k=1}^l \delta_{jk} \dot{q}_k = \dot{q}_j$. Dunque

$$(0, \tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) \xrightarrow{T\Phi} (q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l),$$

ovvero $T\Phi(0, \tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}})$ è l'identità.

Mostriamo infine la proprietà 3, provando che $T\Phi(s_1, T\Phi(s_2, \cdot)) = T\Phi(s_1 + s_2, \cdot)$. Il passaggio $T\Phi(s_1, T\Phi(s_2, \cdot)) = T\Phi(s_2, T\Phi(s_1, \cdot))$ è banale. Dalla (5.97)

$$T\Phi(s_2, \tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) = \left(\Phi(s_2, \tilde{\mathbf{q}}), \sum_{k=1}^l \frac{\partial \Phi_1(s_2, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k, \dots, \sum_{k=1}^l \frac{\partial \Phi_l(s_2, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k \right)$$

e quindi

$$\begin{aligned} & T\Phi\left(s_1, \underbrace{\Phi(s_2, \tilde{\mathbf{q}}), \sum_{k=1}^l \frac{\partial \Phi_1(s_2, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k, \dots, \sum_{k=1}^l \frac{\partial \Phi_l(s_2, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k}_{T\Phi(s_2, \tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}})}\right) \\ &= \left(\Phi(s_1, \Phi(s_2, \tilde{\mathbf{q}})), \sum_{k=1}^l \frac{\partial \Phi_1(s_1, \Phi(s_2, \tilde{\mathbf{q}}))}{\partial q_k} \dot{q}_k, \dots \right. \\ & \quad \left. \dots, \sum_{k=1}^l \frac{\partial \Phi_l(s_1, \Phi(s_2, \tilde{\mathbf{q}}))}{\partial q_k} \dot{q}_k \right), \end{aligned}$$

che, sfruttando la proprietà 3 della definizione 5.8.1, diventa

$$\begin{aligned} &= \left(\Phi(s_1 + s_2, \tilde{\mathbf{q}}), \sum_{k=1}^l \frac{\partial \Phi_1(s_1 + s_2, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k, \dots, \sum_{k=1}^l \frac{\partial \Phi_l(s_1 + s_2, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k \right) \\ &= T\Phi(s_1 + s_2, \tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}). \end{aligned}$$

□

Esempio 5.8.2 Riprendendo l'esempio 5.8.1, il rilevamento della "traslazione" è dato da

$$(s, q_1, \dots, q_l; \dot{q}_1, \dots, \dot{q}_l) \xrightarrow{T\Phi} (q_1 + s, \dots, q_l; \dot{q}_1, \dots, \dot{q}_l).$$

Infatti, dal momento che $\Phi_j(s, \tilde{\mathbf{q}}) = q_j$, se $j = 2, \dots, l$, e $\Phi_1(s, \tilde{\mathbf{q}}) = q_1 + s$, abbiamo

$$\sum_{k=1}^l \frac{\partial \Phi_j(s, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k = \sum_{k=1}^l \delta_{jk} \dot{q}_k = \dot{q}_j, \quad j = 2, \dots, l,$$

e

$$\sum_{k=1}^l \frac{\partial \Phi_1(s, \tilde{\mathbf{q}})}{\partial q_k} \dot{q}_k = \sum_{k=1}^l \frac{\partial (q_1 + s)}{\partial q_k} \dot{q}_k = \dot{q}_1.$$

Definizione 5.8.3 Un gruppo a un parametro Φ si dice una *simmetria per la funzione Lagrangiana* \mathcal{L} se per ogni $s \in I$ e per ogni $(\tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}}) \in \mathcal{S}$ vale

$$\mathcal{L}(\tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}}) = \mathcal{L}(T\Phi(s, \tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}})) , \quad (5.98)$$

cioè, per ogni $s \in I$, e $(\tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}}) \in \mathcal{S}$, abbiamo

$$\begin{aligned} \mathcal{L}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l) &= \mathcal{L}\left(\Phi_1(s, \tilde{\mathbf{q}}), \dots, \Phi_l(s, \tilde{\mathbf{q}}), \sum_{k=1}^l \frac{\partial \Phi_1}{\partial q_k} \dot{q}_k, \dots, \right. \\ &\quad \left. \dots, \sum_{k=1}^l \frac{\partial \Phi_l}{\partial q_k} \dot{q}_k\right). \end{aligned} \quad (5.99)$$

Possiamo ora enunciare il teorema di Noether²³

Teorema 5.8.1 Sia Φ una simmetria per la funzione Lagrangiana \mathcal{L} , allora la funzione $\mathcal{I} : \mathcal{S} \rightarrow \mathbb{R}$ definita da

$$\mathcal{I}(\tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}}) = \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{d\Phi_k}{ds} \Big|_{s=0}, \quad (5.100)$$

è una costante del moto, ovvero

$$\frac{d\mathcal{I}}{dt} = 0.$$

Dim. La dimostrazione si basa sulla possibilità di scambiare le derivate rispetto a s e a t . Queste sono due variabili indipendenti tra loro (t è il tempo, mentre s è il parametro di una trasformazione "geometrica" indipendente dal tempo). Quindi, considerando

$$\tilde{\mathbf{q}}(t) = (q_1(t), \dots, q_l(t)),$$

²³Amalie Emmy Noether (Erlangen, 1882 – Bryn Mawr, Pennsylvania 1935), matematica tedesca.

per ogni componente $\Phi_j, j = 1, \dots, l$, della l -upla $\Phi(s, \tilde{\mathbf{q}}(t))$, avremo

$$\frac{d}{dt} \left(\frac{d\Phi_j(s, \tilde{\mathbf{q}}(t))}{ds} \right) = \frac{d}{ds} \left(\frac{d\Phi_j}{dt} \right) = \frac{d}{ds} \left(\sum_{k=1}^l \frac{\partial \Phi_j}{\partial q_k} \dot{q}_k \right). \quad (5.101)$$

Deriviamo la funzione \mathcal{I} rispetto al tempo

$$\frac{d}{dt} \mathcal{I}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l) = \frac{d}{dt} \left[\sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{d\Phi_k}{ds} \Big|_{s=0} \right], \quad (5.102)$$

e sviluppiamo la derivata del prodotto dentro la somma

$$\sum_{k=1}^l \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) \frac{d\Phi_k}{ds} \Big|_{s=0} + \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{d}{dt} \left(\frac{d\Phi_k}{ds} \Big|_{s=0} \right) \quad (5.103)$$

Nel primo addendo possiamo usare le equazioni di Lagrange per sostituire $\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right)$ con $\frac{\partial \mathcal{L}}{\partial q_k}$, e nel secondo addendo possiamo scambiare le derivate rispetto a s e t come nella (5.101),

$$\begin{aligned} & \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial q_k} \frac{d\Phi_k}{ds} \Big|_{s=0} + \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{d}{ds} \left(\frac{d\Phi_k}{dt} \right) \Big|_{s=0} \\ &= \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial q_k} \frac{d\Phi_k}{ds} \Big|_{s=0} + \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{d}{ds} \left(\sum_{h=1}^l \frac{\partial \Phi_k}{\partial q_h} \dot{q}_h \right) \Big|_{s=0}. \end{aligned}$$

Quest'ultima è la derivata di $\mathcal{L}(T\Phi(s, (q_1, \dots, q_l; \dot{q}_1, \dots, \dot{q}_l)))$ rispetto a s ed è quindi nulla per l'invarianza della Lagrangiana sotto l'azione del gruppo a un parametro. Infatti, sfruttando la (5.99), per ogni $s \in I$,

$$\begin{aligned} 0 &= \frac{d}{ds} \mathcal{L} \left(\Phi_1(s, \tilde{\mathbf{q}}), \dots, \Phi_l(s, \tilde{\mathbf{q}}), \sum_{k=1}^l \frac{\partial \Phi_1}{\partial q_k} \dot{q}_k, \dots, \sum_{k=1}^l \frac{\partial \Phi_1}{\partial q_k} \dot{q}_k \right) \\ &= \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial q_k} \frac{d\Phi_k}{ds} + \sum_{k=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{d}{ds} \left(\sum_{h=1}^l \frac{\partial \Phi_k}{\partial q_h} \dot{q}_h \right). \end{aligned}$$

E quindi, calcolando la derivata per $s = 0$, otteniamo la tesi.

□

5.9 Equilibrio

Nel seguito lavoreremo sotto le seguenti ipotesi:

Hp. 1. Il sistema di N punti materiali è sottoposto a m vincoli olonomi lisci, ed è soggetto soltanto a forze conservative ed alle forze vincolari, cioè $\tilde{\mathbf{S}} \equiv 0$.

Hp. 2. La Lagrangiana non dipende esplicitamente dal tempo. Dalla (4.33), abbiamo dunque

$$\mathcal{L} = \underbrace{\frac{1}{2} \sum_{h,k=1}^l a_{hk}(q_1, \dots, q_l) \dot{q}_h \dot{q}_k}_{T_2} + \underbrace{\sum_{h=1}^l b_h(q_1, \dots, q_l) \dot{q}_h}_{T_1} + T_0(q_1, \dots, q_l) - V(q_1, \dots, q_l). \quad (5.104)$$

Notiamo che tale espressione potrebbe sembrare in contrasto con l'ipotesi 2: infatti, se nella funzione \mathcal{L} compaiono T_1 e T_0 , allora abbiamo a che fare con vincoli mobili e quindi con Lagrangiane dipendenti dal tempo. Così non è poiché esistono casi di vincoli mobili per cui la \mathcal{L} non dipende esplicitamente dal tempo. Questo è il caso dell'esempio 5.7.1 dove, se la circonferenza ruota con velocità costante ω , abbiamo una Lagrangiana che non dipende esplicitamente dal tempo.

Poiché che la Lagrangiana in (5.104) non dipende esplicitamente dal tempo e $\tilde{\mathcal{S}} \equiv 0$, la funzione di Hamilton (5.78) è un integrale primo del moto. Infatti dalla (5.82) deduciamo

$$\frac{d\mathcal{H}}{dt} = 0. \quad (5.105)$$

La forma di \mathcal{H} è data dalla (5.84), cioè

$$\mathcal{H} = \frac{1}{2} \sum_{h,k=1}^l a_{hk}(q_1, \dots, q_l) \dot{q}_h \dot{q}_k - T_0(q_1, \dots, q_l) + V(q_1, \dots, q_l) = T_2 - V_{eff}. \quad (5.106)$$

Ora, cerchiamo (se esistono) soluzioni *costanti* delle equazioni di moto (5.36), ovvero funzioni del tipo

$$q_1(t) \equiv q_{1,e}, \dots, q_l(t) \equiv q_{l,e}, \quad \text{per ogni } t. \quad (5.107)$$

In analogia con la definizione 2.8.1, diamo la seguente:

Definizione 5.9.1 La configurazione $\tilde{\mathbf{q}}_e = (q_{1,e}, \dots, q_{l,e})$ si dice una **configurazione di equilibrio** se una volta posto il sistema in tale configurazione, con $(\dot{q}_1, \dots, \dot{q}_l) \equiv (0, \dots, 0)$, il sistema ivi rimane indefinitamente. In altri termini, $\tilde{\mathbf{q}}_e$ è configurazione di equilibrio se $\tilde{\mathbf{q}}(t) \equiv \tilde{\mathbf{q}}_e$, è soluzione delle equazioni di moto (5.36) con le condizioni iniziali $\tilde{\mathbf{q}}|_{t=0} = \tilde{\mathbf{q}}_e$, e $\dot{\tilde{\mathbf{q}}}|_{t=0} = 0$.

Proposizione 5.9.1 Se le ipotesi **Hp. 1** e **Hp. 2** sono soddisfatte, le configurazioni di equilibrio sono **tutte e sole le soluzioni di**

$$\frac{\partial V_{eff}}{\partial q_k} = \frac{\partial (V - T_0)}{\partial q_k} = 0, \quad k = 1, \dots, l. \quad (5.108)$$

Dim. Ricordando la forma compatta (4.40) della Lagrangiana, abbiamo

$$\mathcal{L} = \frac{1}{2} \tilde{\mathbf{q}}^T \mathbb{A} \tilde{\mathbf{q}} + \tilde{\mathbf{b}} \bullet \tilde{\mathbf{q}} - V_{eff}(\tilde{\mathbf{q}}),$$

dove né \mathbb{A} né $\tilde{\mathbf{b}}$ dipendono esplicitamente dal tempo. Scriviamo adesso le equazioni (5.36)

$$\mathbb{A} \tilde{\mathbf{q}} + \frac{d\mathbb{A}}{dt} \tilde{\mathbf{q}} + \frac{d\tilde{\mathbf{b}}}{dt} - \begin{bmatrix} \frac{1}{2} \tilde{\mathbf{q}}^T \frac{\partial \mathbb{A}}{\partial q_1} \tilde{\mathbf{q}} \\ \vdots \\ \frac{1}{2} \tilde{\mathbf{q}}^T \frac{\partial \mathbb{A}}{\partial q_l} \tilde{\mathbf{q}} \end{bmatrix} - \begin{bmatrix} \frac{\partial \tilde{\mathbf{b}}}{\partial q_1} \bullet \tilde{\mathbf{q}} \\ \vdots \\ \frac{\partial \tilde{\mathbf{b}}}{\partial q_l} \bullet \tilde{\mathbf{q}} \end{bmatrix} + \begin{bmatrix} \frac{\partial V_{eff}}{\partial q_1} \\ \vdots \\ \frac{\partial V_{eff}}{\partial q_l} \end{bmatrix} = 0.$$

In particolare,

$$\frac{d\tilde{\mathbf{b}}}{dt} = \begin{bmatrix} \frac{db_1}{dt} \\ \vdots \\ \frac{db_l}{dt} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^l \frac{\partial b_1}{\partial q_k} \dot{q}_k \\ \vdots \\ \sum_{k=1}^l \frac{\partial b_l}{\partial q_k} \dot{q}_k \end{bmatrix} = \sum_{k=1}^l \frac{\partial \tilde{\mathbf{b}}}{\partial q_k} \dot{q}_k,$$

e quindi

$$\mathbb{A} \tilde{\mathbf{q}} + \frac{d\mathbb{A}}{dt} \tilde{\mathbf{q}} + \sum_{k=1}^l \frac{\partial \tilde{\mathbf{b}}}{\partial q_k} \dot{q}_k - \begin{bmatrix} \frac{1}{2} \tilde{\mathbf{q}}^T \frac{\partial \mathbb{A}}{\partial q_1} \tilde{\mathbf{q}} \\ \vdots \\ \frac{1}{2} \tilde{\mathbf{q}}^T \frac{\partial \mathbb{A}}{\partial q_l} \tilde{\mathbf{q}} \end{bmatrix} - \begin{bmatrix} \frac{\partial \tilde{\mathbf{b}}}{\partial q_1} \bullet \tilde{\mathbf{q}} \\ \vdots \\ \frac{\partial \tilde{\mathbf{b}}}{\partial q_l} \bullet \tilde{\mathbf{q}} \end{bmatrix} + \begin{bmatrix} \frac{\partial V_{eff}}{\partial q_1} \\ \vdots \\ \frac{\partial V_{eff}}{\partial q_l} \end{bmatrix} = 0. \tag{5.109}$$

Ora, se $\tilde{\mathbf{q}} = 0$, il sistema (5.109) si riduce a

$$\begin{bmatrix} \frac{\partial V_{eff}}{\partial q_1} \\ \vdots \\ \frac{\partial V_{eff}}{\partial q_l} \end{bmatrix} = 0.$$

Le configurazioni di equilibrio $\tilde{\mathbf{q}}_e$ sono quindi tutte e sole le soluzioni di (5.108).

□

Ricordiamo poi che il potenziale efficace si riduce al potenziale delle forze direttamente applicate nel caso dei vincoli fissi. In questo caso le configurazioni di equilibrio corrispondono ai punti critici dell'energia potenziale V .

5.9.1 Stabilità

In quello che segue continuiamo a considerare le ipotesi **Hp. 1** e **Hp. 2**, cioè $\tilde{S} = 0$, e Lagrangiane indipendenti dal tempo. Ricordando la sezione 5.4, sappiamo che il sistema (5.109) può essere scritto nella seguente forma compatta

$$\mathbb{A}\tilde{\ddot{\mathbf{q}}} = \tilde{\mathbf{G}}(\tilde{\mathbf{q}}, \tilde{\dot{\mathbf{q}}}), \quad (5.110)$$

dove l'espressione di $\tilde{\mathbf{G}}$ si può ricavare dalla (5.109). Ora, la matrice \mathbb{A} è non singolare (v. teorema 4.1.3), e quindi possiamo ridurre il sistema (5.110), ad un sistema del primo ordine in forma normale. Infatti introducendo

$$\tilde{\boldsymbol{\eta}} = \tilde{\mathbf{q}}, \quad \Leftrightarrow \quad \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_l \end{pmatrix} = \begin{pmatrix} \dot{q}_1 \\ \vdots \\ \dot{q}_l \end{pmatrix},$$

abbiamo

$$\begin{cases} \tilde{\dot{\boldsymbol{\eta}}} = \tilde{\boldsymbol{\eta}}, \\ \tilde{\boldsymbol{\eta}} = \mathbb{A}^{-1}\tilde{\mathbf{G}}(\tilde{\mathbf{q}}, \tilde{\boldsymbol{\eta}}). \end{cases} \quad (5.111)$$

Di conseguenza, ricordando il teorema 2.8.1, abbiamo la seguente

Definizione 5.9.2 *Una configurazione di equilibrio, o punto di equilibrio, $\tilde{\mathbf{q}}_e = (q_{1,e}, \dots, q_{l,e})$ si dirà **stabile** se $(\tilde{\mathbf{q}}_e, \tilde{\boldsymbol{\eta}}_e) = (q_{e1}, \dots, q_{el}, 0, \dots, 0)$ è una posizione di equilibrio stabile nel senso della definizione 2.8.2, per il sistema di equazioni del primo ordine (5.111). Un punto di equilibrio non stabile si dirà **instabile**.*

Come nella sezione 2.8.3, vale il

Teorema 5.9.1 (Principio di Dirichlet) *Se le ipotesi **Hp. 1** e **Hp. 2** sono soddisfatte, e $\tilde{\mathbf{q}}_e = (q_{1,e}, \dots, q_{l,e})$ è un minimo isolato dell'energia potenziale efficace $V_{eff} = -T_0 + V$, allora $\tilde{\mathbf{q}}_e$ è una configurazione di equilibrio stabile.*

Dim. La dimostrazione di questo principio è relativamente semplice: si riduce a far vedere che la funzione Hamiltoniana data da (5.106), vista come funzione di $\tilde{\boldsymbol{\eta}}$ e di $\tilde{\mathbf{q}}$, meno $V_{eff}(\tilde{\mathbf{q}}_e)$, cioè

$$\Lambda(\tilde{\boldsymbol{\eta}}, \tilde{\mathbf{q}}) = \mathcal{H}(\tilde{\boldsymbol{\eta}}, \tilde{\mathbf{q}}) - V_{eff}(\tilde{\mathbf{q}}_e) = \frac{1}{2}\tilde{\boldsymbol{\eta}}^T \mathbb{A}\tilde{\boldsymbol{\eta}} + V_{eff}(\tilde{\mathbf{q}}) - V_{eff}(\tilde{\mathbf{q}}_e),$$

è una funzione di Lyapunov per il sistema (5.111). Si applica dunque il teorema 2.8.1 da cui discende che $(\tilde{\mathbf{q}}_e, \tilde{\boldsymbol{\eta}}_e)$ è punto di equilibrio stabile. Mostriamo dunque che $\Lambda(\tilde{\boldsymbol{\eta}}, \tilde{\mathbf{q}})$ è funzione di Lyapunov.

La condizione che la derivata di $\Lambda(\tilde{\boldsymbol{\eta}}, \tilde{\mathbf{q}})$ lungo le soluzioni sia minore o uguale a zero è conseguenza della conservazione dell'energia (5.105). Resta da far vedere che la funzione Λ ha un minimo isolato in $(q_{e,1}, \dots, q_{e,l}, 0, \dots, 0)$. Anche questo segue facilmente dalla forma della Λ e dall'ipotesi di minimo isolato per V_{eff} . Infatti il primo addendo nella definizione di Λ è una forma quadratica definita positiva nelle η_k (cioè nelle \dot{q}_k) e quindi è nullo soltanto se $\eta_1 = \dots = \eta_l = 0$, mentre $V_{eff}(\tilde{\mathbf{q}}) > V_{eff}(\tilde{\mathbf{q}}_e)$.

Quindi, se prendiamo $(\tilde{\mathbf{q}}, \tilde{\boldsymbol{\eta}})$ in un intorno della configurazione di equilibrio $(\tilde{\mathbf{q}}_e, \tilde{\boldsymbol{\eta}}_e)$, ma diversi da $(\tilde{\mathbf{q}}_e, \tilde{\boldsymbol{\eta}}_e)$, si ha

$$\Lambda(\tilde{\boldsymbol{\eta}}, \tilde{\mathbf{q}}) = \frac{1}{2} \tilde{\boldsymbol{\eta}}^T \mathbb{A} \tilde{\boldsymbol{\eta}} + V_{eff}(\tilde{\mathbf{q}}) - V_{eff}(\tilde{\mathbf{q}}_e) > 0,$$

La funzione Λ ha un minimo isolato in $(q_{1, e}, \dots, q_{l, e}, 0, \dots, 0)$.

□

Esempio 5.9.1 Continuando con l'esempio 5.7.1, vogliamo determinare le configurazioni di equilibrio stabile nel caso in cui la circonferenza ruoti uniformemente, cioè $\frac{d\varphi}{dt} = \omega = \text{costante}$. Abbiamo già osservato che, in questo caso, \mathcal{L} non dipende esplicitamente dal tempo. Sono quindi verificate le ipotesi **Hp.1** e **Hp.2**, per cui le configurazioni di equilibrio sono i massimi e i minimi di

$$V_{eff}(\theta) = V - T_0 = -\frac{\omega^2}{2} \sin^2 \theta - \frac{g}{R} \cos \theta.$$

I punti di equilibrio sono dunque le soluzioni di

$$\frac{dV_{eff}(\theta)}{d\theta} = \omega^2 \sin \theta \left(\frac{g}{\omega^2 R} - \cos \theta \right) = 0.$$

Si distinguono due casi:

- $\frac{g}{\omega^2 R} \geq 1$. I punti di equilibrio sono $\theta = 0, \pi$ (o $\theta = -\pi$). Dall'analisi del segno di V'_{eff} , si deduce che $\theta = 0$, è un minimo isolato e quindi è punto di equilibrio stabile. E' facile poi vedere che $\theta = \pi$ è configurazione di equilibrio instabile.
- $\frac{g}{\omega^2 R} < 1$. Gli zeri di V'_{eff} sono $\theta = 0, \pi, \pm \bar{\theta}$, dove $\bar{\theta} = \arccos\left(\frac{g}{\omega^2 R}\right)$. Le configurazioni di equilibrio stabile corrispondono a $\pm \bar{\theta}$. Applicando poi la definizione 5.9.2, si prova che $\theta = 0, \pi$, sono punti di equilibrio instabile.

5.10 Piccole Oscillazioni

Supponiamo di avere un sistema meccanico con l gradi di libertà, con funzione Lagrangiana della forma²⁴

$$\mathcal{L}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) = \frac{1}{2} \sum_{i,j=1}^l a_{ij}(\tilde{\mathbf{q}}) \dot{q}_i \dot{q}_j - V(\tilde{\mathbf{q}}). \quad (5.112)$$

Supponiamo che la configurazione $\tilde{\mathbf{q}}_e = (q_{1, e}, \dots, q_{l, e})$ sia tale che si abbia

$$\frac{\partial V}{\partial q_k}(\tilde{\mathbf{q}}_e) = 0, \quad k = 1, \dots, l, \quad (5.113)$$

²⁴Se la Lagrangiana ha la forma $\mathcal{L} = T_2 + T_0 - V$, quello che segue vale sostituendo $V_{eff} = -T_0 + V$ al posto dell'energia potenziale V .

e che la matrice hessiana

$$\mathbb{B}(\tilde{\mathbf{q}}_e) = \frac{\partial^2 V(\tilde{\mathbf{q}}_e)}{\partial q_k \partial q_h}, \quad k, h = 1, \dots, l, \quad (5.114)$$

sia definita positiva. Le condizioni (5.113) e (5.114), e il criterio di Dirichlet, ci garantiscono che la configurazione $\tilde{\mathbf{q}}_e$ è una **configurazione di equilibrio stabile**, ovvero che il moto resta confinato in un intorno della configurazione di equilibrio $\tilde{\mathbf{q}}_e$ (e della velocità lagrangiana $\tilde{\mathbf{q}}$ nulla). In altri termini, durante il moto, le norme $\|\tilde{\mathbf{q}}(t) - \tilde{\mathbf{q}}_e\|$ e $\|\dot{\tilde{\mathbf{q}}}(t)\|$ restano “piccole”. Vediamo dunque come sfruttare questo fatto per semplificare le equazioni di Lagrange in vicinanza della configurazione di equilibrio stabile. Dapprima considereremo sistemi con un solo grado di libertà e poi sistemi con $l > 1$, gradi di libertà.

5.10.1 Sistemi con un solo grado di libertà

In questo caso \mathcal{L} è data da

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}a(q)\dot{q}^2 - V(q).$$

La configurazione di equilibrio stabile è $q = q_e$, per cui $V'(q_e) = 0$, e $V''(q_e) > 0$. Consideriamo quindi (q, \dot{q}) “vicini” al punto di equilibrio $(q_e, 0)$, cioè $(q, \dot{q}) \in B_\varepsilon(q_e, 0)$, essendo $B_\varepsilon(q_e, 0)$ l’intorno di $(q_e, 0)$ di raggio ε , come mostrato in figura 5.6. Si introduce quindi la nuova variabile ξ

$$\begin{aligned} q &= q_e + \varepsilon\xi, & \Leftrightarrow & \quad \xi = \frac{q - q_e}{\varepsilon}, \\ \dot{q} &= \varepsilon\dot{\xi}. \end{aligned}$$

Evidentemente $0 \leq |\xi| \leq 1$, mentre ε sarà “piccolo”, cioè $\varepsilon \ll 1$. Scriveremo dunque $\varepsilon = \mathcal{O}(q - q_e)$, mentre $\xi = \mathcal{O}(1)$, che significano: ε è dell’ordine di $(q - q_e)$, e quindi “molto piccolo”, mentre l’ordine di grandezza di ξ è 1.

Adesso sviluppiamo sia $V(q)$ che $a(q)$ in vicinanza di q_e ,

$$\begin{aligned} V(q) &= V(q_e) + \underbrace{V'(q_e)}_{=0}(q - q_e) + \frac{1}{2}\underbrace{V''(q_e)}_{>0}(q - q_e)^2 + \dots \\ &= V(q_e) + \frac{1}{2}V''(q_e)\varepsilon^2\xi^2 + o(\varepsilon^2). \\ a(q) &= a(q_e) + a'(q_e)(q - q_e) + \frac{1}{2}a''(q_e)(q - q_e)^2 + \dots \\ &= a(q_e) + a'(q_e)\varepsilon\xi + \frac{1}{2}a''(q_e)\varepsilon^2\xi^2 + o(\varepsilon^2). \end{aligned}$$

Sostituiamo tali sviluppi nella Lagrangiana

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \left[a(q_e) + a'(q_e)\varepsilon\xi + \frac{1}{2}a''(q_e)\varepsilon^2\xi^2 + o(\varepsilon^2) \right] \varepsilon^2\dot{\xi}^2 \\ &\quad + V(q_e) + \frac{1}{2}V''(q_e)\varepsilon^2\xi^2 + o(\varepsilon^2). \end{aligned}$$

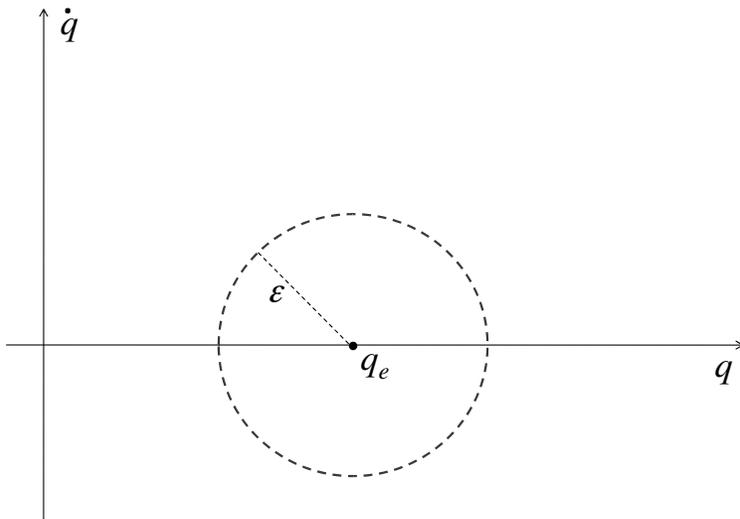


Figura 5.6: L'intorno $B_\varepsilon(q_e, 0)$.

Se trascuriamo i termini $o(\varepsilon^2)$, possiamo approssimare la funzione di Lagrange con

$$\mathcal{L}_{app} = \frac{1}{2}a(q_e)\varepsilon^2\xi^2 - \frac{1}{2}V''(q_e)\varepsilon^2\xi^2 - V(q_e),$$

che definiamo **Lagrangiana approssimata**, o **Lagrangiana delle piccole oscillazioni**. Scrivendo, per tale Lagrangiana, l'equazione di moto (5.36) otteniamo

$$a(q_e)\ddot{\xi} + V''(q_e)\xi = 0, \Rightarrow \ddot{\xi} = -\frac{V''(q_e)}{a(q_e)}\xi,$$

che è l'equazione dell'oscillatore armonico $m\ddot{\xi} - k\xi = 0$, la cui frequenza è

$$\omega^2 = \frac{V''(q_e)}{a(q_e)}.$$

5.10.2 Sistemi con l gradi di libertà

Si procede come nella sezione precedente introducendo lo sviluppo di Taylor di V in un intorno del punto del punto di equilibrio limitandoci soltanto al secondo ordine

$$V(\mathbf{q}) \approx V(\tilde{\mathbf{q}}_e) + \sum_{k=1}^l \underbrace{\frac{\partial V(\tilde{\mathbf{q}}_e)}{\partial q_k}}_{=0} (q_k - q_{k,e}) + \frac{1}{2} \underbrace{(\tilde{\mathbf{q}} - \tilde{\mathbf{q}}_e)^T \mathbb{B}(\tilde{\mathbf{q}}_e) (\tilde{\mathbf{q}} - \tilde{\mathbf{q}}_e)}_{\sum_{k,h=1}^l b_{kh}(\tilde{\mathbf{q}}_e)(q_h - q_{h,e})(q_k - q_{k,e})},$$

dove $\mathbb{B}(\tilde{\mathbf{q}}_e)$ è la matrice Hessiana calcolata in $\tilde{\mathbf{q}}_e$, data dalla (5.114). Per quel che riguarda l'energia cinetica, seguendo lo stesso approccio illustrato nella sezione 5.10.1, ci limitiamo all'ordine "zero", cioè

$$T \approx \frac{1}{2} \underbrace{\sum_{h,k=1}^l a_{hk}(\tilde{\mathbf{q}}_e) \dot{q}_h \dot{q}_k}_{\tilde{\mathbf{q}}^T \mathbb{A}(\tilde{\mathbf{q}}_e) \tilde{\mathbf{q}}} = \frac{1}{2} \sum_{h,k=1}^l a_{hk}(\tilde{\mathbf{q}}_e) \frac{d(q_k - q_{k,e})}{dt} \frac{d(q_h - q_{h,e})}{dt}.$$

Approssimiamo dunque la Lagrangiana (5.112) con la seguente

$$\begin{aligned} \mathcal{L}_{app}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) &= \frac{1}{2} \sum_{h,k=1}^l a_{hk}(\tilde{\mathbf{q}}_e) \frac{d(q_k - q_{k,e})}{dt} \frac{d(q_h - q_{h,e})}{dt} \\ &\quad - \frac{1}{2} \sum_{k,h=1}^l b_{kh}(q_h - q_{h,e})(q_k - q_{k,e}). \end{aligned} \quad (5.115)$$

In particolare, introducendo

$$\tilde{\mathbf{u}}(t) = \tilde{\mathbf{q}}(t) - \tilde{\mathbf{q}}_e = \begin{pmatrix} q_1(t) - q_{1,e} \\ \vdots \\ q_l(t) - q_{l,e} \end{pmatrix}, \Rightarrow \dot{\tilde{\mathbf{u}}} = \frac{d(\tilde{\mathbf{q}} - \tilde{\mathbf{q}}_e)}{dt} = \dot{\tilde{\mathbf{q}}}, \quad (5.116)$$

possiamo riscrivere la (5.115) in forma matriciale compatta

$$\mathcal{L}_{app}(\tilde{\mathbf{u}}, \dot{\tilde{\mathbf{u}}}) = \frac{1}{2} \dot{\tilde{\mathbf{u}}}^T \mathbb{A}(\tilde{\mathbf{q}}_e) \dot{\tilde{\mathbf{u}}} - \frac{1}{2} \tilde{\mathbf{u}}^T \mathbb{B}(\tilde{\mathbf{q}}_e) \tilde{\mathbf{u}}. \quad (5.117)$$

Le equazioni di Lagrange (5.36) sono

$$\mathbb{A} \ddot{\tilde{\mathbf{u}}} + \mathbb{B} \tilde{\mathbf{u}} = 0, \quad (5.118)$$

dove \mathbb{A} e \mathbb{B} sono matrici $l \times l$, simmetriche, definite positive, le cui componenti sono costanti.

Chiameremo **Lagrangiana delle piccole oscillazioni** o **Lagrangiana approssimata** la funzione definita in (5.117) (o (5.115)), mentre le equazioni (5.118) sono usualmente dette **equazioni delle piccole oscillazioni**.

Soluzione delle equazioni delle piccole oscillazioni

Come già visto nella sezione 2.10, dove abbiamo trattato i sistemi lineari bidimensionali di equazioni differenziali a coefficienti costanti, cerchiamo di semplificare il sistema (5.118) applicando un "trucco vecchio quanto la matematica": il cambiamento di variabili. Cerchiamo cioè nuove variabili indipendenti $\tilde{\mathbf{w}}(t)$, che siano una combinazione lineare delle variabili $\tilde{\mathbf{u}}(t)$, ovvero che siano legate alle $\tilde{\mathbf{u}}(t)$ da una trasformazione lineare, ovviamente **invertibile**, del tipo

$$\tilde{\mathbf{u}}(t) = \mathbb{V} \tilde{\mathbf{w}}(t), \quad \Leftrightarrow \quad \tilde{\mathbf{w}}(t) = \mathbb{V}^{-1} \tilde{\mathbf{u}}(t), \quad (5.119)$$

dove \mathbb{V} è una matrice $l \times l$, invertibile, le cui componenti, denotate con v_{ij} , $i, j = 1, \dots, l$, sono numeri reali. Sostituendo nella (5.118) si ha

$$\mathbb{A}\mathbb{V}\tilde{\mathbf{w}} + \mathbb{B}\mathbb{V}\tilde{\mathbf{w}} = 0,$$

che, moltiplicata a sinistra per \mathbb{V}^T , diventa

$$(\mathbb{V}^T\mathbb{A}\mathbb{V})\tilde{\mathbf{w}} + (\mathbb{V}^T\mathbb{B}\mathbb{V})\tilde{\mathbf{w}} = 0. \quad (5.120)$$

Se adesso riusciamo a selezionare \mathbb{V} in modo tale che

$$\mathbb{V}^T\mathbb{A}\mathbb{V} = \mathbb{I}, \quad (5.121)$$

e

$$\mathbb{V}^T\mathbb{B}\mathbb{V} = \mathbb{M} = \begin{pmatrix} \mu_1 & 0 & 0 & \cdots & 0 \\ 0 & \mu_2 & 0 & \cdots & 0 \\ 0 & 0 & \mu_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mu_l \end{pmatrix}, \quad (5.122)$$

con μ_k **positivi**, cioè $\mu_k > 0$, $\forall k = 1, \dots, l$, il sistema (5.120) diventa

$$\tilde{\mathbf{w}} + \mathbb{M}\tilde{\mathbf{w}} = 0,$$

che, scritto per esteso, assume la forma

$$\begin{cases} \ddot{w}_1 + \mu_1 w_1 = 0, \\ \ddot{w}_2 + \mu_2 w_2 = 0, \\ \vdots \\ \ddot{w}_l + \mu_l w_l = 0. \end{cases}$$

Abbiamo quindi ottenuto un sistema di l oscillatori armonici disaccoppiati, le cui soluzioni sono²⁵

$$w_k(t) = a_k \cos(\sqrt{\mu_k}t + \phi_k), \quad k = 1, 2, \dots, l, \quad (5.123)$$

dove a_k , e ϕ_k , sono costanti arbitrarie (il cui valore può essere determinato quando sono note le condizioni iniziali).

Quindi tornando alla (5.119), ovvero alle “vecchie variabili” $\tilde{\mathbf{u}}$, che poi, ricordando la (5.116), altro non sono che $(\tilde{\mathbf{q}}(t) - \tilde{\mathbf{q}}_e)$, abbiamo

$$\begin{aligned} q_i(t) - q_{ie} &= \sum_{k=1}^l v_{ik} w_k(t) \\ &= \sum_{j=1}^l v_{ij} a_k \cos(\sqrt{\mu_k}t + \phi_k), \quad i = 1, \dots, l, \end{aligned}$$

²⁵Una scrittura equivalente della (5.123) è

$$w_k(t) = a_k \cos(\sqrt{\mu_k}t) + b_k \sin(\sqrt{\mu_k}t),$$

con a_k, b_k , costanti.

cioè

$$\tilde{\mathbf{q}}(t) - \tilde{\mathbf{q}}_e = \mathbb{V} \begin{pmatrix} a_1 \cos(\sqrt{\mu_1}t + \phi_1) \\ \vdots \\ a_l \cos(\sqrt{\mu_l}t + \phi_l) \end{pmatrix}.$$

Le variabili $\tilde{\mathbf{w}}(t)$ sono detti **modi normali**, mentre

$$\omega_k = \sqrt{\mu_k}, \quad k = 1, \dots, l, \quad (5.124)$$

sono dette **frequenze**²⁶ **proprie delle piccole oscillazioni**.

Mostriamo adesso che esiste ed è unica la matrice \mathbb{V} che, realizzando le (5.121) e (5.122), permette di passare dalle variabili $(\tilde{\mathbf{q}}(t) - \tilde{\mathbf{q}}_e)$, ai modi normali $\tilde{\mathbf{w}}(t)$ dati dalla (5.123). Poniamoci dunque in \mathbb{R}^l , dove, seguendo la notazione fin qui utilizzata, con $\tilde{\mathbf{v}}$ intendiamo i vettori (cioè le l -uple) colonna

$$\tilde{\mathbf{v}} = \begin{pmatrix} v_1 \\ \vdots \\ v_l \end{pmatrix},$$

e con $\tilde{\mathbf{v}}^T$ i vettori (l -uple) riga, $\tilde{\mathbf{v}}^T = (v_1 \ \cdots \ v_l)$. Dimosteremo che è vero il seguente teorema per le matrici \mathbb{A} e \mathbb{B}

Teorema 5.10.1 (Teorema spettrale) *Sia \mathbb{A} una matrice simmetrica definita positiva e \mathbb{B} una matrice simmetrica. Allora esistono l vettori linearmente indipendenti $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_l$, e l numeri reali μ_1, \dots, μ_l (non necessariamente diversi tra loro) tali che*

$$\mu_k \mathbb{A} \tilde{\mathbf{v}}_k = \mathbb{B} \tilde{\mathbf{v}}_k, \quad k = 1, \dots, l, \quad (5.125)$$

e²⁷

$$\tilde{\mathbf{v}}_h \bullet \mathbb{A} \tilde{\mathbf{v}}_k = \tilde{\mathbf{v}}_h^T \mathbb{A} \tilde{\mathbf{v}}_k = \delta_{hk}, \quad \forall h, k = 1, \dots, l. \quad (5.126)$$

Se anche la matrice \mathbb{B} è definita positiva, allora i numeri μ_k sono tutti positivi.

I numeri μ_k e i vettori $\tilde{\mathbf{v}}_k$ sono detti rispettivamente gli **autovalori** e gli **autovettori di \mathbb{B} relativi ad \mathbb{A}** . Il teorema è una generalizzazione del teorema spettrale per le matrici simmetriche (e infatti rientra in questa versione se $\mathbb{A} = \mathbb{I}$). Ciò che si perde rispetto al caso in cui $\mathbb{A} = \mathbb{I}$, è il fatto che la base di autovettori non è più, in generale, ortonormale. Infatti non vale $\tilde{\mathbf{v}}_h \bullet \tilde{\mathbf{v}}_k = \delta_{hk}$, bensì la (5.126).

In accordo con il teorema spettrale possiamo facilmente provare che esiste ma matrice \mathbb{V} che realizza le (5.121) e (5.122). Abbiamo infatti

Teorema 5.10.2 *Se $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_l$, sono gli autovettori di \mathbb{B} relativi ad \mathbb{A} , allora, ponendo*

$$\mathbb{V} = (\tilde{\mathbf{v}}_1 \ \tilde{\mathbf{v}}_2 \ \cdots \ \tilde{\mathbf{v}}_l), \quad (5.127)$$

\mathbb{V} realizza le (5.121) e (5.122).

²⁶In realtà, se volessimo esser precisi, si tratta di pulsazioni e non di frequenze. Le frequenze ν_k sono date da $\omega_k/2\pi$.

²⁷Con \bullet si intende l'usuale prodotto scalare, componente per componente, in \mathbb{R}^l , cioè

$$\tilde{\mathbf{z}} \bullet \tilde{\mathbf{p}} = \tilde{\mathbf{z}}^T \tilde{\mathbf{p}} = \tilde{\mathbf{p}}^T \tilde{\mathbf{z}}.$$

Al solito δ_{hk} denota la δ di Dirac: $\delta_{hk} = 1$ se $h = k$, $\delta_{hk} = 0$, altrimenti.

Dim. Le colonne della matrice \mathbb{V} sono i vettori colonna (cioè le l -uple colonna) $\tilde{\mathbf{v}}_k$, $k = 1, \dots, l$. La matrice \mathbb{V}^T sarà dunque formata dai vettori riga

$$\mathbb{V}^T = \begin{pmatrix} \tilde{\mathbf{v}}_1^T \\ \tilde{\mathbf{v}}_2^T \\ \vdots \\ \tilde{\mathbf{v}}_l^T \end{pmatrix}.$$

Eseguendo il prodotto righe per colonne abbiamo

$$\mathbb{V}^T \mathbb{A} \mathbb{V} = \begin{pmatrix} \tilde{\mathbf{v}}_1^T \mathbb{A} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_1^T \mathbb{A} \tilde{\mathbf{v}}_2 & \cdots & \tilde{\mathbf{v}}_1^T \mathbb{A} \tilde{\mathbf{v}}_l \\ \tilde{\mathbf{v}}_2^T \mathbb{A} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2^T \mathbb{A} \tilde{\mathbf{v}}_2 & \cdots & \tilde{\mathbf{v}}_2^T \mathbb{A} \tilde{\mathbf{v}}_l \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{v}}_l^T \mathbb{A} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_l^T \mathbb{A} \tilde{\mathbf{v}}_2 & \cdots & \tilde{\mathbf{v}}_l^T \mathbb{A} \tilde{\mathbf{v}}_l \end{pmatrix} \stackrel{(5.126)}{=} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

ed anche

$$\begin{aligned} \mathbb{V}^T \mathbb{B} \mathbb{V} &= \begin{pmatrix} \tilde{\mathbf{v}}_1^T \mathbb{B} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_1^T \mathbb{B} \tilde{\mathbf{v}}_2 & \cdots & \tilde{\mathbf{v}}_1^T \mathbb{B} \tilde{\mathbf{v}}_l \\ \tilde{\mathbf{v}}_2^T \mathbb{B} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2^T \mathbb{B} \tilde{\mathbf{v}}_2 & \cdots & \tilde{\mathbf{v}}_2^T \mathbb{B} \tilde{\mathbf{v}}_l \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{v}}_l^T \mathbb{B} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_l^T \mathbb{B} \tilde{\mathbf{v}}_2 & \cdots & \tilde{\mathbf{v}}_l^T \mathbb{B} \tilde{\mathbf{v}}_l \end{pmatrix} \stackrel{(5.125)}{=} \\ &= \begin{pmatrix} \mu_1 \tilde{\mathbf{v}}_1^T \mathbb{A} \tilde{\mathbf{v}}_1 & \mu_2 \tilde{\mathbf{v}}_1^T \mathbb{A} \tilde{\mathbf{v}}_2 & \cdots & \mu_l \tilde{\mathbf{v}}_1^T \mathbb{A} \tilde{\mathbf{v}}_l \\ \mu_1 \tilde{\mathbf{v}}_2^T \mathbb{A} \tilde{\mathbf{v}}_1 & \mu_2 \tilde{\mathbf{v}}_2^T \mathbb{A} \tilde{\mathbf{v}}_2 & \cdots & \mu_l \tilde{\mathbf{v}}_2^T \mathbb{A} \tilde{\mathbf{v}}_l \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1 \tilde{\mathbf{v}}_l^T \mathbb{A} \tilde{\mathbf{v}}_1 & \mu_2 \tilde{\mathbf{v}}_l^T \mathbb{A} \tilde{\mathbf{v}}_2 & \cdots & \mu_l \tilde{\mathbf{v}}_l^T \mathbb{A} \tilde{\mathbf{v}}_l \end{pmatrix} \\ &= \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_l \end{pmatrix}. \end{aligned}$$

Da notare che, siccome la matrice \mathbb{B} è definita positiva (è la matrice hessiana di un punto di minimo), ogni μ_k , $k = 1, \dots, l$, è **positivo**. La matrice \mathbb{V} , soddisfacendo sia la (5.121) che la (5.122), è dunque la matrice che cercavamo.

□

Il teorema spettrale ci fornisce anche il “**metodo pratico**” per determinare μ_k , $k = 1, \dots, l$, e quindi calcolare le frequenze delle piccole oscillazioni tramite la (5.124), ed i modi normale, ovvero i vettori colonna \mathbf{v}_k , $k = 1, \dots, l$. Infatti, la (5.125) può esser così riscritta

$$(\mathbb{B} - \mu_k \mathbb{A}) \tilde{\mathbf{v}}_k = 0,$$

dove sappiamo che $\tilde{\mathbf{v}}_k \neq 0$, $k = 1, \dots, l$. Di conseguenza i valori μ_k si ottengono risolvendo l’equazione

$$\det(\mathbb{B} - \mu \mathbb{A}) = 0, \tag{5.128}$$

anche detta **equazione secolare**. Una volta determinati i μ_k (che sappiamo essere tutti reali), i vettori colonna

$$\tilde{\mathbf{v}}_k = \begin{pmatrix} v_{k,1} \\ v_{k,2} \\ \vdots \\ v_{k,l} \end{pmatrix}, \quad k = 1, \dots, l,$$

si ottengono risolvendo il sistema lineare

$$(\mathbb{B} - \mu_k \mathbb{A}) \begin{pmatrix} v_{k,1} \\ \vdots \\ v_{k,l} \end{pmatrix} = 0, \quad (5.129)$$

con la condizione supplementare $\tilde{\mathbf{v}}_k \bullet \mathbb{A} \tilde{\mathbf{v}}_k = 1$, $k = 1, \dots, l$. Assemblando poi gli l vettori colonna $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_l$, si forma la matrice \mathbb{V} secondo la (5.127). I modi normali $\tilde{\mathbf{w}}$ si ottengono poi applicando la (5.119).

Dimostrazione del Teorema spettrale

Prima di passare alla dimostrazione vera e propria del teorema spettrale è opportuno premettere il seguente

Lemma 5.10.1 *Se le matrici \mathbb{A} e \mathbb{B} soddisfano le ipotesi del teorema spettrale, allora l'equazione in μ (5.128) ha soltanto radici reali. Se poi \mathbb{B} è definita positiva allora le radici di (5.128) sono tutte positive.*

Dim. Il teorema fondamentale dell'algebra ci dice che (5.128) ammette l soluzioni. Vediamo che queste non possono essere complesse. Supponiamo infatti che μ sia una soluzione complessa, e quindi $\text{Im}(\mu) \neq 0$, di (5.128). Siccome l'equazione (5.128) è a coefficienti reali anche²⁸ $\bar{\mu}$ è soluzione. Sia $\tilde{\mathbf{n}}$ la l -upla complessa, $\tilde{\mathbf{n}} = \text{Re } \tilde{\mathbf{n}} + i \text{Im } \tilde{\mathbf{n}}$, corrispondente a μ , cioè soluzione di

$$\mathbb{B} \tilde{\mathbf{n}} = \mu \mathbb{A} \tilde{\mathbf{n}}. \quad (5.130)$$

Siccome \mathbb{A} e \mathbb{B} sono reali, considerando i complessi coniugati, abbiamo che $\tilde{\bar{\mathbf{n}}} = \text{Re } \tilde{\mathbf{n}} - i \text{Im } \tilde{\mathbf{n}}$, risolve $\mathbb{B} \tilde{\bar{\mathbf{n}}} = \bar{\mu} \mathbb{A} \tilde{\bar{\mathbf{n}}}$. Consideriamo adesso

$$\begin{aligned} \tilde{\mathbf{n}} \bullet \mathbb{A} \tilde{\bar{\mathbf{n}}} &= (\text{Re } \tilde{\mathbf{n}} + i \text{Im } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Re } \tilde{\mathbf{n}} - i \text{Im } \tilde{\mathbf{n}}) = \\ &= (\text{Re } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Re } \tilde{\mathbf{n}}) + (\text{Im } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Im } \tilde{\mathbf{n}}) + \\ &\quad i [(\text{Im } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Re } \tilde{\mathbf{n}}) - (\text{Re } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Im } \tilde{\mathbf{n}})]. \end{aligned}$$

Siccome \mathbb{A} è simmetrica, $(\text{Im } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Re } \tilde{\mathbf{n}}) = \mathbb{A}^T (\text{Im } \tilde{\mathbf{n}}) \bullet (\text{Re } \tilde{\mathbf{n}}) = \mathbb{A} (\text{Im } \tilde{\mathbf{n}}) \bullet (\text{Re } \tilde{\mathbf{n}})$, e dunque il termine in parentesi quadra si annulla. Concludiamo pertanto che

$$\tilde{\mathbf{n}} \bullet \mathbb{A} \tilde{\bar{\mathbf{n}}} = \underbrace{(\text{Re } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Re } \tilde{\mathbf{n}})}_{>0 \text{ perchè } \mathbb{A} \text{ è definita positiva}} + \underbrace{(\text{Im } \tilde{\mathbf{n}}) \bullet \mathbb{A} (\text{Im } \tilde{\mathbf{n}})}_{>0 \text{ perchè } \mathbb{A} \text{ è definita positiva}} > 0. \quad (5.131)$$

²⁸Con $\bar{\mu}$ intendiamo il complesso coniugato di μ , cioè $\bar{\mu} = \text{Re}(\mu) - i \text{Im}(\mu)$.

Osserviamo adesso che se μ_1, μ_2 , con $\mu_1 \neq \mu_2$, sono due soluzioni di (5.128) e $\tilde{\mathbf{n}}_1, \tilde{\mathbf{n}}_2$, sono le due corrispondenti l -uple,

$$\mathbb{B} \tilde{\mathbf{n}}_1 = \mu_1 \mathbb{A} \tilde{\mathbf{n}}_1, \quad \mathbb{B} \tilde{\mathbf{n}}_2 = \mu_2 \mathbb{A} \tilde{\mathbf{n}}_2,$$

abbiamo

$$\tilde{\mathbf{n}}_1 \bullet \mathbb{A} \tilde{\mathbf{n}}_2 = \tilde{\mathbf{n}}_2 \bullet \mathbb{A} \tilde{\mathbf{n}}_1 = 0.$$

Infatti, sfruttando la simmetria di \mathbb{A} e \mathbb{B} ,

$$\mu_1 \tilde{\mathbf{n}}_2 \bullet \mathbb{A} \tilde{\mathbf{n}}_1 = \tilde{\mathbf{n}}_2 \bullet \mathbb{B} \tilde{\mathbf{n}}_1 \underset{\mathbb{B}=\mathbb{B}^T}{=} \mathbb{B} \tilde{\mathbf{n}}_2 \bullet \tilde{\mathbf{n}}_1 = \mu_2 \mathbb{A} \tilde{\mathbf{n}}_2 \bullet \tilde{\mathbf{n}}_1 \underset{\mathbb{A}=\mathbb{A}^T}{=} \mu_2 \tilde{\mathbf{n}}_2 \bullet \mathbb{A} \tilde{\mathbf{n}}_1,$$

che ci darebbe l'assurdo $\mu_1 = \mu_2$, se non fosse $\tilde{\mathbf{n}}_1 \bullet \mathbb{A} \tilde{\mathbf{n}}_2 = 0$. Ma allora, tornando alla (5.131), abbiamo ottenuto un assurdo: due soluzioni $\tilde{\mathbf{n}}$ e $\tilde{\bar{\mathbf{n}}}$, relative a μ e $\bar{\mu}$, ovviamente fra loro diversi, danno $\tilde{\mathbf{n}} \bullet \mathbb{A} \tilde{\bar{\mathbf{n}}} > 0$, anziché $\tilde{\mathbf{n}} \bullet \mathbb{A} \tilde{\bar{\mathbf{n}}} = 0$. Concludiamo quindi che una qualunque soluzione μ di (5.128) non può essere complessa.

Mostriamo infine che, se \mathbb{B} è definita positiva e se μ è soluzione di (5.128), abbiamo $\mu > 0$. Sia infatti se $\tilde{\mathbf{n}}$ è il vettore colonna (ovviamente non nullo) relativo a μ , cioè $\mathbb{B} \tilde{\mathbf{n}} = \mu \mathbb{A} \tilde{\mathbf{n}}$. Otteniamo

$$\underbrace{\tilde{\mathbf{n}}^T \mathbb{B} \tilde{\mathbf{n}}}_{>0} = \mu \underbrace{\tilde{\mathbf{n}}^T \mathbb{A} \tilde{\mathbf{n}}}_{>0}, \quad \implies \quad \mu > 0.$$

□

Passiamo adesso alla **dimostrazione del teorema spettrale**. Si parte dalla (5.128) e si considera una radice μ , che, in virtù del lemma precedente, sappiamo essere reale. Sia quindi $\tilde{\mathbf{v}} \neq 0$ il corrispondente vettore colonna di \mathbb{R}^l soluzione di

$$\mathbb{B} \tilde{\mathbf{v}} = \mu \mathbb{A} \tilde{\mathbf{v}}. \tag{5.132}$$

Vogliamo provare che esistono $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_l$, vettori linearmente indipendenti che soddisfano la (5.132) in corrispondenza delle soluzioni μ (tutte reali) della (5.128).

Osserviamo che la matrice \mathbb{A} , essendo simmetrica e definita positiva, ammette una base di l autovettori $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_l$, di \mathbb{A} , cioè

$$\begin{aligned} \mathbb{A} \tilde{\mathbf{a}}_i &= \rho_i \tilde{\mathbf{a}}_i, \quad \text{con } \rho_i \in \mathbb{R} \text{ e } \rho_i > 0, \quad i = 1, \dots, l, \\ \tilde{\mathbf{a}}_i \bullet \tilde{\mathbf{a}}_j &= \tilde{\mathbf{a}}_i^T \tilde{\mathbf{a}}_j = \delta_{ij}, \quad \forall i, j = 1, \dots, l. \end{aligned}$$

Costruiamo quindi la matrice \mathbb{P} le cui colonne sono i vettori $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_l$,

$$\mathbb{P} = (\tilde{\mathbf{a}}_1 \quad \tilde{\mathbf{a}}_2 \quad \cdots \quad \tilde{\mathbf{a}}_l), \quad \mathbb{P}^T = \begin{pmatrix} \tilde{\mathbf{a}}_1^T \\ \tilde{\mathbf{a}}_2^T \\ \vdots \\ \tilde{\mathbf{a}}_l^T \end{pmatrix}.$$

Evidentemente $\mathbb{P}^T \mathbb{P} = \mathbb{P} \mathbb{P}^T = \mathbb{I}$ e

$$\mathbb{P}^T \mathbb{A} \mathbb{P} = \mathbb{D}, \quad \text{dove } \mathbb{D} = \begin{pmatrix} \rho_1 & 0 & \cdots & 0 \\ 0 & \rho_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \rho_l \end{pmatrix}, \tag{5.133}$$

dove tutti i ρ_i sono positivi. Riscrivendo quindi la (5.132) otteniamo

$$\mu \underbrace{\mathbb{A} \mathbb{P} \mathbb{P}^T}_{\mathbb{I}} \tilde{\mathbf{v}} = \underbrace{\mathbb{B} \mathbb{P} \mathbb{P}^T}_{\mathbb{I}} \tilde{\mathbf{v}}, \quad \implies \quad \mu \underbrace{\mathbb{P}^T \mathbb{A} \mathbb{P}}_{\mathbb{D}} \mathbb{P}^T \tilde{\mathbf{v}} = \mathbb{P}^T \mathbb{B} \mathbb{P} \mathbb{P}^T \tilde{\mathbf{v}},$$

che, ponendo

$$\mathbb{P}^T \mathbb{B} \mathbb{P} = \mathbb{B}_1, \quad \text{e} \quad \tilde{\mathbf{V}} = \mathbb{P}^T \tilde{\mathbf{v}}, \quad (5.134)$$

riscriveremo come

$$\mu \mathbb{D} \tilde{\mathbf{V}} = \mathbb{B}_1 \tilde{\mathbf{V}}. \quad (5.135)$$

Definiamo adesso

$$\mathbb{D}^{1/2} = \begin{pmatrix} \sqrt{\rho_1} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \sqrt{\rho_l} \end{pmatrix}, \quad \mathbb{D}^{-1/2} = \begin{pmatrix} \frac{1}{\sqrt{\rho_1}} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\rho_l}} \end{pmatrix},$$

ed osserviamo che $\mathbb{D}^{1/2} \mathbb{D}^{-1/2} = \mathbb{D}^{-1/2} \mathbb{D}^{1/2} = \mathbb{I}$, e che $\mathbb{D}^{-1/2} \mathbb{D} \mathbb{D}^{-1/2} = \mathbb{I}$. Se moltiplichiamo dunque a sinistra la (5.135) per $\mathbb{D}^{-1/2}$ ed inseriamo fra \mathbb{B}_1 e $\tilde{\mathbf{V}}$ la matrice identità scritta come $\mathbb{I} = \mathbb{D}^{-1/2} \mathbb{D}^{1/2}$, si ottiene

$$\mu \mathbb{D}^{-1/2} \mathbb{D} \tilde{\mathbf{V}} = \mathbb{D}^{-1/2} \mathbb{B}_1 \underbrace{\mathbb{D}^{-1/2} \mathbb{D}^{1/2}}_{\mathbb{I}} \tilde{\mathbf{V}}, \quad (5.136)$$

che potremo scrivere come $\mu \mathbb{D}^{1/2} \tilde{\mathbf{V}} = \mathbb{B}_2 \mathbb{D}^{1/2} \tilde{\mathbf{V}}$, dove

$$\mathbb{B}_2 = \mathbb{D}^{-1/2} \mathbb{B}_1 \mathbb{D}^{-1/2} \stackrel{(5.134)}{=} \mathbb{D}^{-1/2} \mathbb{P}^T \mathbb{B} \mathbb{P} \mathbb{D}^{-1/2}. \quad (5.137)$$

Quindi, introducendo

$$\tilde{\mathbf{W}} = \mathbb{D}^{1/2} \tilde{\mathbf{V}}, \quad \Leftrightarrow \quad \tilde{\mathbf{V}} = \mathbb{D}^{-1/2} \tilde{\mathbf{W}},$$

la (5.136) si riscrive così

$$\mathbb{B}_2 \tilde{\mathbf{W}} = \mu \tilde{\mathbf{W}}.$$

Ora, tenendo presente la definizione (5.137), è facile provare che \mathbb{B}_2 è simmetrica. Possiamo quindi invocare il teorema spettrale per matrici simmetriche e concludere che esiste una base $\{\tilde{\mathbf{W}}_1, \dots, \tilde{\mathbf{W}}_l\}$ di autovettori a due a due ortogonali cui corrispondono l autovalori μ_k , reali. Scriveremo dunque

$$\mathbb{B}_2 \tilde{\mathbf{W}}_k = \mu_k \tilde{\mathbf{W}}_k, \quad k = 1, \dots, l, \quad \text{con} \quad \tilde{\mathbf{W}}_k \bullet \tilde{\mathbf{W}}_h = \delta_{kh}. \quad (5.138)$$

Chiaramente i vettori $\{\tilde{\mathbf{W}}_1, \dots, \tilde{\mathbf{W}}_l\}$ non sono i vettori $\{\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_l\}$ che stiamo cercando. Facciamo vedere che questi si ottengono a partire da $\{\tilde{\mathbf{W}}_1, \dots, \tilde{\mathbf{W}}_l\}$. Infatti, ripercorrendo all'indietro i passaggi abbiamo

$$\mathbb{D}^{-1/2} \mathbb{P}^T \mathbb{B} \mathbb{P} \mathbb{D}^{-1/2} \tilde{\mathbf{W}}_k = \mu_k \tilde{\mathbf{W}}_k, \quad \Rightarrow \quad \mathbb{B} \mathbb{P} \mathbb{D}^{-1/2} \tilde{\mathbf{W}}_k = \mu_k \underbrace{\mathbb{P}}_{\mathbb{D} \mathbb{D}^{-1/2}} \mathbb{D}^{1/2} \tilde{\mathbf{W}}_k,$$

cioè

$$\mathbb{B} \mathbb{P} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k = \mu_k \mathbb{P} \mathbb{D} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k = \mu_k \mathbb{P} \underbrace{\mathbb{D}^{-1/2} \mathbb{P}^T \mathbb{P}}_{\mathbb{I}} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k .$$

da cui, ricordando la (5.133) che scriviamo come $\mathbb{P} \mathbb{D} \mathbb{P}^T = \mathbb{A}$, si ottiene

$$\mathbb{B} \left(\mathbb{P} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k \right) = \mu_k \mathbb{A} \left(\mathbb{P} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k \right) .$$

Quindi gli l vettori linearmente indipendenti $\widetilde{\mathbf{v}}_k, k = 1, \dots, l$, che soddisfano la (5.125) sono dati da

$$\widetilde{\mathbf{v}}_k = \mathbb{P} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k, \quad k = 1, 2, \dots, l . \tag{5.139}$$

Rimane da dimostrare la (5.126) e che $\mu_k > 0$, per ogni $k = 1, \dots, l$, se \mathbb{B} è definita positiva. Per quel che riguarda la (5.126) abbiamo

$$\begin{aligned} \widetilde{\mathbf{v}}_h \bullet \mathbb{A} \widetilde{\mathbf{v}}_k &\stackrel{(5.139)}{=} \mathbb{P} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_h \bullet \mathbb{A} \mathbb{P} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k \\ &= \widetilde{\mathbf{W}}_h \bullet \underbrace{\left(\mathbb{D}^{-1/2} \right)^T}_{\mathbb{D}^{-1/2}} \underbrace{\mathbb{P}^T \mathbb{A} \mathbb{P}}_{\mathbb{D}} \mathbb{D}^{-1/2} \widetilde{\mathbf{W}}_k = \widetilde{\mathbf{W}}_h \bullet \underbrace{\mathbb{D}^{-1/2} \mathbb{D} \mathbb{D}^{-1/2}}_{\mathbb{I}} \widetilde{\mathbf{W}}_k \stackrel{(5.138)}{=} \delta_{hk} . \end{aligned}$$

Infine, per dimostrare che $\mu_k > 0$, per ogni $k = 1, \dots, l$, quando \mathbb{B} è definita positiva, è sufficiente applicare il lemma 5.10.1.

□

Esempio 5.10.1 Consideriamo²⁹ un punto materiale di massa m vincolato senza attrito al piano e soggetto all'energia potenziale di tipo elastico della forma

$$V(x, y) = \frac{k}{2} (x^2 + y^2) - \hat{\alpha} xy,$$

con $k > 0, \hat{\alpha} > 0$, costanti e dove $k > \hat{\alpha}$. Le posizioni di equilibrio si ottengono imponendo

$$\nabla V(x, y) = 0, \quad \Rightarrow \quad \begin{cases} kx - \hat{\alpha}y = 0, \\ ky - \hat{\alpha}x = 0, \end{cases} \quad \Rightarrow \quad P_e \equiv (0, 0) .$$

Se calcoliamo la matrice hessiana

$$\mathbb{B}(P_e) = \begin{pmatrix} k & -\hat{\alpha} \\ -\hat{\alpha} & k \end{pmatrix},$$

deduciamo che P_e è un punto di equilibrio stabile poiché entrambi gli autovalori di $\mathbb{B}(P_e)$ sono positivi (se $k > \hat{\alpha}$). Scrivendo le equazioni di moto otteniamo

$$\begin{cases} \ddot{x} + \omega_o^2 x - \alpha y = 0, \\ \ddot{y} + \omega_o^2 y - \alpha x = 0, \end{cases} \tag{5.140}$$

²⁹Questo esempio è tratto da L.D. Landau, E. M. Lifshits, Fisica Teorica 1 - Meccanica, Editori Riuniti, 1976.

dove $\omega_o^2 = k/m$, e $\alpha = \hat{\alpha}/m$. Ovviamente $\omega_o^2 > \alpha$. Il sistema (5.140) può riscriversi come la (5.118) se

$$\tilde{\mathbf{u}} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbb{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} \omega_o^2 & -\alpha \\ -\alpha & \omega_o^2 \end{pmatrix}.$$

Le frequenze proprie (5.124) si ottengono essenzialmente risolvendo l'equazione secolare (5.128), cioè

$$\det \begin{pmatrix} \omega_o^2 - \mu & -\alpha \\ -\alpha & \omega_o^2 - \mu \end{pmatrix} = 0, \quad \Rightarrow \quad \mu = \begin{cases} \mu_1 = \omega_o^2 - \alpha, \\ \mu_2 = \omega_o^2 + \alpha, \end{cases}$$

da cui $\omega_1 = \sqrt{\omega_o^2 - \alpha}$, $\omega_2 = \sqrt{\omega_o^2 + \alpha}$. Quindi, ricordando la (5.123), i modi normali sono

$$\tilde{\mathbf{w}} = \begin{pmatrix} a_1 \cos(\sqrt{\omega_o^2 - \alpha}t + \phi_1) \\ a_2 \cos(\sqrt{\omega_o^2 + \alpha}t + \phi_2) \end{pmatrix},$$

con $a_k, \phi_k, k = 1, 2$, le solite costanti di integrazione da calcolarsi tramite le condizioni iniziali. Per determinare i vettori $\tilde{\mathbf{v}}_k, k = 1, 2$, risolviamo la (5.129), cioè

$$\begin{pmatrix} \omega_o^2 & -\alpha \\ -\alpha & \omega_o^2 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \mu_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix},$$

$$\begin{pmatrix} \omega_o^2 & -\alpha \\ -\alpha & \omega_o^2 \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = \mu_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix},$$

con la condizione supplementare

$$\begin{pmatrix} v_{k1} \\ v_{k2} \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mathbb{A}} \begin{pmatrix} v_{k1} \\ v_{k2} \end{pmatrix} = v_{k1}^2 + v_{k2}^2 = 1, \quad k = 1, 2,$$

ottenendo

$$\tilde{\mathbf{v}}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \tilde{\mathbf{v}}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad \Rightarrow \quad \mathbb{V} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Le soluzioni nelle variabili dipendenti originarie x e y sono

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} a_1 \cos(\sqrt{\omega_o^2 - \alpha}t + \phi_1) \\ a_2 \cos(\sqrt{\omega_o^2 + \alpha}t + \phi_2) \end{pmatrix}.$$

E' interessante trovare l'espressione dei modi normali in funzione delle variabili x e y . Dalla (5.119) otteniamo

$$\begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix} = \mathbb{V}^{-1} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$

e quindi

$$w_1(t) = \frac{x(t) + y(t)}{\sqrt{2}}, \quad w_2(t) = \frac{x(t) - y(t)}{\sqrt{2}}.$$

5.10.3 Esempio: la catena di oscillatori

In questa sezione vogliamo studiare i modi normali e le frequenze proprie di oscillazione per un sistema meccanico caratterizzato da una Lagrangiana quadratica (per cui non ci sarà bisogno di approssimare la Lagrangiana ma solo di riscriverla utilizzando coordinate lagrangiane centrate nell'equilibrio).

Consideriamo un sistema di N punti materiali $\{P_i, m_i\}$, $i = 1, \dots, N$, che si muovono senza attrito su una retta orizzontale³⁰ e sono collegati l'uno all'altro da molle le cui masse e lunghezza a riposo sono trascurabili. Questa "catena di oscillatori" è fissata ai punti fissi A e B , posti a distanza L l'uno dall'altro, come mostrato in figura 5.7. La forza totale F_i agente sull' i -esimo punto materiale è data da³¹

$$F_i = \underbrace{-k_i(x_i - x_{i-1})}_{\substack{\text{forza su } P_i \\ \text{esercitata da } P_{i-1}}} + \underbrace{k_{i+1}(x_{i+1} - x_i)}_{\substack{\text{forza su } P_i \\ \text{esercitata da } P_{i+1}}}, \quad i = 1, \dots, N. \quad (5.141)$$

dove:

- k_j , $j = 1, \dots, N + 1$, è il coefficiente elastico (o **rigidezza**) della j -esima molla. In particolare, la molla che collega il primo punto materiale della catena, cioè P_1 , al punto A ha rigidezza k_1 , mentre quella che la collega l'ultimo punto materiale P_N a B ha rigidezza k_{N+1} .
- x_r , $r = 1, \dots, N$, indica l'ascissa del punto r -esimo punto materiale. Inoltre, $x_{N+1} = x_B = L$ e $x_0 = x_A$ sono, rispettivamente, le ascisse dei punti fissi B e A a cui la catena è collegata (vedi figura 5.7).

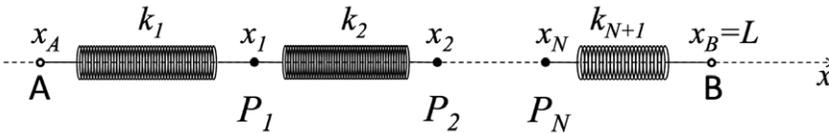


Figura 5.7: Catena di N punti materiali mobili P_1, \dots, P_N , di ascisse x_1, \dots, x_N . La catena è collegata ai punti fissi A e B , di ascisse x_A e $x_B = L$, da molle di rigidezza k_1 e k_{N+1} , rispettivamente. Inoltre i punti materiali mobili sono collegati fra loro da molle le cui rigidezze sono k_2, \dots, k_N . Tutte le molle hanno lunghezza a riposo e masse trascurabili.

La Lagrangiana del sistema è data da

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^N m_i \dot{x}_i^2 - \frac{1}{2} \sum_{i=0}^N k_{i+1} (x_{i+1} - x_i)^2.$$

³⁰Per cui la forza peso non ha nessuna influenza.

³¹ P_0 corrisponde al punto fisso A , mentre P_{N+1} al punto fisso B .

Cerchiamo adesso gli eventuali punti di equilibrio. Questi corrispondono alle configurazioni del sistema nelle quali la risultante delle forze agenti su ciascun punto è nulla. Imponiamo pertanto

$$\begin{cases} F_1 = 0, \\ \vdots \\ F_N = 0, \end{cases} \Rightarrow \begin{cases} -k_1(x_1 - x_0) - k_2(x_1 - x_2) = 0, \\ \vdots \\ -k_N(x_N - x_{N-1}) - k_{N+1}(x_N - L) = 0, \end{cases} \quad (5.142)$$

che scritto in forma esplicita diventa

$$\underbrace{\begin{pmatrix} -(k_1 + k_2) & k_2 & \cdot & 0 \\ k_2 & -(k_2 + k_3) & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & -(k_N + k_{N+1}) \end{pmatrix}}_{\mathbb{M}} \underbrace{\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} -k_1 x_0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ -k_{N+1} L \end{pmatrix}}_{\mathbf{b}}.$$

Tale sistema ammette una e una sola soluzione poiché la matrice \mathbb{M} è invertibile³². La soluzione, che corrisponde all'unica configurazione di equilibrio cercata, è dunque

$$\mathbf{x}^e = \mathbb{M}^{-1} \mathbf{b}. \quad (5.143)$$

Eseguiamo ora il cambio di coordinate

$$q_i = x_i - x_i^e, \quad i = 1, \dots, N, \quad (5.144)$$

dove $x_i^e, i = 1, \dots, N$, è la posizione di equilibrio dell' i -esimo punto materiale (ottenuta dalla (5.143)). Dalla (5.144) otteniamo $\dot{x}_i = \dot{q}_i$, mentre la configurazione di equilibrio corrisponderà a $q_i = 0, i = 1, \dots, N$. L'energia potenziale V nelle nuove variabili diventa

$$\begin{aligned} \hat{V}(q_1, \dots, q_N) &= \frac{1}{2} \sum_{i=0}^N k_{i+1} [(q_{i+1} + x_{i+1}^e) - (q_i + x_i^e)]^2 \\ &= \frac{1}{2} \sum_{i=0}^N k_{i+1} (q_{i+1} - q_i)^2 + \frac{1}{2} \sum_{i=0}^N k_{i+1} (x_{i+1}^e - x_i^e)^2 \\ &\quad + \sum_{i=0}^N k_{i+1} (q_{i+1} - q_i) (x_{i+1}^e - x_i^e), \end{aligned} \quad (5.145)$$

dove abbiamo posto $q_0 = q_{N+1} = 0, x_0^e = x_0$ e $x_{N+1}^e = L$.

³²E' facile provare che \mathbb{M} è invertibile essendo essa una matrice a bande e simmetrica. La dimostrazione si fa per induzione. Definiamo con \mathbb{M}_N la matrice $N \times N$. Per $N = 1$ è banale mostrare l'invertibilità, perché $\mathbb{M}_1 = -(k_1 + k_2)$. Supponiamo ora che \mathbb{M}_N sia invertibile e dimostriamo che lo è anche \mathbb{M}_{N+1} . Ricordando che una matrice è invertibile se e solo se ha determinante non nullo, abbiamo

$$\det(\mathbb{M}_{N+1}) = -(k_N + k_{N+1}) \det(\mathbb{M}_N) \neq 0,$$

ed è dunque chiaro che \mathbb{M}_{N+1} è invertibile.

Se adesso calcoliamo $\frac{\partial \hat{V}}{\partial q_i}$ per ogni $i = 1, \dots, N$, si ottiene

$$\frac{\partial \hat{V}}{\partial q_i} = \underbrace{k_i (q_i - q_{i-1}) - k_{i+1} (q_{i+1} - q_i)}_{\substack{\text{termini che derivano dal primo} \\ \text{addendo della (5.145)}}} + \underbrace{k_i (x_i^e - x_{i-1}^e) - k_{i+1} (x_{i+1}^e - x_i^e)}_{\substack{\text{termini che derivano dal terzo} \\ \text{addendo della (5.145)}}}.$$

Tuttavia $k_i (x_i^e - x_{i-1}^e) - k_{i+1} (x_{i+1}^e - x_i^e) = 0$, poiché $x_i^e, i = 1, \dots, N$, sono proprio le soluzioni del sistema (5.142). Di conseguenza il terzo addendo della (5.145) così come il secondo (che è costante) è del tutto ininfluenza ai fini delle equazioni di Lagrange. Possiamo dunque riscrivere la Lagrangiana come

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^N m_i \dot{q}_i^2 - \frac{1}{2} \sum_{i=0}^N k_{i+1} (q_{i+1} - q_i)^2,$$

dove, lo ricordiamo ancora, $q_0 = q_{N+1} = 0$. Questa Lagrangiana ci dà il seguente sistema di equazioni di Lagrange³³

$$m_i \ddot{q}_i = -k_{i-1} (q_i - q_{i-1}) + k_i (q_{i+1} - q_i), \quad i = 1, \dots, N,$$

che è appunto un sistema di equazioni lineari del secondo ordine lineari reciprocamente accoppiate. L'idea è quindi quella di disaccoppiare le equazioni tramite un opportuno cambio di variabili, giungendo così ad N equazioni del tipo oscillatore armonico.

Se ci limitiamo a considerare il caso in cui tutte le masse siano uguali a m , e tutte le molle abbiano la stessa costante elastica k , allora la Lagrangiana si semplifica e (utilizzando la funzione $L = \mathcal{L}/m$ invece della Lagrangiana \mathcal{L}) otteniamo

$$L = \frac{1}{2} \sum_{i=1}^N \dot{q}_i^2 - \frac{\omega^2}{2} \sum_{i=0}^N (q_{i+1} - q_i)^2,$$

dove, al solito, $\omega^2 = k/m$. Le equazioni di Lagrange prendono la forma

$$\ddot{q}_i = -\omega^2 (q_{i-1} - 2q_i + q_{i+1}), \quad i = 1, \dots, N.$$

Le matrici \mathbb{A} e \mathbb{B} che intervengono nella scrittura di L nella forma matriciale (5.117) sono rispettivamente la matrice identità e la matrice

$$\mathbb{B} = -\omega^2 \begin{pmatrix} -2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 1 & -2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 & -2 & 1 & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 & -2 & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 & -2 \end{pmatrix}.$$

³³Con le condizioni aggiuntive $q_0 = q_{N+1} = 0$.

Quindi la determinazione dei modi normali e delle loro frequenze è ridotta alla ricerca degli autovettori e degli autovalori della matrice \mathbb{B} .

E' facile verificare che gli N autovettori sono dati da

$$(\tilde{\mathbf{v}}_s)_r = \sin\left(\frac{sr\pi}{N+1}\right), \quad r = 1, \dots, N, \quad s = 1, \dots, N, \quad (5.146)$$

dove $(\tilde{\mathbf{v}}_s)_r$ è la r -esima componente del s -esimo autovettore. Il corrispondente autovalore è

$$\lambda_s = 2\omega^2 \left(1 - \cos\left(\frac{s\pi}{N+1}\right)\right). \quad (5.147)$$

Lasciamo come esercizio la verifica che i vettori (5.146) sono effettivamente gli autovettori di \mathbb{B} con autovalori λ_s dati dalla (5.147). Suggerimento: usare le formule di addizione del seno e calcolare

$$-\sin\left(\frac{s(r-1)\pi}{N+1}\right) + 2\sin\left(\frac{sr\pi}{N+1}\right) - \sin\left(\frac{s(r+1)\pi}{N+1}\right).$$

Nota 5.10.1 *Se si scorrono le componenti degli autovettori dalla prima alla N -esima, si osserva che nel primo autovettore ($s = 1$) tutte le componenti hanno lo stesso segno, mentre nel secondo si hanno metà componenti positive e metà negative (se N è pari, se $N = 2M + 1$ la componente M -esima è nulla) e in generale il numero di cambiamenti di segno delle componenti è uguale a $s - 1$.*

*Inoltre il primo autovalore è il più piccolo degli autovalori, che sono una successione crescente, $\lambda_1 < \lambda_2 < \dots < \lambda_N$. Questo significa che le frequenze di oscillazione aumentano all'aumentare di k , ovvero all'aumentare del numero di cambi di segno delle componenti.*³⁴

Per chiarire la dinamica di questo sistema, vediamo il caso semplice di due soli punti. In questo caso gli autovettori sono semplicemente

$$\tilde{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ con autovalore } \omega^2, \text{ e } \tilde{\mathbf{v}}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ con autovalore } 3\omega^2.$$

Nel modo normale associato al primo autovettore, le due masse si muovono oscillando nella stessa direzione e con la stessa frequenza ed ampiezza, mantenendo quindi inalterata la distanza reciproca (uguale alla distanza all'equilibrio). Nel secondo modo invece le due masse oscillano ancora con la stessa frequenza ed ampiezza, ma in due direzioni opposte: in questo caso il centro di massa delle molle resta fermo durante il moto. Il moto generico del sistema è dunque una sovrapposizione (combinazione lineare) di questi due moti.

³⁴Si veda anche la sezione 44 di F. Gantmacher, *Lectures in Analytical Mechanics*, MIR 1975.

Capitolo 6

Cinematica dei Sistemi Rigidi

6.1 Introduzione

In questo capitolo descriveremo come si muove un corpo rigido nello spazio. Un **corpo**, o **sistema, rigido** è un insieme di punti materiali vincolati a mantenere inalterate le loro mutue distanze.

Lo scopo del capitolo è duplice: da un lato vogliamo determinare la relazione fondamentale che lega le velocità di punti differenti che “partecipano di uno stesso moto rigido” (ovvero si muovono mantenendo inalterate le mutue distanze) rispetto a un *osservatore* fisso; dall’altro vogliamo trovare le relazioni che legano le osservazioni che *due* osservatori, in moto tra loro, fanno dei moti di altri “corpi” (punti o corpi rigidi). Quest’ultimo problema prende il nome di *cinematica relativa*.

Con l’espressione *osservatore* ci riferiamo a un sistema cartesiano ortogonale, con orientamento fissato, cioè nel quale sono determinati: un’origine, un’unità di misura per le lunghezze, tre assi ortogonali orientati (ovvero tra i quali si sia fissato un *ordine*, per esempio gli assi x , y e z disposti come i diti pollice, indice e medio della mano destra¹, come specificato nella sezione 1.5.1) e un *orologio* per misurare il tempo.

6.2 Moti rigidi

Sia fissato un osservatore, ovvero un sistema di riferimento (SdR) cartesiano ortogonale levogiro centrato in Ω , i cui assi coordinati sono ξ , η e ζ . Indichiamo con $\Sigma = (\Omega, \xi, \eta, \zeta)$ tale SdR, oppure con (Ω, e_1, e_2, e_3) , essendo e_i , $i = 1, 2, 3$, i versori degli assi coordinati, ovvero i **versori di una terna ortonormale levogira**. L’osservatore Σ viene detto **osservatore fisso**², o **sistema di riferimento (SdR) fisso**, o **terna fissa**. Viene poi dato un sistema rigido, che per brevità chiameremo \mathfrak{B} , di cui vogliamo studiare il moto rispetto all’osservatore Σ .

¹Ricordiamo che, in questo caso, il sistema di riferimento viene detto *destrorso*, o *levogiro*. Si ricordi inoltre che un riferimento destrorso e uno non destrorso, cioè sinistrorso, non possono essere sovrapposti mantenendo l’ordine degli assi.

²Non indichiamo mai esplicitamente la scelta dell’orologio per un osservatore. In questo paragrafo la scelta dell’orologio non è molto importante in quanto l’osservatore fisso è il solo che vede i punti in movimento. Dovremo tornare su questo problema nel caso della cinematica relativa.

Per prima cosa eliminiamo dalla nostra analisi il caso “degenere” in cui tutti i punti del sistema rigido siano allineati (o ci siano solo due punti).

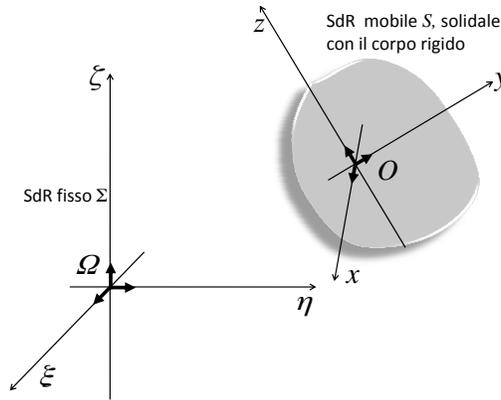


Figura 6.1: Sistema di riferimento - SdR - fisso Σ , e SdR solidale S .

Consideriamo quindi un sistema rigido che abbia almeno tre punti non allineati. Siano essi P_1 , P_2 , e P_3 . Al corpo \mathfrak{B} è possibile associare una **terna ortonormale levogira** $S = (O, x, y, z)$ di versori \mathbf{i} , \mathbf{j} , \mathbf{k} , costruita nel seguente modo: poniamo $O = P_1$, $\mathbf{i} = \text{vers}(P_2 - P_1)$, $\mathbf{j} = \text{vers}(\mathbf{i} \wedge (P_3 - P_1))$ e infine $\mathbf{k} = \mathbf{i} \wedge \mathbf{j}$. Il sistema di riferimento S , viene anche detto **sistema di riferimento (SdR) solidale**, o **terna solidale**, o anche **osservatore solidale**.

Ora, la geometria di \mathfrak{B} è nota e pertanto conosciamo la posizione di ogni suo punto rispetto al sistema solidale. Quindi ogni $P \in \mathfrak{B}$ è individuato dalla terna $\mathbf{x} \in \mathbb{R}^3$, data da

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

i cui elementi rappresentano le coordinate di P rispetto ad S , ovvero le componenti del vettore $(P - O)$ rispetto a S , $\mathbf{x} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Con \mathbf{x}^T si denota, come al solito, la terna trasposta, cioè

$$\mathbf{x}^T = (x, y, z).$$

Notiamo che una volta determinata la terna solidale, il numero di punti di cui è composto \mathfrak{B} non ha più importanza: ogni punto P viene ora individuato dalle sue coordinate \mathbf{x} , rispetto alla terna solidale³. Quest'ultime inoltre sono **costanti nel tempo**. Quindi la posizione dei punti di \mathfrak{B} nello spazio (e non rispetto alla terna solidale) è determinata dalla posizione della terna solidale stessa. In altre parole: **il moto di un sistema rigido nello spazio è equivalente al moto di una terna di riferimento**. Infatti, siccome i punti che partecipano al moto rigido di \mathfrak{B} hanno coordinate (x, y, z) costanti

³Questa semplice osservazione ha una conseguenza importante: la cinematica dei sistemi rigidi discreti e dei sistemi rigidi continui può essere descritta allo stesso modo.

nel SdR solidale, è sufficiente conoscere la posizione di S rispetto a Σ per localizzare ogni punto del sistema rigido. Abbiamo quindi ridotto il problema del moto rigido al moto relativo di due sistemi di riferimento nello spazio euclideo tridimensionale, ovvero a situare ad ogni istante la terna mobile S rispetto alla terna fissa Σ .

A sua volta questo problema si scompone naturalmente nella localizzazione dell'origine O della terna mobile S rispetto alla terna fissa, e all'orientazione dei versori di S rispetto ai versori della terna fissa Σ .

6.2.1 Primo caso: $\Omega \equiv O$

Consideriamo dapprima il caso in cui le origini dei due SdR O e Ω rimangono coincidenti. Le coordinate del generico punto $P \in \mathfrak{B}$, saranno espresse dalla terna x nel SdR S , e dalla terna $\xi \in \mathbb{R}^3$,

$$\xi = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix},$$

nel SdR Σ , ovvero

$$(P - O) \longrightarrow \begin{cases} \xi = \xi e_1 + \eta e_2 + \zeta e_3, & \text{rispetto al SdR fisso } \Sigma, \\ x = xi + yj + zk, & \text{rispetto al SdR solidale } S. \end{cases}$$

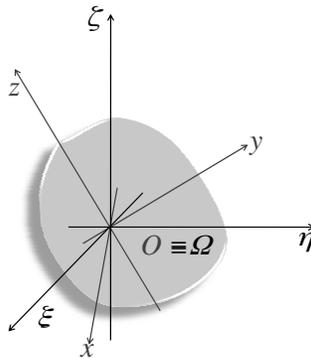


Figura 6.2: Primo caso: $\Omega \equiv O$.

Indichiamo con

$$\xi = F(x), \tag{6.1}$$

l'applicazione da \mathbb{R}^3 in \mathbb{R}^3 che trasforma le coordinate dal sistema S al riferimento fisso⁴ Σ . Siccome il corpo è rigido, la distanza fra due punti qualsiasi di \mathfrak{B} deve essere

⁴Le coordinate dei punti nel sistema S , sono note dal momento che \mathfrak{B} è dato (cioè si conosce la geometria del corpo). Non sono invece note le coordinate ξ dei punti rispetto al SdR Σ .

la stessa in entrambe i riferimenti, $|\xi_1 - \xi_2| = |\mathbf{x}_1 - \mathbf{x}_2|$, cioè

$$|F(\mathbf{x}_1) - F(\mathbf{x}_2)| = |\mathbf{x}_1 - \mathbf{x}_2|, \quad (6.2)$$

ed inoltre

$$F(\mathbf{0}) = \mathbf{0}, \quad (6.3)$$

dal momento che O rimane sempre coincidente con Ω .

Definizione 6.2.1 *Un'applicazione da \mathbb{R}^3 in \mathbb{R}^3 , che verifica le (6.2) e (6.3), si dice isometria ortogonale.*

Abbiamo quindi questo importante risultato

Teorema 6.2.1 *Se F è un isometria ortogonale, allora:*

1. F è un'applicazione lineare, ovvero possiamo far corrispondere ad F una matrice \mathbb{A} , tale che

$$F(\mathbf{x}) = \mathbb{A}\mathbf{x}.$$

2. La matrice \mathbb{A} è ortogonale (si veda la sezione 1.3), cioè $\mathbb{A}^T = \mathbb{A}^{-1}$, e di conseguenza

$$\det \mathbb{A} = \pm 1. \quad (6.4)$$

Dim⁵. E' facile provare che l'isometria F conserva il prodotto scalare. Infatti,

$$|F(\mathbf{x}_1) - F(\mathbf{x}_2)|^2 = |\mathbf{x}_1 - \mathbf{x}_2|^2,$$

implica

$$\underbrace{|F(\mathbf{x}_1)|^2}_{|\mathbf{x}_1|^2} + \underbrace{|F(\mathbf{x}_2)|^2}_{|\mathbf{x}_2|^2} - 2F(\mathbf{x}_1) \cdot F(\mathbf{x}_2) = |\mathbf{x}_1|^2 + |\mathbf{x}_2|^2 - 2\mathbf{x}_1 \cdot \mathbf{x}_2,$$

e dunque $F(\mathbf{x}_1) \cdot F(\mathbf{x}_2) = \mathbf{x}_1 \cdot \mathbf{x}_2$.

La linearità si prova mostrando che $\forall \mathbf{x}_1$ e \mathbf{x}_2 di \mathbb{R}^3 , e per ogni coppia di numeri reali m, n , sia ha

$$F(m\mathbf{x}_1 + n\mathbf{x}_2) = mF(\mathbf{x}_1) + nF(\mathbf{x}_2),$$

ovvero che

$$|F(m\mathbf{x}_1 + n\mathbf{x}_2) - mF(\mathbf{x}_1) - nF(\mathbf{x}_2)| = 0.$$

La dimostrazione di questo passaggio viene lasciata per esercizio. Quindi se F è lineare e $F(\mathbf{0}) = \mathbf{0}$, allora

$$F(\mathbf{x}) = \mathbb{A}\mathbf{x},$$

dove \mathbb{A} è un matrice 3×3 . Notiamo inoltre che la matrice \mathbb{A} è sicuramente invertibile. Infatti, siccome $|\mathbb{A}\mathbf{x}| = |\mathbf{x}|$, abbiamo $\mathbb{A}\mathbf{x} = \mathbf{0}$ se e solo se $\mathbf{x} = \mathbf{0}$.

Il punto 2 si ottiene provando che

$$\mathbb{A}^T \mathbb{A} = \mathbb{I}, \quad (6.5)$$

⁵La dimostrazione del presente teorema è ripresa da Giuseppe Benfatto, *Note del Corso di Meccanica Analitica*, Università di Roma "Tor Vergata", a.a. 2007-08.

essendo \mathbb{I} la matrice identità. Per la dimostrazione⁶ della (6.5) rimandiamo alla sezione 1.3. Di conseguenza, moltiplicando a destra per \mathbb{A}^{-1} ambo i membri della (6.5), abbiamo $\mathbb{A}^T = \mathbb{A}^{-1}$. Infine, sempre dalla (6.5)

$$1 = \det \mathbb{I} = \det (\mathbb{A}^T \mathbb{A}) = \underbrace{\det \mathbb{A}^T}_{\det \mathbb{A}} \det \mathbb{A} = (\det \mathbb{A})^2,$$

da cui deriva banalmente la (6.4).

□

La condizione $F(\mathbf{x}_1) \cdot F(\mathbf{x}_2) = \mathbf{x}_1 \cdot \mathbf{x}_2$, comporta la conservazione degli angoli. Infatti, se θ_1 è l'angolo fra $F(\mathbf{x}_1)$, $F(\mathbf{x}_2)$, e θ_2 è l'angolo fra \mathbf{x}_1 , \mathbf{x}_2 , abbiamo

$$\left. \begin{aligned} F(\mathbf{x}_1) \cdot F(\mathbf{x}_2) &= \underbrace{|F(\mathbf{x}_1)| |F(\mathbf{x}_2)|}_{|\mathbf{x}_1| |\mathbf{x}_2|} \cos \theta_1 \\ \mathbf{x}_1 \cdot \mathbf{x}_2 &= |\mathbf{x}_1| |\mathbf{x}_2| \cos \theta_2 \end{aligned} \right\} \Rightarrow \cos \theta_1 = \cos \theta_2,$$

da cui $\theta_1 = \pm \theta_2$. E dunque, considerando gli angoli privi di segno, $\theta_1 = \theta_2$.

Vediamo adesso come determinare la matrice \mathbb{A} . Consideriamo le immagini mediante F delle terne rappresentanti i versori \mathbf{i} , \mathbf{j} e \mathbf{k} (ovvero le componenti in Σ dei versori di S)

$$F \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{i}} = \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix}, \quad F \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{j}} = \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix}, \quad F \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\mathbf{k}} = \begin{pmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{pmatrix}.$$

Ora, data la linearità di F abbiamo

$$\begin{aligned} F(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) &= xF(\mathbf{i}) + yF(\mathbf{j}) + zF(\mathbf{k}) \\ &= x \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} + y \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} + z \begin{pmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \end{aligned}$$

Quindi

$$\mathbb{A} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix}. \quad (6.6)$$

⁶La dimostrazione è immediata se ricordiamo che $F(\mathbf{x}_2) \cdot F(\mathbf{x}_2) = \mathbf{x}_1 \cdot \mathbf{x}_2$. Abbiamo infatti

$$\mathbb{A}\mathbf{x}_1 \cdot \mathbb{A}\mathbf{x}_2 = \mathbb{A}^T \mathbb{A}\mathbf{x}_1 \cdot \mathbf{x}_2 = \mathbf{x}_1 \cdot \mathbf{x}_2,$$

da cui discende (6.5).

Ne segue che la relazione fra terne (6.1) si potrà esplicitamente scrivere come

$$\xi = \mathbb{A}x, \quad \Leftrightarrow \quad \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Il cambiamento dal SdR solidale S al SdR fisso Σ , è **completamente descritto dalla matrice** \mathbb{A} . In generale le componenti di \mathbb{A} dipenderanno dal tempo, dal momento che l'orientazione dei versori i, j e k rispetto a Σ può variare nel tempo. Scriveremo pertanto $\mathbb{A} = \mathbb{A}(t)$, intendendo

$$\mathbb{A}(t) = \begin{pmatrix} \alpha_1(t) & \alpha_2(t) & \alpha_3(t) \\ \beta_1(t) & \beta_2(t) & \beta_3(t) \\ \gamma_1(t) & \gamma_2(t) & \gamma_3(t) \end{pmatrix} \quad (6.7)$$

Osserviamo che la condizione (6.5) rappresenta un “vincolo” sulle nove componenti della matrice \mathbb{A} (tale concetto verrà approfondito nella sezione 6.2.3). Infatti i versori devono soddisfare la condizione di ortonormalità, ovvero le equazioni

$$i \cdot i = 1, \quad j \cdot j = 1, \quad k \cdot k = 1, \quad i \cdot j = 0, \quad j \cdot k = 0, \quad k \cdot i = 0, \quad (6.8)$$

che si traduce nelle sei⁷ condizioni date dalla (6.5), ovvero

$$\begin{aligned} \mathbb{A}i \cdot \mathbb{A}j &= 0, & \mathbb{A}j \cdot \mathbb{A}k &= 0, & \mathbb{A}k \cdot \mathbb{A}i &= 0, \\ \mathbb{A}i \cdot \mathbb{A}i &= 1, & \mathbb{A}j \cdot \mathbb{A}j &= 1, & \mathbb{A}k \cdot \mathbb{A}k &= 1. \end{aligned} \quad (6.9)$$

Com'è noto, le matrici 3×3 che soddisfano la condizione (6.5), ovvero $\mathbb{A}^T = \mathbb{A}^{-1}$, sono dette **ortogonali**.

Ora, siccome stiamo studiando moti continui dello spazio, possiamo escludere le matrici a determinante negativo (che rappresentano un moto rigido più un ribaltamento speculare: quest'ultimo infatti non è riconducibile con continuità, attraverso moti rigidi, all'identità, che, a sua volta, indica lo “stato di partenza”)⁸ e limitarci alle matrici a determinante uguale a uno,

$$\det \mathbb{A} = +1. \quad (6.10)$$

Vediamo adesso un'importante teorema, usualmente noto come **teorema di Eulero**.

Teorema 6.2.2 *Sia \mathbb{A} una matrice ortogonale, $\mathbb{A} \neq \mathbb{I}$. Allora esiste $x \in \mathbb{R}^3$, $x \neq 0$, definita a meno di una costante moltiplicativa (ovvero esiste una direzione individuata dal vettore associato alla terna x) per cui*

$$\mathbb{A}x = x. \quad (6.11)$$

⁷La (6.5) dà luogo a sei equazioni indipendenti, e non nove, dal momento che $\mathbb{A}^T \mathbb{A}$ è simmetrica.

⁸In altre parole, e in modo più formale: prendiamo la posizione a un certo istante come “configurazione di riferimento”. Possiamo sempre orientare gli assi del sistema fisso in modo che questi coincidano con gli assi solidali. Avremo quindi $\mathbb{A}(0) = \mathbb{I}$. La (6.4) ci dice che, al variare del tempo $\mathbb{A}(t)$ può assumere solo i valori 1 e -1 . Ma il determinante è una funzione continua delle componenti della matrice, pensata come un elemento di $\mathbb{R}^{3 \times 3}$, quindi non può passare con continuità dal valore iniziale 1 al valore -1 .

Dim. Innanzitutto osserviamo che \mathbb{A} non ammette autovalori nulli⁹. Proviamo adesso che **se λ è un autovalore reale di \mathbb{A}** , ovvero se $\mathbb{A}z = \lambda z$ con $\lambda \in \mathbb{R}$, **allora $|\lambda| = 1$** . Dalla (6.5) infatti

$$|z|^2 = z \cdot z = z \cdot \underbrace{\mathbb{A}^T \mathbb{A}}_{\mathbb{I}} z = \mathbb{A} z \cdot \mathbb{A} z = \lambda^2 |z|^2 .$$

Abbiamo quindi dimostrato che **tutti gli autovalori reali di \mathbb{A} hanno modulo unitario**.

Supponiamo adesso che $\lambda_i, i = 1, 2, 3$, siano i tre autovalori di \mathbb{A} . La (6.10) comporta

$$1 = \det \mathbb{A} = \lambda_1 \lambda_2 \lambda_3 . \tag{6.12}$$

Si hanno quindi due casi:

1. Due autovalori sono complessi, diciamo λ_1 e λ_2 , mentre $\lambda_3 \in \mathbb{R}$ e, per quanto visto, $|\lambda_3| = 1$. Siccome gli autovalori sono radici dell'equazione a coefficienti reali $\det(\mathbb{A} - \lambda \mathbb{I}) = 0$, λ_1 e λ_2 sono uno il complesso coniugato dell'altro. Ne segue $\lambda_1 \lambda_2 = |\lambda_1|^2$. Riscrivendo la (6.12) abbiamo $|\lambda_1|^2 \lambda_3 = 1$, e dunque, per evitare contraddizioni nei segni, deve essere $\lambda_3 = 1$. Evidentemente i tre autovettori non possono essere tutti e tre complessi per il fatto che, come già detto, le radici di $\det(\mathbb{A} - \lambda \mathbb{I}) = 0$ sono coppie di numeri complessi coniugati.
2. Tutti e tre gli autovalori sono reali e inoltre $|\lambda_i| = 1, i = 1, 2, 3$. In tal caso necessariamente almeno uno deve essere uguale a +1.

Dunque, in entrambi i casi esiste un autovalore $\lambda = 1$, cui corrisponde (a meno di una costante moltiplicativa) un autovettore x che soddisfa la (6.11).

□

In sostanza il teorema di Eulero afferma che esiste una retta r , eventualmente dipendente dal tempo, i cui punti sono invarianti per l'applicazione di \mathbb{A} : ovvero un qualunque vettore $(P - O) \in r$ è rappresentato in S ed in Σ dalla medesima terna. Non solo, ma valgono anche le seguenti proprietà:

- (a) Detto P_o un qualsiasi punto di r , individuato dalla terna x_o in S , il piano $\pi_{\perp r}$ per P_o ortogonale ad r è descritto nel sistema S dalla equazione $(x - x_o) \cdot \nu = 0$, essendo ν il versore di r . Consideriamo adesso $\mathbb{A}x$, e mostriamo che soddisfa la stessa equazione, cioè $(\mathbb{A}x - x_o) \cdot \nu = 0$. Infatti¹⁰

$$0 = \underbrace{\mathbb{A}^T \mathbb{A}}_{\mathbb{I}} (x - x_o) \cdot \nu = \mathbb{A} (x - x_o) \cdot \underbrace{\mathbb{A} \nu}_{\nu} = (\mathbb{A}x - \mathbb{A}x_o) \cdot \nu = (\mathbb{A}x - x_o) \cdot \nu .$$

Questo vuol dire che, comunque si scelga una generica terna x di S , $\mathbb{A}x$ è giace sullo stesso piano ortogonale ad r su cui giace x .

⁹Se \mathbb{A} avesse autovalori nulli la (6.10) sarebbe contraddetta.

¹⁰Si ricordi che $\mathbb{A}x_o = x_o$, come del resto $\mathbb{A}\nu = \nu$.

- (b) Consideriamo ancora il generico punto P_o della retta r , ed il solito piano $\pi_{\perp r}$ passante per P_o . Sia q una generica retta giacente su tale piano e passante per P_o , la cui direzione è individuata, in S , dal versore \mathbf{q} . Quindi, nel SdR S , l'equazione parametrica della retta q è $\mathbf{x} - \mathbf{x}_o = \mu \mathbf{q}$, essendo $\mu \in \mathbb{R}$, il parametro. Proviamo adesso che q è trasformata, nel SdR Σ , in un'altra retta. Infatti

$$\mathbf{x} - \mathbf{x}_o = \mu \mathbf{q}, \Rightarrow \underbrace{\mathbb{A}^T \mathbb{A}}_{\mathbb{I}} (\mathbf{x} - \mathbf{x}_o) = \mu \mathbf{q}.$$

Quindi moltiplicando a sinistra per \mathbb{A} si ha

$$\mathbb{A}(\mathbf{x} - \mathbf{x}_o) = \mu \mathbb{A} \mathbf{q}, \Rightarrow \underbrace{\mathbb{A} \mathbf{x}}_{\xi} - \underbrace{\mathbb{A} \mathbf{x}_o}_{\xi_o} = \mu \mathbb{A} \mathbf{q},$$

che è l'equazione parametrica, nel SdR Σ , di una retta di direzione $\mathbb{A} \mathbf{q}$.

- (c) Mostriamo infine che, data una qualsiasi retta q di $\pi_{\perp r}$ per P_o , l'angolo fra q e $\mathbb{A} \mathbf{q}$ è indipendente dalla retta selezionata. Infatti, consideriamo un'altra retta p , diversa da q , giacente sempre sul piano $\pi_{\perp r}$ e passante per P_o . Sia β l'angolo che p forma con la retta q . Ora, siccome l'applicazione \mathbb{A} conserva gli angoli, l'angolo fra $\mathbb{A} \mathbf{q}$ e $\mathbb{A} \mathbf{p}$ è sempre β . Di conseguenza, riferendoci alla figura 6.3 se θ_q è l'angolo fra la retta q e la retta $\mathbb{A} \mathbf{q}$, e θ_p è quello fra p e $\mathbb{A} \mathbf{p}$, abbiamo $\theta_q + \beta = \theta_p + \beta, \Rightarrow \theta_p = \theta_q$. Si conclude quindi che, dato un generico \mathbf{x} , l'angolo fra $(\mathbf{x} - \mathbf{x}_o)$ e $\mathbb{A}(\mathbf{x} - \mathbf{x}_o)$, è indipendente da \mathbf{x} .

Quindi, volendo riformulare alla luce delle proprietà (a), (b) e (c) l'enunciato del teorema di Eulero si potrebbe renderlo come: **dati due sistemi di riferimento** (nello specifico il SdR S ed il SdR Σ) **è sempre possibile sovrapporre il primo al secondo con un'unica rotazione intorno ad un opportuno asse**. Infatti: (i) l'applicazione rappresentata dalla matrice \mathbb{A} lascia invariate lunghezze ed angoli; (ii) le terne \mathbf{x} (ovvero le coordinate dei punti) che giacciono su r non vengono mutate dalla rotazione; (iii) le coordinate dei punti che non stanno sulla retta r sono comunque trasportate in altre terne che giacciono nel medesimo piano ortogonale ad r ; (iv) le rette del piano $\pi_{\perp r}$, passanti per il punto P_o sono "mappate" in altre rette; (v) l'angolo, misurato sul piano $\pi_{\perp r}$, fra \mathbf{x} e $\mathbb{A} \mathbf{x}$, è indipendente da \mathbf{x} .

L'asse di rotazione si individua determinando l'autovettore di \mathbb{A} relativo all'autovalore $\lambda = 1$, mentre l'angolo di rotazione si determina calcolando la traccia della matrice¹¹ \mathbb{A} . Infatti, tornando alla notazione del teorema di Eulero, se $\lambda_3 = 1$, gli altri due autovalori di \mathbb{A} , saranno¹² $\lambda_1 = e^{i\alpha}$, e $\lambda_2 = e^{-i\alpha}$, con $\alpha \in \mathbb{R}$. Abbiamo quindi¹³

$$\text{tr} \mathbb{A} = \sum_{i=1}^3 \lambda_i = 1 + 2 \cos \alpha. \quad (6.13)$$

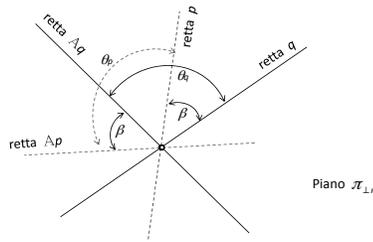
¹¹Riferendoci alla (6.6), $\text{tr} \mathbb{A} = \alpha_1 + \beta_2 + \gamma_3$.

¹²Ricordiamo che la formula di Eulero per i numeri complessi di modulo unitario afferma che, per ogni numero reale α , si ha

$$e^{i\alpha} = \cos \alpha + i \sin \alpha.$$

Osserviamo anche che il caso $\lambda_2 = \lambda_1 = -1$, corrisponde a $\alpha = \pi$.

¹³Si ricorda che gli autovalori, la traccia ed il determinante della matrice \mathbb{A} , risultano indipendenti dalla base. Se infatti avessimo scelto un diverso SdR fisso Σ' ed un diverso SdR mobile S' , avremmo ottenuto una matrice \mathbb{A}' diversa da \mathbb{A} (ma ad essa collegata tramite una trasformazione di similitudine). Tuttavia gli autovalori di \mathbb{A}' sono gli stessi di \mathbb{A} , ed inoltre $\det \mathbb{A} = \det \mathbb{A}'$, $\text{tr} \mathbb{A} = \text{tr} \mathbb{A}'$.



$$\theta_p + \beta = \theta_q + \beta, \quad \Rightarrow \quad \theta_p = \theta_q$$

Figura 6.3: L'angolo fra due rette viene conservato.

Se adesso si sceglie come SdR fisso quello in cui la direzione della retta invariante r coincide con l'asse ζ , la matrice \mathbb{A} è data da

$$\mathbb{A} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.14}$$

essendo θ , l'angolo di cui la coppia di vettori \hat{i}, \hat{j} è ruotata rispetto ai vettori e_1, e_2 . Calcolando la traccia di \mathbb{A} data da (6.14), e confrontando con (6.13), otteniamo

$$\alpha = \theta, \tag{6.15}$$

e quindi α va identificato con l'angolo di rotazione. Osserviamo però che per **definire univocamente l'angolo di rotazione θ è necessario definire una convenzione**. Noi adotteremo la seguente: *scelto il verso positivo della retta invariante r come quello di \mathbf{k} , se la rotazione avviene in senso antiorario rispetto al verso di \mathbf{k} , allora θ è positivo, altrimenti è negativo*.

Va infine osservato che l'uguaglianza (6.15), data l'invarianza della traccia, assume carattere generale e non si riferisce soltanto al caso in cui la retta r è parallela a \mathbf{k} .

Nota 6.2.1 *In definitiva il teorema di Eulero dà anche la regola per passare, mediante semplici operazioni, dal SdR Σ al SdR S . Lavorando nel SdR Σ , si individua la retta r (autodirezione relativa all'autovalore $\lambda = 1$) ed un qualsiasi piano $\pi_{\perp r}$. Poi, selezionato un qualsiasi punto $P \in \pi_{\perp r}$, si ruota P attorno ad r dell'angolo α (che essenzialmente si trova calcolando $\text{tr} \mathbb{A}$). In questo movimento P trascina con sé il sistema rigido e quindi anche gli assi cartesiani. Al termine della rotazione il SdR ottenuto è proprio S .*

6.2.2 Secondo caso: $\Omega \neq O$

Consideriamo ora il caso particolare in cui¹⁴ O , origine del SdR S , e Ω origine del SdR Σ , non sono coincidenti. Il generico punto $P \in \mathfrak{B}$, è individuato, rispetto al SdR S dal vettore $(P - O)$, la cui espressione è data dalla terna \boldsymbol{x} . Lo stesso punto P viene individuato, rispetto al SdR Σ dal vettore $(P - \Omega)$, la cui espressione, sempre relativa a Σ , è data dalla terna $\boldsymbol{\xi}$. Consideriamo poi il SdR Σ' , i cui assi si mantengono sempre paralleli a quelli di Σ , e la cui origine coincide con O (v. figura 6.4). Denotiamo con $\boldsymbol{\xi}_P$ la terna che rappresenta il vettore $(P - O)$ rispetto al SdR Σ' . La relazione che lega $\boldsymbol{\xi}_P$ e \boldsymbol{x} sarà sempre

$$\boldsymbol{\xi}_P = \mathbb{A}\boldsymbol{x}, \quad \Leftrightarrow \quad \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

dove \mathbb{A} è la matrice di rotazione di cui abbiamo parlato nella precedente sezione. Quindi, la relazione vettoriale

$$(P - \Omega) = (P - O) + (O - \Omega), \quad (6.16)$$

si traduce nella seguente relazione fra terne

$$\boldsymbol{\xi} = \boldsymbol{\xi}_O + \boldsymbol{\xi}_P = \boldsymbol{\xi}_O + \mathbb{A}\boldsymbol{x}, \quad (6.17)$$

dove $\boldsymbol{\xi}_O = \begin{pmatrix} \xi_O \\ \eta_O \\ \zeta_O \end{pmatrix}$, è la terna che rappresenta il vettore $(O - \Omega)$ nel SdR Σ . Osserviamo che nella (6.17) $\boldsymbol{\xi}_O = \boldsymbol{\xi}_O(t)$, come $\mathbb{A} = \mathbb{A}(t)$, nel senso della (6.7), mentre \boldsymbol{x} non dipende dal tempo (essendo la terna che, nel SdR S solidale col corpo rigido, rappresenta il vettore $(P - O)$).

6.2.3 Gradi di libertà di un corpo rigido

Dato un corpo rigido \mathfrak{B} , col termine *gradi di libertà intendiamo il numero di variabili indipendenti necessarie per determinare univocamente la sua posizione nello spazio* (si tratta quindi del numero di gradi libertà anche secondo la definizione data nella sezione 4.1). Quindi, siccome è nota la struttura geometrica di \mathfrak{B} , ovvero sono note, rispetto ad un SdR solidale, tutte le coordinate \boldsymbol{x} dei punti che lo compongono, dalla (6.17) si deduce che determinare la posizione di un qualsiasi punto di \mathfrak{B} rispetto ad un SdR fisso Σ è necessario:

- conoscere la posizione di almeno un punto O solidale con \mathfrak{B} ; ovvero la terna $\boldsymbol{\xi}_O$, e questi sono **tre parametri indipendenti**;
- conoscere la rotazione descritta dalla matrice \mathbb{A} .

Per quanto riguarda la matrice \mathbb{A} , a priori sembrerebbe che fossero necessari 9 parametri per determinarla. Tuttavia, è ben noto che per descrivere una rotazione sono sufficienti solo tre parametri: l'angolo di rotazione (un parametro); e l'asse di rotazione (due parametri). Quest'ultimo è determinato infatti dal versore che ne dà la direzione. Abbiamo infatti il seguente teorema

¹⁴Si ricorda che O è solidale con il corpo rigido \mathfrak{B} .

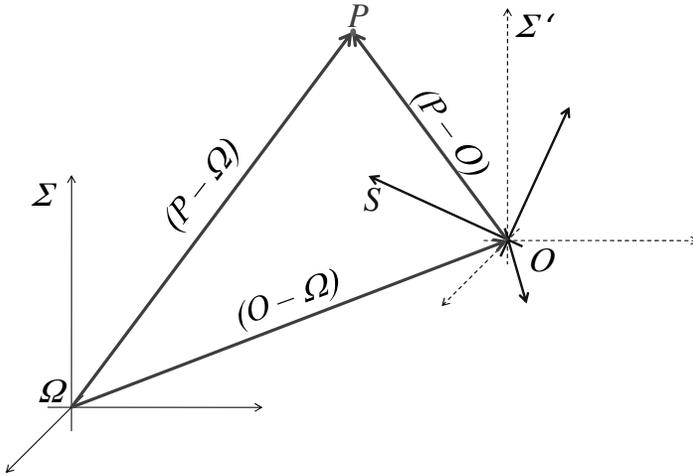


Figura 6.4: Secondo caso: $\Omega \neq O$. Il SdR Σ' ha l'origine in O , ma mantiene gli assi costantemente paralleli a quelli di Σ .

Teorema 6.2.3 *La matrice \mathbb{A} può essere espressa in termini di soli tre parametri indipendenti.*

Dim. Le 9 componenti della matrice di rotazione \mathbb{A} si trovano esplicitate nella formula (6.6). Di esse però non tutte sono indipendenti: ovvero ben 6 possono essere espresse in termini di soli tre parametri indipendenti. Infatti \mathbb{A} deve soddisfare la relazione (6.5), che quindi rappresenta un **vincolo** per le 9 componenti. In pratica, la (6.9) è un insieme di sei vincoli indipendenti (si veda la sezione 4.1 del capitolo 4) che legano fra loro le componenti $\alpha_i, \beta_i, \gamma_i, i = 1, 2, 3$. Tali equazioni consentono di esprimere sei coefficienti in termini di soli tre parametri indipendenti (che, come visto nella sezione 4.1, sono detti parametri lagrangiani). Per semplicità svolgiamo la dimostrazione nel caso in cui \mathbb{A} sia una matrice 2×2 , cioè nel caso in cui si abbia a che fare con un corpo rigido piano che si muove sul piano. In questo semplice caso

$$\mathbb{A} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix},$$

dove $-1 \leq \alpha_i \leq 1, i = 1, 2$, e $-1 \leq \beta_i \leq 1, i = 1, 2$. Ovviamente

$$\det \mathbb{A} = \alpha_1\beta_2 - \alpha_2\beta_1 > 0, \tag{6.18}$$

mentre la condizione (6.5) si traduce in

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1^2 + \beta_1^2 & \alpha_1\alpha_2 + \beta_1\beta_2 \\ \alpha_1\alpha_2 + \beta_1\beta_2 & \alpha_2^2 + \beta_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

ovvero nelle seguenti 3 equazioni che i coefficienti $\alpha_1, \alpha_2, \beta_1$ e β_2 devono soddisfare

$$\begin{cases} \alpha_1^2 + \beta_1^2 = 1, \\ \alpha_1\alpha_2 + \beta_1\beta_2 = 0, \\ \alpha_2^2 + \beta_2^2 = 1. \end{cases} \quad (6.19)$$

Ora, applicando il teorema delle funzioni implicite al sistema (6.19) possiamo stabilire quando è possibile utilizzare una sola variabile per esplicitare le altre 3 (ovvero quando le tre equazioni di (6.19) sono **indipendenti**). Leggendo (6.19) come

$$\begin{cases} F_1(\alpha_1, \alpha_2, \beta_1, \beta_2) = 0, \\ F_2(\alpha_1, \alpha_2, \beta_1, \beta_2) = 0, \\ F_3(\alpha_1, \alpha_2, \beta_1, \beta_2) = 0, \end{cases}$$

se

$$\begin{pmatrix} \frac{\partial F_1}{\partial \alpha_1} & \frac{\partial F_1}{\partial \beta_1} & \frac{\partial F_1}{\partial \alpha_2} & \frac{\partial F_1}{\partial \beta_2} \\ \frac{\partial F_2}{\partial \alpha_1} & \frac{\partial F_2}{\partial \beta_1} & \frac{\partial F_2}{\partial \alpha_2} & \frac{\partial F_2}{\partial \beta_2} \\ \frac{\partial F_3}{\partial \alpha_1} & \frac{\partial F_3}{\partial \beta_1} & \frac{\partial F_3}{\partial \alpha_2} & \frac{\partial F_3}{\partial \beta_2} \end{pmatrix} = \begin{pmatrix} 2\alpha_1 & 2\beta_1 & 0 & 0 \\ \alpha_2 & \beta_2 & \alpha_1 & \beta_1 \\ 0 & 0 & 2\alpha_2 & 2\beta_2 \end{pmatrix}, \quad (6.20)$$

ha rango massimo, cioè 3, allora (almeno localmente) sarà possibile esplicitare tre variabili in funzione della quarta. Se consideriamo¹⁵ $\alpha_2 \neq 0$, il determinante delle prime tre colonne di (6.20) dà

$$\det \begin{pmatrix} 2\alpha_1 & 2\beta_1 & 0 \\ \alpha_2 & \beta_2 & \alpha_1 \\ 0 & 0 & 2\alpha_2 \end{pmatrix} = 4\alpha_2 \det \mathbb{A} \neq 0. \quad (6.18)$$

Infatti, possiamo riscrivere (6.19) come

$$\begin{cases} \beta_1^2 = 1 - \alpha_1^2, \\ \alpha_2^2 \alpha_1^2 = \beta_1^2 \beta_2^2, \\ \beta_2^2 = 1 - \alpha_2^2, \end{cases} \Rightarrow \begin{cases} \beta_1^2 = 1 - \alpha_1^2, \\ \alpha_2^2 = 1 - \alpha_1^2, \\ \beta_2^2 = \alpha_1^2, \end{cases} \quad (6.21)$$

e quindi usare α_1 come parametro indipendente ed esprimere gli altri 3 in funzione di esso. Di fatto, al posto di α_1 è molto più pratico usare l'angolo di rotazione θ , $\theta \in [0, 2\pi)$, come variabile (o parametro lagrangiano) tramite cui esprimere $\alpha_i, \beta_i, i = 1, 2$. Utilizzando dunque l'angolo θ , otteniamo $\alpha_1 = \cos \theta$, $\beta_1 = \sin \theta$, $\alpha_2 = -\sin \theta$, e $\beta_2 = \cos \theta$, e quindi

$$\mathbb{A} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (6.22)$$

Nel caso in cui \mathbb{A} sia una matrice 3×3 , si procede nello stesso modo (adesso i calcoli sono decisamente più complicati). Si individuano tre parametri indipendenti che, come

¹⁵Il caso $\alpha_2 = 0$ implica, in virtù del sistema (6.19),

$$\mathbb{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

visto nel capitolo 4, sono tre parametri lagrangiani e si esprimono tutti i coefficienti $\alpha_i, \beta_i, \gamma_i, i = 1, 2, 3$, in funzione di essi. Di fatto, cioè, non si adoperano tre dei nove coefficienti $\alpha_i, \beta_i, \gamma_i, i = 1, 2, 3$, in quanto, come nella (6.21), i legami funzionali sono di tipo quadratico e quindi difficilmente utilizzabili. Si preferisce quindi usare altri tre parametri, i cosiddetti **angoli di Eulero** (che saranno illustrati nella sezione 6.4), ed esprimere tramite essi i coefficienti della matrice \mathbb{A} .

□

In sostanza quanto fino ad ora detto ci porta a concludere che **un sistema un sistema rigido che si muove nello spazio ha 6 gradi di libertà**: le tre coordinate dell'origine O del sistema mobile; ed i tre parametri indipendenti tramite cui si definisce la matrice \mathbb{A} . Infatti, come abbiamo visto, le nove componenti della matrice \mathbb{A} sono vincolate a “muoversi” sulla sottovarietà determinata dalle equazioni (6.5) (o meglio nella componente connessa che soddisfa anche alla disequazione $\det \mathbb{A} > 0$). Questo insieme, che ha la struttura di gruppo rispetto al prodotto tra matrici, si indica con la sigla¹⁶ $SO(3)$,

$$SO(3) = \left\{ \mathbb{M} \in \mathbb{R}^{3 \times 3} \mid \mathbb{M}\mathbb{M}^T = \mathbb{I}, \det \mathbb{M} = +1 \right\}.$$

In definitiva, l'insieme delle possibili “configurazioni” che un sistema di riferimento “mobile” può assumere rispetto a un sistema di riferimento “fisso” è parametrizzato dai punti dell'insieme $\mathbb{R}^3 \times SO(3)$, che risulta essere una varietà differenziabile di dimensione 6.

Nota 6.2.2 *Il prodotto tra matrici non è commutativo. Infatti, date due matrici \mathbb{A} e \mathbb{B} , in generale $\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$. Questo è vero anche per due matrici generiche di $SO(3)$. Come si interpreta ciò in termini di moti rigidi?*

6.3 Formula fondamentale del moto rigido

La velocità di un corpo rigido varia, in generale, da punto a punto. Per tal motivo si parla di **campo di velocità** o **atto di moto**. L'atto di moto è infatti definito come *un'applicazione che ad ogni punto del corpo associa la sua velocità*. Rifacendoci a quanto detto nell'introduzione, la *velocità delle particelle di \mathfrak{B} è quella misurata dall'osservatore solidale con il SdR fisso Σ centrato in Ω* . Quindi, considerando la (6.17) e derivandola abbiamo

$$\frac{d}{dt} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \frac{d}{dt} \begin{pmatrix} \xi_O \\ \eta_O \\ \zeta_O \end{pmatrix} + \frac{d}{dt} \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{6.23}$$

dove

$$\frac{d}{dt} \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} = \begin{pmatrix} \dot{\alpha}_1(t) & \dot{\alpha}_2(t) & \dot{\alpha}_3(t) \\ \dot{\beta}_1(t) & \dot{\beta}_2(t) & \dot{\beta}_3(t) \\ \dot{\gamma}_1(t) & \dot{\gamma}_2(t) & \dot{\gamma}_3(t) \end{pmatrix},$$

¹⁶Il numero 3 sta a indicare la dimensione della matrice (e quindi dello spazio su cui agisce come matrice di trasformazione), O sta per “ortogonale” e significa che le matrici devono soddisfare la (6.5) e infine S sta per “speciale” indicando che si prendono solo le matrici con determinante positivo.

e dove $\frac{d\mathbf{x}}{dt} = 0$, dal momento che \mathbf{x} rappresenta la terna delle coordinate del punto P rispetto al SdR solidale S (che quindi non variano nel tempo).

Rimarchiamo ancora che la derivata a primo membro della (6.23) è la velocità del punto P misurata da Σ , e ugualmente la derivata di $(\xi_O, \eta_O, \zeta_O)^T$ è la velocità del punto O rispetto al SdR Σ .

Possiamo ora riscrivere la relazione (6.23) inserendo tra $\frac{d\mathbb{A}}{dt}$ e il vettore $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ la matrice identità scritta nella forma $\mathbb{A}^T \mathbb{A}$, cosicché abbiamo

$$\frac{d}{dt} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \frac{d}{dt} \begin{pmatrix} \xi_O \\ \eta_O \\ \zeta_O \end{pmatrix} + \left(\frac{d\mathbb{A}}{dt} \mathbb{A}^T \right) \underbrace{\mathbb{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\xi_P}. \quad (6.24)$$

Ricordiamo infatti che $\xi_P = \mathbb{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ è la terna delle componenti del vettore geometrico $(P - O)$ espresse rispetto al SdR Σ (ovvero le differenze tra le coordinate di P e quelle di O calcolate da Σ). Osserviamo inoltre che se si sceglie un punto diverso da O come origine del riferimento solidale (mantenendo però gli stessi versori $\mathbf{i}, \mathbf{j}, \mathbf{k}$) la matrice $\frac{d\mathbb{A}}{dt} \mathbb{A}^T$ che compare nella formula non cambia. Possiamo quindi dire che la distribuzione di velocità di tutti i punti di un sistema rigido si ottiene dalla formula (6.24) quando si conosca la velocità di un punto solidale e la matrice $\frac{d\mathbb{A}}{dt} \mathbb{A}^T$.

Osserviamo infine una caratteristica fondamentale della matrice $\frac{d\mathbb{A}}{dt} \mathbb{A}^T$. Derivando la relazione (6.5) si ha

$$\frac{d\mathbb{A}}{dt} \mathbb{A}^T + \mathbb{A} \frac{d\mathbb{A}^T}{dt} = \frac{d\mathbb{A}}{dt} \mathbb{A}^T + \mathbb{A} \left(\frac{d\mathbb{A}}{dt} \right)^T = 0. \quad (6.25)$$

Ricordando che $\mathbb{A} \mathbb{B}^T = (\mathbb{B} \mathbb{A}^T)^T$, possiamo riscrivere la (6.25) come

$$\frac{d\mathbb{A}}{dt} \mathbb{A}^T = - \left(\frac{d\mathbb{A}}{dt} \mathbb{A} \right)^T, \quad (6.26)$$

ovvero **la matrice $\frac{d\mathbb{A}}{dt} \mathbb{A}^T$ è una matrice antisimmetrica**. Indichiamola con

$$\frac{d\mathbb{A}}{dt} \mathbb{A}^T = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \quad (6.27)$$

e riscriviamo la (6.24) come

$$\begin{pmatrix} \dot{\xi} \\ \dot{\eta} \\ \dot{\zeta} \end{pmatrix} = \begin{pmatrix} \dot{\xi}_O \\ \dot{\eta}_O \\ \dot{\zeta}_O \end{pmatrix} + \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \begin{pmatrix} \xi_P \\ \eta_P \\ \zeta_P \end{pmatrix}. \quad (6.28)$$

Se adesso **estriamo dalla matrice** (6.27) **la terna**¹⁷

$$\underline{\omega} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}, \quad (6.29)$$

e ricordiamo la definizione del prodotto vettoriale fra terne (1.15), è immediato verificare che il risultato del prodotto tra la matrice antisimmetrica $\frac{d\mathbb{A}}{dt} \mathbb{A}^T$ e la terna ξ_p nella (6.28) è identico al prodotto vettoriale fra le terne $\underline{\omega}$ e ξ_p , cioè

$$\left(\frac{d\mathbb{A}}{dt} \mathbb{A}^T \right) \xi_p = \underline{\omega} \wedge \xi_p. \quad (6.30)$$

Possiamo quindi riscrivere la (6.28) così:

$$\dot{\xi} = \dot{\xi}_O + \underline{\omega} \wedge \xi_p. \quad (6.31)$$

A questo punto però vengono spontanee le seguenti domande¹⁸:

- i. La terna $\underline{\omega}$, o meglio le componenti della “colonna” (6.29) rappresentano veramente le componenti di un vettore? In altre parole, esiste un vettore ω , le cui componenti sono date dalla terna $\underline{\omega}$? E, nel caso in cui il vettore ω esistesse, gli elementi della terna (6.29) sono le componenti del vettore ω rispetto a quale SdR?
- ii. Ammesso che esista il vettore ω , questo è intrinseco all’atto di moto oppure dipende dai SdR Σ , S , e dalle origini Ω , O ?

Se ω esiste ed è un vettore che dipende soltanto dal moto del sistema rigido, allora (6.31) può essere scritta nella notazione “vettoriale” (la cui forma cioè è indipendente dal SdR¹⁹)

$$\frac{d(P - \Omega)}{dt} = \frac{d(O - \Omega)}{dt} + \omega \wedge (P - O), \quad (6.32)$$

detta **formula fondamentale del moto rigido**. Il vettore ω è detto **velocità angolare del corpo rigido**, o anche **velocità angolare istantanea del corpo rigido**.

Teorema 6.3.1 *La velocità di un punto P del sistema rigido \mathfrak{B} rispetto al SdR Σ è data dalla formula (6.32), dove ω è un vettore (o meglio uno “pseudovettore”), le cui componenti nel SdR Σ vengono estratte dalla matrice (6.27) secondo la (6.29). Il vettore ω può dipendere, in generale, dal tempo, ma è univocamente determinato dal moto del corpo \mathfrak{B} rispetto a Σ .*

¹⁷Sottolineiamo ancora che $\underline{\omega}$ denota semplicemente tre numeri messi in colonna. Questo aspetto delicato verrà chiarito nel seguito.

¹⁸Si veda anche: Giuseppe Benfatto, *Note del Corso di Meccanica Analitica*, Università di Roma “Tor Vergata”, a.a. 2007-08.

¹⁹La (6.32) non è una relazione vettoriale in senso stretto. Essa infatti mantiene la sua forma solo per cambiamenti di coordinate ortogonali (cioè se solo se si passa da un sistema ortonormale a un altro sempre ortonormale); questo perché sia la definizione di ω sia quella di prodotto vettoriale fanno intervenire la struttura metrica dello spazio euclideo, ovvero la definizione di ortogonalità, e l’orientazione, cioè l’ordine in cui si considerano i versori della terna. Spesso, per essere più corretti, si usa l’espressione “pseudovettoriale” per riferirsi a relazioni che si conservano solo per trasformazioni ortogonali con determinante positivo.

Dim. La dimostrazione di questo teorema fondamentale verrà svolta nei seguenti passi:

1. Dapprima mostriamo che i tre elementi della (6.29) rappresentano le componenti, rispetto al SdR Σ , di un vettore che appunto abbiamo indicato con ω . Facciamo poi vedere che tale vettore non dipende né da Σ né dal centro Ω .
2. Proviamo poi che ω è indipendente sia dal SdR S che da O .

Passo 1. Ricordiamo che una terna ordinata di \mathbb{R}^3 rappresenta effettivamente le componenti di un vettore se, nel cambiamento da un SdR ad un altro, soddisfa una ben precisa regola di trasformazione. In sostanza se \mathbf{u} e \mathbf{u}' sono due terne che rappresentano rispettivamente nei SdR Σ e Σ' lo stesso vettore allora

$$\mathbf{u}' = \mathbb{Q}\mathbf{u}, \quad (6.33)$$

dove \mathbb{Q} è la matrice del cambio di base da Σ a Σ' . Viceversa se vale la (6.33) allora le terne \mathbf{u} e \mathbf{u}' sono, rispettivamente nel SdR Σ e nel SdR Σ' , le terne delle componenti del medesimo vettore²⁰. Quindi per dimostrare che le componenti della (6.29) sono effettivamente le componenti di un vettore dobbiamo provare che $\underline{\omega}$ soddisfa la condizione di trasformazione (6.33) quando si considera un diverso SdR. In quest'ottica consideriamo, come in figura 6.5, un altro **SdR fisso** Σ' **centrato in** $\Omega' \neq \Omega$. Sia \mathbb{Q} la matrice del cambiamento di riferimento da Σ a Σ' . Abbiamo

$$P - \Omega' = (\Omega - \Omega') + (P - \Omega) = (\Omega - \Omega') + (O - \Omega) + (P - O). \quad (6.34)$$

Se denotiamo con ξ' la terna delle componenti del vettore $(P - \Omega')$ rispetto al SdR Σ' , questa può essere espressa in due modi equivalenti

$$\xi' = \begin{cases} \xi'_\Omega + \mathbb{Q}(\xi_O + \underbrace{\mathbb{A}\mathbf{x}}_{\xi_p}) = \xi'_\Omega + \mathbb{Q}\xi_O + \mathbb{Q}\xi_p, \\ \xi'_\Omega + \mathbb{Q}\xi_O + \xi'_p, \end{cases} \quad (6.35)$$

dove ξ'_p è la terna delle componenti del vettore $(P - O)$ rispetto a Σ' , mentre, lo stesso vettore $(P - O)$ è rappresentato dalla terna ξ_p nel SdR Σ . Di conseguenza, le terne ξ'_Ω e ξ_p sono collegate dalla relazione

$$\xi'_p = \mathbb{Q}\xi_p. \quad (6.36)$$

Osserviamo poi che, se \mathbf{x} è la terna delle componenti del vettore $(P - O)$ rispetto al SdR solidale S , le terne ξ'_p e ξ_p sono legate ad \mathbf{x} tramite

$$\xi'_p = \mathbb{B}\mathbf{x}, \quad \xi_p = \mathbb{A}\mathbf{x}, \quad (6.37)$$

dove \mathbb{B} è la matrice di rotazione di S rispetto al SdR fisso Σ' , mentre, come sappiamo, \mathbb{A} è la matrice di rotazione fra S e Σ . Infine, si rimarca che nella (6.35) né \mathbb{Q} né ξ'_Ω (terna delle componenti del vettore $(\Omega - \Omega')$ rispetto al SdR Σ') dipendono dal tempo.

²⁰In caso contrario, cioè se $\underline{\omega}$ non rappresentasse alcun vettore, ci ritroveremmo con importanti limitazioni all'uso della (6.31). Per esempio, la somma di $\underline{\omega}_1$ e $\underline{\omega}_2$, ricavate in due riferimenti diversi, non risulterebbe definita.

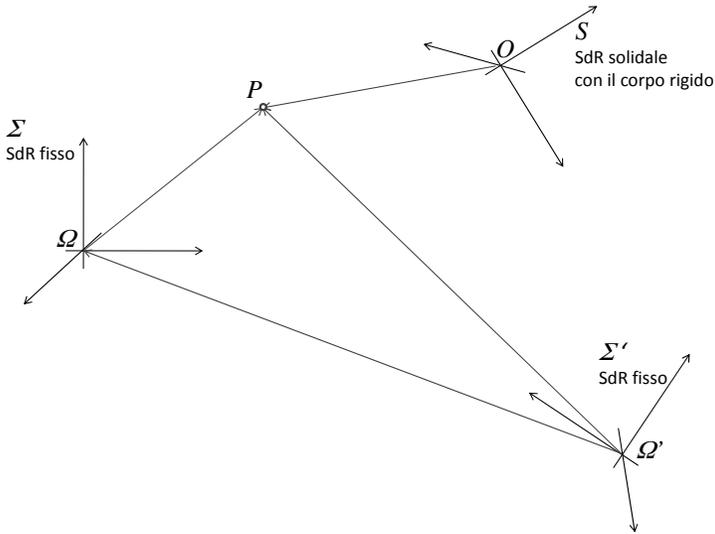


Figura 6.5: Il primo SdR fisso Σ è centrato in Ω , il secondo SdR fisso Σ' centrato in $\Omega' \neq \Omega$.

Calcoliamo adesso la velocità di P rispetto a Σ' derivando la (6.35). Evidentemente $\frac{d\xi'_\Omega}{dt} = 0$, e $\frac{dQ}{dt} = 0$, in quanto Σ e Σ' sono fissi. Abbiamo dunque

$$\frac{d\xi'}{dt} = \begin{cases} Q \frac{d\xi_O}{dt} + Q \frac{d\xi_p}{dt} = Q \frac{d\xi_O}{dt} + Q (\underline{\omega} \wedge \xi_p) , \\ Q \frac{d\xi_O}{dt} + \frac{d\xi'_p}{dt} \stackrel{(6.37)}{=} Q \frac{d\xi_O}{dt} + \underbrace{\left(\frac{d\mathbb{B}}{dt} \mathbb{B}^T \right)}_{\underline{\omega}' \wedge \xi'_p} \underbrace{(\mathbb{B}x)}_{\xi'_p} = Q \frac{d\xi_O}{dt} + \underline{\omega}' \wedge \underbrace{Q \xi_p}_{\xi'_p} , \end{cases}$$

dove $\underline{\omega}'$ è la “colonna” associata alla matrice antisimmetrica $\frac{d\mathbb{B}}{dt} \mathbb{B}^T$. Confrontando adesso le due espressioni otteniamo

$$Q (\underline{\omega} \wedge \xi_p) = \underline{\omega}' \wedge Q \xi_p .$$

Ora se $Q \in SO(3)$, cioè se Q è una matrice ortogonale con determinante uguale a +1,

vale la seguente uguaglianza²¹

$$\mathbb{Q}(\underline{\omega} \wedge \underline{\xi}_p) = \mathbb{Q}\underline{\omega} \wedge \mathbb{Q}\underline{\xi}_p, \quad (6.38)$$

da cui deduciamo

$$\underline{\omega}' = \mathbb{Q}\underline{\omega}, \quad (6.39)$$

ovvero che $\underline{\omega}'$ e $\underline{\omega}$ trasformano come specificato nella (6.33). Possiamo quindi affermare che $\underline{\omega}'$ e $\underline{\omega}$ sono le terne che rappresentano le componenti del vettore ω nei SdR Σ' e Σ , , rispettivamente. *Di fatto la prova dell'esistenza del vettore ω si basa sulla (6.38), che però non è vera per tutte le trasformazioni invertibili, ma soltanto per quelle di $SO(3)$. Per tal motivo non possiamo propriamente affermare che le terne $\underline{\omega}'$ e $\underline{\omega}$ rappresentano le componenti di un vettore ω . Per esser più precisi infatti parliamo di "pseudovettore" ω .*

Inoltre, il fatto che $\underline{\omega}$ si trasformi secondo la (6.33) implica che il vettore ω non dipende dal particolare SdR fisso (purché ortogonale) né dal punto dove lo stesso è centrato. Infatti considerando un SdR fisso Σ' , diverso da Σ , abbiamo ottenuto una terna $\underline{\omega}'$ che è collegata ad $\underline{\omega}$ dalla (6.39). Ciò vuol dire che le due terne sono le **componenti dello stesso vettore**. Quest'ultimo dunque non varia se si passa dal SdR Σ centrato in O , al SdR Σ' centrato in O' .

Resta da provare che ω non dipende dal SdR S , né dal centro O . Questo verrà fatto nel passo successivo.

Passo 2. Mostriamo per prima cosa che ω non dipende dal SdR solidale. Faremo tale dimostrazione provando che $\underline{\omega}$ non dipende dal SdR solidale. Consideriamo infatti un altro riferimento solidale S' , diverso da S , sempre centrato in O . Indichiamo con \mathbf{x} e \mathbf{x}' le terne delle componenti di $(P - O)$ nei SdR S e S' . Quindi, ricordando la (6.17), scriveremo

$$\underline{\xi}(t) = \underline{\xi}_O(t) + \underline{\xi}_p(t), \quad \text{dove} \quad \underline{\xi}_p(t) = \begin{cases} \mathbb{A}(t)\mathbf{x}, & \text{SdR } S, \\ \mathbb{C}(t)\mathbf{x}', & \text{SdR } S', \end{cases} \quad (6.40)$$

dove \mathbb{C} è la matrice di rotazione tra il SdR fisso Σ e il SdR solidale S' . Derivando rispetto al tempo la (6.40) abbiamo

$$\frac{d\underline{\xi}}{dt} = \frac{d\underline{\xi}_O}{dt} + \frac{d\underline{\xi}_p}{dt}, \quad \text{dove} \quad \frac{d\underline{\xi}_p}{dt} = \begin{cases} \frac{d\mathbb{A}}{dt}\mathbf{x} = \left(\frac{d\mathbb{A}}{dt}\mathbb{A}^T\right) \underbrace{\mathbb{A}\mathbf{x}}_{\underline{\xi}_p}, & \text{SdR } S, \\ \frac{d\mathbb{C}}{dt}\mathbf{x}' = \left(\frac{d\mathbb{C}}{dt}\mathbb{C}^T\right) \underbrace{\mathbb{C}\mathbf{x}'}_{\underline{\xi}_p}, & \text{SdR } S'. \end{cases}$$

²¹Questo è un punto delicato, infatti la (6.38) vale soltanto per le matrici di $SO(3)$. In generale, se \mathbb{F} è una matrice invertibile, abbiamo

$$\mathbb{F}\mathbf{u} \wedge \mathbb{F}\mathbf{w} = (\det \mathbb{F}) (\mathbb{F}^{-1})^T (\mathbf{u} \wedge \mathbf{w}).$$

E' chiaro che tale formula si riduce alla (6.38) nel caso in cui $\mathbb{F} \in SO(3)$, dal momento che $\det \mathbb{F} = 1$ e $(\mathbb{F}^{-1})^T = \mathbb{F}$.

Dal confronto si ottiene $\frac{d\mathbb{A}}{dt}\mathbb{A}^T = \frac{d\mathbb{C}}{dt}\mathbb{C}^T$, e quindi rammentando la (6.27),

$$\frac{d\mathbb{C}}{dt}\mathbb{C}^T = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

Si conclude pertanto che il vettore ω non dipende dal SdR solidale dal momento che due SdR solidali distinti hanno dato origine alla stessa terna $\underline{\omega}$.

Terminiamo la dimostrazione con la verifica che il vettore ω è indipendente dal punto solidale O . Lavoriamo direttamente con la formula vettoriale (6.32), considerando un altro punto solidale O' , diverso da O , e su di esso centriamo un SdR solidale, come mostrato in figura 6.6.

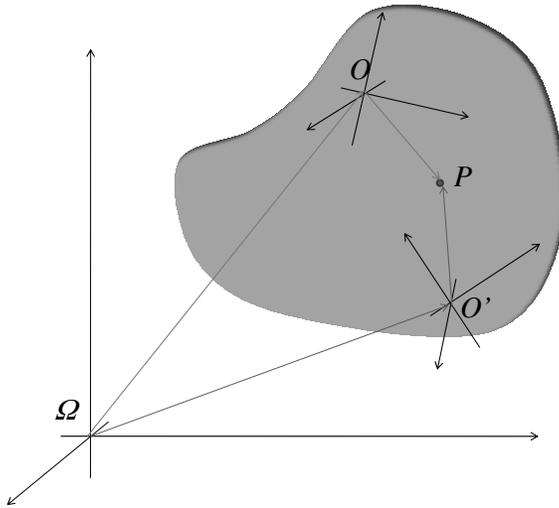


Figura 6.6: Si considerano due SdR solidali. Il primo centrato in O , il secondo centrato in $O' \neq O$.

Se indichiamo con ω' la velocità angolare relativa alla scelta di O' , vogliamo provare che $\omega' = \omega$. A tal fine consideriamo dapprima il centro O scrivendo

$$\frac{d(P - \Omega)}{dt} = \frac{d(O - \Omega)}{dt} + \omega \wedge (P - O). \tag{6.41}$$

Se invece consideriamo il centro O' , scriveremo

$$\frac{d(P - \Omega)}{dt} = \frac{d(O' - \Omega)}{dt} + \omega' \wedge (P - O').$$

Ora O' è solidale con il corpo rigido e dunque

$$\frac{d(O' - \Omega)}{dt} = \frac{d(O - \Omega)}{dt} + \omega \wedge (O' - O),$$

mentre

$$(P - O) = (O' - O) + (P - O').$$

Quindi

$$\frac{d(P - O)}{dt} = \begin{cases} \frac{d(O - \Omega)}{dt} + \boldsymbol{\omega} \wedge (P - O) & = \frac{d(O - \Omega)}{dt} + \boldsymbol{\omega} \wedge (O' - O) \\ & + \boldsymbol{\omega} \wedge (P - O'), \\ \frac{d(O' - \Omega)}{dt} + \boldsymbol{\omega}' \wedge (P - O') & = \frac{d(O - \Omega)}{dt} + \boldsymbol{\omega} \wedge (O' - O) \\ & + \boldsymbol{\omega}' \wedge (P - O'). \end{cases}$$

Per confronto si ottiene

$$(\boldsymbol{\omega} - \boldsymbol{\omega}') \wedge (P - O') = 0,$$

che, dovendo valere per ogni punto P , implica $\boldsymbol{\omega} - \boldsymbol{\omega}' = 0$, ovvero il vettore $\boldsymbol{\omega}$ non dipende neanche dal centro O del SdR solido.

□

Nota 6.3.1 Osserviamo che la formula (6.50) implica che la distanza fra i punti del sistema rigido rimane invariata. Infatti, se P_1 e P_2 sono due punti del sistema rigido

$$\begin{aligned} \frac{d(P_1 - P_2)^2}{dt} &= 2(P_1 - P_2) \cdot \frac{d}{dt} [(P_1 - O) - (P_2 - O)] = \\ &= 2(P_1 - P_2) \cdot [\boldsymbol{\omega} \wedge (P_1 - O) - \boldsymbol{\omega} \wedge (P_2 - O)] = \\ &= 2(P_1 - P_2) \cdot [\boldsymbol{\omega} \wedge (P_1 - P_2)] = 0. \end{aligned}$$

Nota 6.3.2 Vale la pena anche illustrare brevemente un approccio, diverso da quello seguito, che porta ad introdurre il vettore $\boldsymbol{\omega}$. Tale approccio ha il pregio di condurre direttamente al vettore $\boldsymbol{\omega}$, evitando il passaggio dalle terne, ma, a nostro avviso, ha il difetto di non mettere subito in evidenza due importanti fatti: il primo è che, in realtà, $\boldsymbol{\omega}$ è uno pseudovettore mentre il secondo è la stretta relazione fra $\boldsymbol{\omega}$ e la matrice \mathbb{A} . Vediamo comunque come si procede secondo questa seconda via. Si lavora nel SdR S e qui si osserva che

$$\mathbf{i} \cdot \mathbf{i} = 1, \quad \mathbf{j} \cdot \mathbf{j} = 1, \quad \mathbf{k} \cdot \mathbf{k} = 1, \quad (6.42)$$

$$\mathbf{i} \cdot \mathbf{j} = 0, \quad \mathbf{i} \cdot \mathbf{k} = 0, \quad \mathbf{j} \cdot \mathbf{k} = 0. \quad (6.43)$$

Se adesso consideriamo i versori \mathbf{i} , \mathbf{j} , \mathbf{k} , nel SdR Σ , questi variano nel tempo essendo solidali col corpo in movimento. Derivando rispetto al tempo le relazioni di (6.42) otteniamo

$$\frac{d\mathbf{i}}{dt} \cdot \mathbf{i} = 0, \quad \implies \quad \frac{d\mathbf{i}}{dt} = \boldsymbol{\omega}_1 \wedge \mathbf{i}, \quad (6.44)$$

$$\frac{d\mathbf{j}}{dt} \cdot \mathbf{j} = 0, \quad \implies \quad \frac{d\mathbf{j}}{dt} = \boldsymbol{\omega}_2 \wedge \mathbf{j}, \quad (6.45)$$

$$\frac{d\mathbf{k}}{dt} \cdot \mathbf{k} = 0, \quad \implies \quad \frac{d\mathbf{k}}{dt} = \boldsymbol{\omega}_3 \wedge \mathbf{k}, \quad (6.46)$$

dove ω_i , $i = 1, 2, 3$, sono vettori a priori diversi fra loro. Derivando poi la prima relazione di (6.43) otteniamo

$$0 = \frac{di}{dt} \cdot \mathbf{j} + \mathbf{i} \cdot \frac{dj}{dt} \stackrel{(6.44)}{=} (\omega_1 \wedge \mathbf{i}) \cdot \mathbf{j} + (\omega_2 \wedge \mathbf{j}) \cdot \mathbf{i}, \quad (6.45)$$

ed espressioni analoghe per le altre due relazioni. Scambiando l'ordine del prodotto scalare e del prodotto vettoriale e raccogliendo abbiamo

$$0 = (\omega_1 - \omega_2) \cdot (\mathbf{i} \wedge \mathbf{j}) = (\omega_1 - \omega_2) \cdot \mathbf{k}, \Rightarrow \omega_1 \cdot \mathbf{k} = \omega_2 \cdot \mathbf{k}.$$

Eseguito lo stesso procedimento con le espressioni ottenute derivando $\mathbf{i} \cdot \mathbf{k} = 0$, e $\mathbf{j} \cdot \mathbf{k} = 0$, abbiamo

$$\omega_2 \cdot \mathbf{i} = \omega_3 \cdot \mathbf{i}, \quad \omega_3 \cdot \mathbf{j} = \omega_1 \cdot \mathbf{j}.$$

Come si vede, per ognuno dei vettori ω_i resta non soggetta a condizioni solo una componente: $\omega_1 \cdot \mathbf{i}$, $\omega_2 \cdot \mathbf{j}$, $\omega_3 \cdot \mathbf{k}$. Ora nulla vieta di giocare sull'arbitrarietà di queste componenti e sceglierle in modo che

$$\begin{aligned} \omega_1 \cdot \mathbf{i} &= \omega_2 \cdot \mathbf{i} = \omega_3 \cdot \mathbf{i}, \\ \omega_2 \cdot \mathbf{j} &= \omega_3 \cdot \mathbf{j} = \omega_1 \cdot \mathbf{j}, \\ \omega_3 \cdot \mathbf{k} &= \omega_1 \cdot \mathbf{k} = \omega_2 \cdot \mathbf{k}. \end{aligned}$$

Ma queste relazioni equivalgono a dire che esiste un unico vettore, appunto il vettore velocità angolare, tale che

$$\omega = \omega_1 = \omega_2 = \omega_3.$$

In particolare, ricordando (6.44), (6.45) e (6.46) otteniamo una relazione che fornisce ω in funzione delle derivate dei versori del SdR S ,

$$\mathbf{i} \wedge \frac{di}{dt} + \mathbf{j} \wedge \frac{dj}{dt} + \mathbf{k} \wedge \frac{dk}{dt} = 2\omega.$$

A questo punto, ricordando che $(P - O) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, abbiamo

$$\frac{d(P - O)}{dt} = x \frac{di}{dt} + y \frac{dj}{dt} + z \frac{dk}{dt} = \omega \wedge \underbrace{(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})}_{(P-O)}$$

da cui la (6.41) discende facilmente.

La comparsa del vettore (o “pseudovettore”) ω ha un che di artificioso. Tuttavia il suo significato cinematico può essere reso più chiaro. Per semplificare la trattazione consideriamo il caso in cui il moto sia una **rotazione attorno ad un asse fisso** che, per esempio, coincide con l'asse ζ nel sistema fisso Σ e con l'asse delle z nel sistema solidale S . In questo moto tutti i punti che giacciono sull'asse delle z sono fermi mentre gli altri punti descrivono, rispetto al SdR Σ , delle circonferenze che giacciono su piani perpendicolari all'asse ζ ed hanno centro su di esso. Si indica $\theta(t)$ l'angolo

di rotazione²² (con segno). La matrici $\mathbb{A}(t)$ e $\frac{d\mathbb{A}(t)}{dt}$, associate a questo moto rigido hanno la forma (si veda anche (6.22))

$$\mathbb{A}(t) = \begin{pmatrix} \cos \theta(t) & -\sin \theta(t) & 0 \\ \sin \theta(t) & \cos \theta(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \frac{d\mathbb{A}(t)}{dt} = \begin{pmatrix} -\dot{\theta} \sin \theta(t) & -\dot{\theta} \cos \theta(t) & 0 \\ \dot{\theta} \cos \theta(t) & -\dot{\theta} \sin \theta(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (6.47)$$

È immediato verificare che

$$\frac{d\mathbb{A}(t)}{dt} \mathbb{A}(t) = \begin{pmatrix} 0 & -\dot{\theta} & 0 \\ \dot{\theta} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

e pertanto il vettore velocità angolare è rappresentato, nel SdR²³ Σ , dalla seguente terna

$$\underline{\omega} = \begin{pmatrix} 0 \\ 0 \\ \dot{\theta} \end{pmatrix}, \quad \Leftrightarrow \quad \omega = \dot{\theta} e_3. \quad (6.48)$$

Concludiamo che ω ha per direzione l'asse di rotazione, per modulo il valore assoluto della velocità angolare (rapporto tra angolo percorso e tempo di percorrenza) e per verso quello di e_3 (ovvero ζ crescenti) o $-e_3$ (ζ decrescenti) a seconda che la rotazione avvenga concordemente o non con l'orientazione degli assi (ovvero portando e_1 su e_2 , o viceversa). La rotazione si dirà uniforme se vengono descritti angoli uguali in tempi uguali cioè se $\theta(t) = \omega t + \theta_o$, con ω costante.

Se la rotazione avviene **attorno ad un asse fisso individuato dal versore n** , la generalizzazione della (6.48) è

$$\omega = \dot{\theta} n. \quad (6.49)$$

Qualche considerazione sulle notazioni. Sovente, e lo faremo anche noi nel seguito, la formula vettoriale (6.32) viene, per brevità, denotata così

$$v(P) = v(O) + \omega \wedge (P - O), \quad (6.50)$$

dove: $v(P) = (P - \Omega)$ **rappresenta la velocità del punto P rispetto all'osservatore fisso**; $v(O) = (O - \Omega)$ **denota la velocità, sempre rispetto al medesimo osservatore fisso, del punto O solidale con il corpo rigido \mathcal{B}** ; ω **rappresenta il vettore (o meglio lo "pseudovettore") velocità angolare**.

Riprendendo le frasi iniziali di questo paragrafo, la (6.50) mette bene in luce quanto affermato: la velocità dei punti che costituiscono un sistema rigido varia da punto a punto. In particolare, ricordando la definizione di atto di moto come l'applicazione che ad ogni punto del sistema associa la sua velocità (rispetto ad un determinato SdR fisso), la (6.50) è proprio l'**atto di moto** del sistema rigido.

Per concludere, consideriamo la terna delle componenti di ω rispetto al SdR S . Indichiamo tale terna con $\underline{\omega}_S$, che è così collegata ad $\underline{\omega}$:

$$\underline{\omega} = \mathbb{A} \underline{\omega}_S, \quad \Leftrightarrow \quad \underline{\omega}_S = \mathbb{A}^T \underline{\omega}. \quad (6.51)$$

²²Cioè l'angolo di cui la coppia di versori i, j è ruotata rispetto ai versori e_1, e_2 .

²³Notiamo che in questo caso particolare $\omega = \omega k$, ovvero ω è rappresentato dalla stessa terna anche nel SdR S .

Ci chiediamo adesso, se $\frac{d\mathbb{A}}{dt}\mathbb{A}^T$ è la matrice antisimmetrica che dà origine alla terna $\underline{\omega}$, qual'è la matrice antisimmetrica che dà luogo alla terna $\underline{\omega}_S$? La risposta a tale domanda sarà utile quando calcoleremo proprio $\underline{\omega}_S$ a partire dagli angoli di Eulero nella sezione 6.4. Premettiamo il seguente:

Lemma 6.3.1 *Se \mathbb{S} è una matrice antisimmetrica cui corrisponde il vettore²⁴ \mathbf{w} , allora per ogni matrice $\mathbb{B} \in SO(3)$, al vettore $\mathbb{B}\mathbf{w}$, corrisponde la matrice antisimmetrica $\mathbb{B}\mathbb{S}\mathbb{B}^T$.*

Dim. Si ricorda che, \forall terna $\mathbf{y} \in \mathbb{R}^3$, vale la seguente uguaglianza $\mathbb{S}\mathbf{y} = \mathbf{w} \wedge \mathbf{y}$. Considerando adesso la matrice antisimmetrica $\mathbb{B}\mathbb{S}\mathbb{B}^T$, avremo

$$\mathbb{B}\mathbb{S}\mathbb{B}^T\mathbf{y} = \mathbb{B}\left(\mathbb{S}\mathbb{B}^T\mathbf{y}\right) = \mathbb{B}\left(\mathbf{w} \wedge \mathbb{B}^T\mathbf{y}\right) \stackrel{(6.38)}{=} \mathbb{B}\mathbf{w} \wedge \underbrace{\mathbb{B}\mathbb{B}^T}_{\mathbb{I}}\mathbf{y} = \mathbb{B}\mathbf{w} \wedge \mathbf{y}.$$

Quindi, $\forall \mathbf{y} \in \mathbb{R}^3$, abbiamo dimostrato $\mathbb{B}\mathbb{S}\mathbb{B}^T\mathbf{y} = \mathbb{B}\mathbf{w} \wedge \mathbf{y}$, ovvero che alla matrice antisimmetrica $\mathbb{B}\mathbb{S}\mathbb{B}^T$ è associato il vettore (o meglio lo “pseudovettore”) $\mathbb{B}\mathbf{w}$ e viceversa.

□

Applicando il lemma alla matrice antisimmetrica $\frac{d\mathbb{A}}{dt}\mathbb{A}^T$, e considerando come matrice di $SO(3)$ la matrice \mathbb{A}^T , abbiamo che al vettore $\mathbb{A}^T\underline{\omega}$, ovvero ad $\underline{\omega}_S$, è associato alla matrice

$$\underbrace{\mathbb{A}^T}_{\mathbb{B}} \underbrace{\left(\frac{d\mathbb{A}}{dt}\mathbb{A}^T\right)}_{\mathbb{S}} \underbrace{(\mathbb{A}^T)^T}_{\mathbb{B}^T} = \mathbb{A}^T \frac{d\mathbb{A}}{dt} \underbrace{\mathbb{A}^T\mathbb{A}}_{\mathbb{I}} = \mathbb{A}^T \frac{d\mathbb{A}}{dt}.$$

Il tutto quindi può essere riassunto in questo semplice schema

$$\begin{aligned} \frac{d\mathbb{A}}{dt}\mathbb{A}^T &\longrightarrow \underline{\omega} \\ \mathbb{A}^T \frac{d\mathbb{A}}{dt} &\longrightarrow \underline{\omega}_S \end{aligned} \quad \text{dove } \underline{\omega}_S = \mathbb{A}^T\underline{\omega}. \tag{6.52}$$

Notiamo che tale risultato ribadisce ancora la coerenza intrinseca dalla (6.23). Infatti esprimiamo tale formula nel SdR S , cioè

$$\mathbb{A}^T\dot{\xi} = \mathbb{A}^T\dot{\xi}_O + \underbrace{\mathbb{A}^T \frac{d\mathbb{A}}{dt}}_{\underline{\omega}_S \wedge} \mathbf{x},$$

che appunto fornisce, nel SdR S , le componenti del vettore velocità di P misurata dall'osservatore Σ . Infatti ad $\mathbb{A}^T\dot{\mathbb{A}}$ associamo proprio la terna $\underline{\omega}_S$ che, coerentemente, è costituita dalle componenti di $\underline{\omega}$ rispetto al SdR S .

²⁴Inteso come terna di \mathbb{R}^3 .

Nota 6.3.3 Il teorema di Eulero, come abbiamo sottolineato, afferma che è sempre possibile sovrapporre il SdR S al SdR Σ (o meglio al SdR Σ') eseguendo un'unica rotazione intorno ad un'opportuna retta. Ci chiediamo adesso: in che relazione stanno l'asse di rotazione ed il vettore ω ? Se $\omega = \omega(t) \mathbf{n}$, con \mathbf{n} costante nel tempo allora ω è parallelo all'asse di rotazione del teorema di Eulero. Infatti, è sufficiente selezionare il SdR Σ ed il SdR solidale S in modo che i versori degli assi ζ e z coincidano con \mathbf{n} , per ottenere la matrice \mathbb{A} nella forma (6.47), da cui ricavre $\omega = \omega(t) \mathbf{e}_3$, ed $\mathbf{e}_3 = \mathbf{k}$, cioè \mathbf{e}_3 è l'autovettore di \mathbb{A} relativo all'autovalore $\lambda = 1$ (e quindi è proprio l'asse di rotazione del teorema di Eulero). Se invece ω non ha direzione costante, allora, in generale, ω non è parallelo all'asse di rotazione del teorema di Eulero. Questo discende direttamente dalla (6.52): le terne delle componenti del vettore ω rispetto al SdR Σ e al SdR S sono diverse. Quindi ω non è, in generale, autovettore relativo all'autovalore $\lambda = 1$ della matrice \mathbb{A} .

6.4 Angoli di Eulero

Abbiamo visto che il moto rigido è parametrizzato dai “punti” dello spazio $\mathbb{R}^3 \times SO(3)$. I primi tre parametri, in \mathbb{R}^3 , sono le coordinate dell'origine O del sistema solidale rispetto al SdR fisso Σ . Gli altri tre parametri dovranno determinare l'orientamento degli assi di S rispetto agli assi del sistema Σ , o meglio, la posizione dei tre versori di S rispetto a Σ' , che è il SdR con origine in O ed assi costantemente paralleli a quelli di Σ (v. figura 6.4). Il moto di S rispetto a Σ' è quindi equivalente a quello di un moto rigido con un punto che resta fisso. Un tale moto prende il nome di **precessione**. Fissiamo quindi una posizione mutua di S e Σ' e supponiamo che, in questa configurazione, gli assi dei due sistemi non siano sovrapposti e, in particolare, non coincidano i due assi ζ e z (vedi figura 6.7). Vogliamo far vedere che possiamo riportare il sistema “fisso” Σ' a coincidere con quello “mobile” S mediante tre opportune rotazioni: gli angoli di queste rotazioni forniranno i parametri cercati.

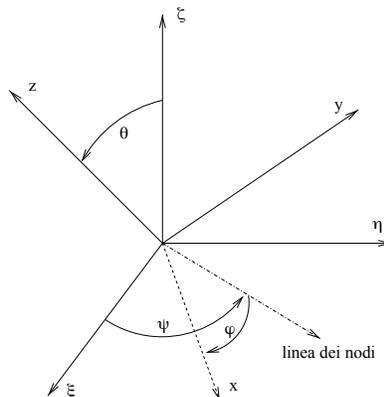


Figura 6.7: Angoli di Eulero.

Poiché gli assi ζ e z non coincidono, i piani $\zeta = 0$ e $z = 0$ si intersecano in una retta detta **linea dei nodi**. Sia \mathbf{n} il versore di questa retta, orientato in modo che la terna \mathbf{k} , \mathbf{n} e \mathbf{e}_3 (cioè i versori dell'asse z , della linea dei nodi, e dell'asse delle ζ) sia positivamente orientata. Ovviamente la linea dei nodi sta sia sul piano (ξ, η) , che sul piano (x, y) . Pertanto indichiamo con ψ l'angolo che \mathbf{e}_1 , cioè l'asse ξ , forma con \mathbf{n} , e con ϕ l'angolo $\hat{\mathbf{i}}$, cioè l'asse x , forma con la linea dei nodi. Sia poi θ l'angolo formato fra gli assi z e ζ .

Nota 6.4.1 *La dimostrazione della possibilità di rappresentare una generica rotazione fra due SdR tramite la composizione ordinata di 3 rotazioni elementari è dovuta, ancora una volta, ad Eulero. I tre angoli di cui ruotare sono detti **angoli di Eulero**. Tuttavia oltre a dover specificare i valori dei 3 angoli è necessario specificare quali siano gli assi attorno ai quali devono essere effettuate le rotazioni elementari. La scelta di tali assi è libera, cosicché esistono diverse possibilità. Nella presente sezione abbiamo scelto come assi quelli individuati dai versori \mathbf{k} , \mathbf{n} ed \mathbf{e}_3 . Essa viene anche detta ZXZ, dal nome degli assi intorno ai quali, di volta in volta, si effettuano le singole rotazioni. Infatti, come spiegheremo in dettaglio nel seguito, partendo dal SdR Σ , si ruota dapprima attorno ad \mathbf{e}_3 (l'asse “ z ” di Σ), poi una volta ottenuto il nuovo SdR si ruota attorno a quello che sarebbe l'asse x di tale nuovo SdR. Si ottiene così un altro SdR e , come ultimo passo, si ruota attorno all'asse z di quest'ultimo SdR. Questo è il motivo per cui tale scelta degli assi di rotazione è convenzionalmente chiamata ZXZ. Le altre possibili scelte sono XZX, XYX, YXY, YZY, ZYZ, XZY, XYZ, YXZ, YZX, ZYX e ZXY. Le 12 sequenze sono ottenute tramite tutte le disposizioni possibili in cui assi uguali non sono consecutivi.*

Vediamo come determinare \mathbb{A}^T , così che $\mathbf{x} = \mathbb{A}^T \boldsymbol{\xi}$, che corrisponde, come detto, a passare dalla terna Σ' alla terna S . Tale passaggio si fa mediante l'applicazione di tre successive rotazioni (cui corrispondono tre distinte matrici di rotazione):

1. **Prima rotazione** (in senso antiorario, scelto come positivo) **di un angolo ψ , attorno all'asse \mathbf{e}_3 , in modo che \mathbf{e}_1 vada a coincidere con \mathbf{n}** . Si passa quindi dal SdR $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ al SdR $(\mathbf{e}'_1 \equiv \mathbf{n}, \mathbf{e}'_2, \mathbf{e}'_3 \equiv \mathbf{e}_3)$, la cui matrice è data da²⁵

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \longrightarrow (\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3) \quad \text{matrice } \mathbb{A}_\psi,$$

$$(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3) \longrightarrow (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \quad \text{matrice } \mathbb{A}_\psi^T,$$

con

$$\mathbb{A}_\psi = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbb{A}_\psi^T = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. **Seconda rotazione di un angolo θ , detto angolo di nutazione** (sempre in verso antiorario e quindi positivo) **attorno all'asse \mathbf{e}'_1 che coincide con linea dei nodi**.

²⁵Si ricordi che secondo la nostra convenzione

$$(\text{terna in } \Sigma) = \mathbb{A} (\text{terna in } S).$$

Si passa da $(e'_1 \equiv n, e'_2, e'_3)$ al SdR $(e''_1 \equiv e'_1, e''_2, e''_3)$ ottenendo la seguente matrice

$$(e'_1, e'_2, e'_3) \longrightarrow (e''_1, e''_2, e''_3) \quad \text{matrice } \mathbb{A}_\theta,$$

$$(e''_1, e''_2, e''_3) \longrightarrow (e'_1, e'_2, e'_3) \quad \text{matrice } \mathbb{A}_\theta^T,$$

con

$$\mathbb{A}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, \quad \mathbb{A}_\theta^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}.$$

3. **La terza rotazione, di un angolo ϕ , detto angolo di rotazione propria** (sempre in verso positivo antiorario) **attorno al vettore e''_3** (che rimane quindi invariato). Si passa da (e''_1, e''_2, e''_3) al SdR S , ovvero $(i, j, k \equiv e''_3)$,

$$(e''_1, e''_2, e''_3) \longrightarrow (i, j, k) \quad \text{matrice } \mathbb{A}_\phi,$$

$$(i, j, k) \longrightarrow (e''_1, e''_2, e''_3) \quad \text{matrice } \mathbb{A}_\phi^T,$$

con

$$\mathbb{A}_\phi = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbb{A}_\phi^T = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matrice \mathbb{A}^T (che consente di trasformare le terne di Σ' in terne di S) è dunque

$$\mathbb{A}^T = \mathbb{A}_\phi^T \mathbb{A}_\theta^T \mathbb{A}_\psi^T,$$

mentre la matrice che permette di passare dalle terne di S a quelle espresse in Σ' (cioè la matrice \mathbb{A}) è data da

$$\mathbb{A} = (\mathbb{A}^T)^T = (\mathbb{A}_\phi^T \mathbb{A}_\theta^T \mathbb{A}_\psi^T)^T = \mathbb{A}_\psi \mathbb{A}_\theta \mathbb{A}_\phi,$$

cioè²⁶

$$\mathbb{A} = \begin{pmatrix} \cos \phi \cos \psi - \cos \theta \sin \phi \sin \psi & -\cos \psi \sin \phi - \cos \theta \cos \phi \sin \psi & \sin \theta \sin \psi \\ \cos \phi \sin \psi + \cos \theta \cos \psi \sin \phi & \cos \theta \cos \phi \cos \psi - \sin \phi \sin \psi & -\cos \psi \sin \theta \\ \sin \theta \sin \phi & \cos \phi \sin \theta & \cos \theta \end{pmatrix}.$$

Da essa poi si ricava l'espressione del vettore ω nel SdR solidale S : la terna²⁷ $\underline{\omega}_S$. Ricordando poi lo schema (6.52), alla terna $\underline{\omega}_S$ è associata alla matrice antisimmetrica

²⁶Tra gli angoli di Eulero e le orientazioni del corpo rigido esiste una corrispondenza solo **localmente** biunivoca. Infatti, si osservi che quando l'asse ζ coincide con z ($\theta = 0$ e $\theta = \pi$) non si può più definire la linea dei nodi perché il piano solidale x, y e il piano fisso ξ, η sono sovrapposti; in tale caso (che escludiamo dalle nostre considerazioni) è sufficiente l'unico parametro ψ per descrivere l'orientazione reciproca delle due terne. Questo "difetto" è inevitabile, qualsiasi sia il sistema di parametri si scelga per determinare la posizione di S , in quanto l'insieme $SO(3)$, i cui punti sono in corrispondenza 1-1 con le posizioni di S , non ammette sistemi coordinate globali (come accade ad esempio per la superficie di una sfera).

²⁷L'espressione di ω in S , cioè la terna $\underline{\omega}_S$, risulta particolarmente utile nell'ambito della dinamica dove è vien fatto uso delle componenti del momento angolare rispetto ad una base solidale e dove l'energia cinetica viene calcolata lavorando nel sistema S .

$\mathbb{A}^T \frac{d\mathbb{A}}{dt}$. Cominciamo quindi col calcolare

$$\frac{d\mathbb{A}}{dt} = \frac{d\mathbb{A}_\psi}{dt} \mathbb{A}_\theta \mathbb{A}_\phi + \mathbb{A}_\psi \frac{d\mathbb{A}_\theta}{dt} \mathbb{A}_\phi + \mathbb{A}_\psi \mathbb{A}_\theta \frac{d\mathbb{A}_\phi}{dt}. \quad (6.53)$$

E' facile provare, svolgendo esplicitamente i calcoli, che

$$\frac{d\mathbb{A}_\psi}{dt} = \mathbb{A}_\psi \mathbb{J}_{\dot{\psi}}, \quad \text{dove} \quad \mathbb{J}_{\dot{\psi}} = \begin{pmatrix} 0 & -\dot{\psi} & 0 \\ \dot{\psi} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\frac{d\mathbb{A}_\theta}{dt} = \mathbb{A}_\theta \mathbb{J}_{\dot{\theta}}, \quad \text{dove} \quad \mathbb{J}_{\dot{\theta}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\dot{\theta} \\ 0 & \dot{\theta} & 0 \end{pmatrix},$$

$$\frac{d\mathbb{A}_\phi}{dt} = \mathbb{A}_\phi \mathbb{J}_{\dot{\phi}}, \quad \text{dove} \quad \mathbb{J}_{\dot{\phi}} = \begin{pmatrix} 0 & -\dot{\phi} & 0 \\ \dot{\phi} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In particolare, le tre matrici $\mathbb{J}_{\dot{\psi}}$, $\mathbb{J}_{\dot{\theta}}$ e $\mathbb{J}_{\dot{\phi}}$ sono matrici antisimmetriche cui possiamo associare, in base alla regola (6.27), le seguenti terne

$$\mathbb{J}_{\dot{\psi}} \rightarrow \underline{\dot{\psi}} = \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix}, \quad \mathbb{J}_{\dot{\theta}} \rightarrow \underline{\dot{\theta}} = \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbb{J}_{\dot{\phi}} \rightarrow \underline{\dot{\phi}} = \begin{pmatrix} 0 \\ 0 \\ \dot{\phi} \end{pmatrix}. \quad (6.54)$$

Quindi, sfruttando la notazione appena introdotta e la (6.53), abbiamo

$$\begin{aligned} \mathbb{A}^T \frac{d\mathbb{A}}{dt} &= (\mathbb{A}_\psi \mathbb{A}_\theta \mathbb{A}_\phi)^T \left[\mathbb{A}_\psi \mathbb{J}_{\dot{\psi}} \mathbb{A}_\theta \mathbb{A}_\phi + \mathbb{A}_\psi \mathbb{A}_\theta \mathbb{J}_{\dot{\theta}} \mathbb{A}_\phi + \mathbb{A}_\psi \mathbb{A}_\theta \mathbb{A}_\phi \mathbb{J}_{\dot{\phi}} \right] \\ &= \underbrace{\mathbb{A}_\phi^T \mathbb{A}_\theta^T}_{(\mathbb{A}_\theta \mathbb{A}_\phi)^T} \mathbb{J}_{\dot{\psi}} \mathbb{A}_\theta \mathbb{A}_\phi + \mathbb{A}_\phi^T \mathbb{J}_{\dot{\theta}} \mathbb{A}_\phi + \mathbb{J}_{\dot{\phi}}. \end{aligned}$$

Ora, rammentando il lemma 6.3.1 dimostrato nella precedente sezione, otteniamo le seguenti associazioni tra matrici antisimmetriche e terne

$$(\mathbb{A}_\theta \mathbb{A}_\phi)^T \mathbb{J}_{\dot{\psi}} \left((\mathbb{A}_\theta \mathbb{A}_\phi)^T \right)^T \longrightarrow (\mathbb{A}_\theta \mathbb{A}_\phi)^T \underline{\dot{\psi}} = \mathbb{A}_\phi^T \mathbb{A}_\theta^T \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix},$$

$$\mathbb{A}_\phi^T \mathbb{J}_{\dot{\theta}} \left(\mathbb{A}_\phi^T \right)^T \longrightarrow \mathbb{A}_\phi^T \underline{\dot{\theta}} = \mathbb{A}_\phi^T \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix},$$

mentre alla matrice $\mathbb{J}_{\dot{\phi}}$ associamo direttamente la terna $\underline{\dot{\phi}}$ come specificato nella (6.54). Abbiamo quindi

$$\underline{\omega}_S = \mathbb{A}_\phi^T \mathbb{A}_\theta^T \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} + \mathbb{A}_\phi^T \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} \dot{\theta} \cos \phi + \dot{\psi} \sin \theta \sin \phi \\ \dot{\psi} \cos \phi \sin \theta - \dot{\theta} \sin \phi \\ \dot{\phi} + \dot{\psi} \cos \theta \end{pmatrix}. \quad (6.55)$$

Procedendo in modo simile si trovano le componenti di ω nel SdR Σ , ovvero la terna $\underline{\omega}$. Infatti

$$\begin{aligned} \frac{d\mathbb{A}}{dt} \mathbb{A}^T &= \left(\underbrace{\frac{d\mathbb{A}_\psi}{dt} \mathbb{A}_\theta \mathbb{A}_\phi}_{\mathbb{A}_\psi \mathbb{J}_\psi} + \mathbb{A}_\psi \underbrace{\frac{d\mathbb{A}_\theta}{dt} \mathbb{A}_\phi}_{\mathbb{A}_\theta \mathbb{J}_\theta} + \mathbb{A}_\psi \mathbb{A}_\theta \underbrace{\frac{d\mathbb{A}_\phi}{dt}}_{\mathbb{A}_\phi \mathbb{J}_\phi} \right) \mathbb{A}_\phi^T \mathbb{A}_\theta^T \mathbb{A}_\psi^T \\ &= \mathbb{A}_\psi \mathbb{J}_\psi \mathbb{A}_\psi^T + \mathbb{A}_\psi \mathbb{A}_\theta \mathbb{J}_\theta \mathbb{A}_\theta^T \mathbb{A}_\psi^T + \mathbb{A}_\psi \mathbb{A}_\theta \mathbb{A}_\phi \mathbb{J}_\phi \mathbb{A}_\phi^T \mathbb{A}_\theta^T \mathbb{A}_\psi^T, \end{aligned}$$

a cui si associa la terna

$$\underline{\omega} = \mathbb{A}_\psi \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} + \mathbb{A}_\psi \mathbb{A}_\theta \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix} + \mathbb{A}_\psi \mathbb{A}_\theta \mathbb{A}_\phi \begin{pmatrix} 0 \\ 0 \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} \dot{\theta} \cos \psi + \dot{\phi} \sin \theta \sin \psi \\ \dot{\theta} \sin \psi - \dot{\phi} \cos \psi \sin \theta \\ \dot{\psi} + \dot{\phi} \cos \theta \end{pmatrix}. \quad (6.56)$$

Notiamo che avremmo ottenuto la stessa espressione se avessimo usato la relazione $\underline{\omega} = \mathbb{A} \underline{\omega}_S$.

Nota 6.4.2 Riprendendo quanto illustrato nella nota 6.3.3, rimarchiamo ancora come, nel caso generale, ω non sia parallelo all'asse di rotazione del teorema di Eulero. Confrontando le terne (6.55) e (6.56) risulta evidente che le componenti del vettore ω sono diverse nei due SdR. Quindi ω non è, in generale, un autovettore della matrice \mathbb{A} relativo all'autovalore $\lambda = 1$. Lo diventa quando $\underline{\omega} = \underline{\omega}_S$, cioè quando, per esempio, $\theta \equiv 0$, $\phi \equiv 0$ e la rotazione, di un angolo ψ , avviene attorno all'asse e_3 , che così rimane coincidente con l'asse z del SdR S .

6.5 Asse istantaneo di moto, rigate del moto

Vogliamo ora determinare come è fatto il campo di velocità, o meglio l'atto di moto, determinato dalla formula (6.50). Fissiamo un istante t e indichiamo con $v(P)$ la velocità di un punto solidale P a questo istante. Moltiplicando scalarmente la (6.50) per il vettore $\omega/|\omega|$ otteniamo

$$v(P) \cdot \frac{\omega}{|\omega|} = v(O) \cdot \frac{\omega}{|\omega|}, \quad (6.57)$$

che ci dice che la componente della velocità nella direzione di ω è la stessa per tutti i punti solidali, ovvero per ogni $P \in \mathfrak{B}$,

$$v_{\parallel} = \left(v(P) \cdot \frac{\omega}{|\omega|} \right) \frac{\omega}{|\omega|}.$$

Quindi, per ogni punto P vale la seguente scomposizione nella direzione parallela ad ω (individuata dal versore $\frac{\omega}{|\omega|}$) ed in quella ortogonale

$$v(P) = \underbrace{v_{\parallel}}_{\text{uguale } \forall P} + v(P)_{\perp}, \quad \text{dove } v(P)_{\perp} = v(P) - v_{\parallel}, \quad (6.58)$$

o meglio, ricordando ancora la (6.50),

$$\mathbf{v}(P)_{\perp} = \mathbf{v}(O)_{\perp} + \underbrace{\boldsymbol{\omega} \wedge (P - O)}_{\text{vettore } \perp \boldsymbol{\omega}}, \quad (6.59)$$

con $\mathbf{v}(O)_{\perp} = \mathbf{v}(O) - \mathbf{v}(O)_{\parallel} = \mathbf{v}(O) - \left(\mathbf{v}(O) \cdot \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|} \right) \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|}$, componente di $\mathbf{v}(O)$ ortogonale ad $\boldsymbol{\omega}$. Al variare di P solo la componente di $\mathbf{v}(P)$ ortogonale alla direzione di $\boldsymbol{\omega}$, ovvero il vettore $\mathbf{v}(P)_{\perp}$, può cambiare. Mostriamo che esiste una intera retta, **detta asse istantaneo del moto**, di punti solidali²⁸ Q tali che $\mathbf{v}(Q)_{\perp} = 0$, e quindi $\mathbf{v}(Q)$ si riduce a

$$\mathbf{v}(Q) = \mathbf{v}_{\parallel}, \quad \forall Q \in \text{asse istantaneo del moto.} \quad (6.60)$$

Nota 6.5.1 *Si osserva che la definizione di asse istantaneo di moto coincide con la definizione di asse centrale data nella sezione 1.5.2.*

Scrivendo le (6.58) (6.59) per tali punti abbiamo

$$\begin{aligned} \mathbf{v}(Q) &= \mathbf{v}_{\parallel} + \mathbf{v}(Q)_{\perp} \\ &= \underbrace{\mathbf{v}_{\parallel}}_{\text{componente } \parallel \boldsymbol{\omega}} + \underbrace{\mathbf{v}(O)_{\perp} + \boldsymbol{\omega} \wedge (Q - O)}_{\text{componente } \perp \boldsymbol{\omega}}, \end{aligned}$$

e, imponendo $\mathbf{v}(Q)_{\perp} = 0$, otteniamo

$$\mathbf{v}(O)_{\perp} + \boldsymbol{\omega} \wedge (Q - O) = 0. \quad (6.61)$$

Possiamo riscrivere l'equazione (6.61) considerando la rappresentazione matriciale della velocità angolare

$$\begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \begin{pmatrix} \xi_Q \\ \eta_Q \\ \zeta_Q \end{pmatrix} = - \begin{pmatrix} v_1(O)_{\perp} \\ v_2(O)_{\perp} \\ v_3(O)_{\perp} \end{pmatrix}, \quad (6.62)$$

dove $\begin{pmatrix} \xi_Q \\ \eta_Q \\ \zeta_Q \end{pmatrix}$ sono le componenti del vettore $(Q - O)$ rispetto al SdR Σ e $\begin{pmatrix} v_1(O)_{\perp} \\ v_2(O)_{\perp} \\ v_3(O)_{\perp} \end{pmatrix}$

quelle del vettore $\mathbf{v}(O)_{\perp}$, nel medesimo SdR. Ora, il sistema (6.62) può essere visto come un sistema lineare non omogeneo di tre equazioni (6.62) nelle incognite ξ_Q , η_Q , ζ_Q . La matrice del sistema ha rango 2 (salvo quando $\boldsymbol{\omega} = 0$) e la condizione $\boldsymbol{\omega} \cdot \mathbf{v}(O)_{\perp} = 0$, garantisce che anche la matrice completa

$$\begin{pmatrix} 0 & -\omega_3 & \omega_2 & -v_1(O)_{\perp} \\ \omega_3 & 0 & -\omega_1 & -v_2(O)_{\perp} \\ -\omega_2 & \omega_1 & 0 & -v_3(O)_{\perp} \end{pmatrix}$$

²⁸In generale non è detto che i punti Q facciano parte del corpo rigido \mathfrak{B} . Può accadere che $Q \notin \mathfrak{B}$, pur restando solidali con \mathfrak{B} .

abbia ancora rango 2. Infatti, comunque si selezioni una matrice 3×3 che contenga l'ultima colonna, quando se ne calcola il determinante si ottiene $\omega_i (\omega \cdot v(O)_\perp)$, $i = 1, 2, 3$, e quindi si annulla. Ne segue che (6.62) ammette ∞^1 soluzioni: ovvero una retta. In particolare, se consideriamo due soluzioni $(Q_1 - O)$ e $(Q_2 - O)$, il vettore $(Q_1 - Q_2)$ giace sulla retta. Ma siccome

$$\begin{cases} v(O)_\perp + \omega \wedge (Q_1 - O) = 0, \\ v(O)_\perp + \omega \wedge (Q_2 - O) = 0, \end{cases} \quad \xRightarrow{\text{sottraendo}} \quad \omega \wedge (Q_1 - Q_2) = 0,$$

otteniamo che l'asse istantaneo di moto è parallelo a ω .

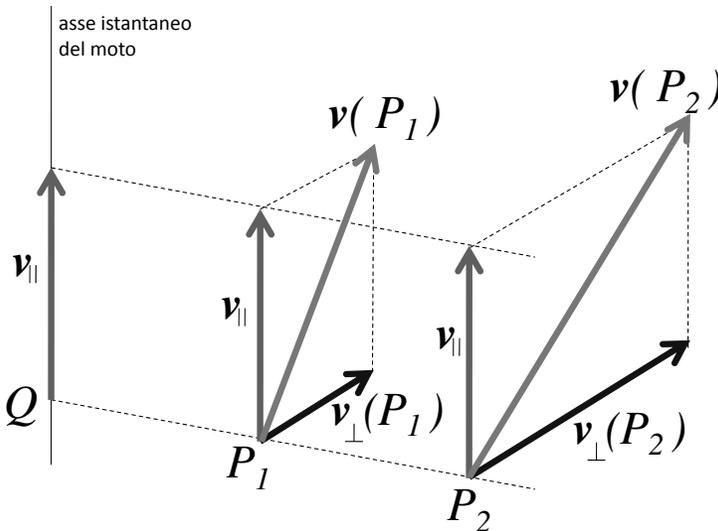


Figura 6.8: Il punto Q sta sull'asse istantaneo del moto. Per ogni punto P vale la scomposizione $v(P) = v_{||} + v_{\perp}(P)$. La componente $v_{||}$ è uguale per tutti i punti ed è pari alla velocità dei giacenti sull'asse istantaneo del moto.

Sull'asse istantaneo di moto la velocità rispetto a Σ assume il minimo modulo e coincide con $v_{||}$, ovvero ha la stessa direzione dell'asse istantaneo di moto, come mostrato in fig. 6.8. Quindi l'asse del moto “scivola” lungo la sua giacitura, o meglio la velocità dei punti materiali che ivi giacciono è parallela all'asse. Infatti, dalla (6.58) si ha

$$|v(P)|^2 = |v_{||}|^2 + |v(P)_{\perp}|^2,$$

poiché $v_{||} \cdot v(P)_{\perp} = 0$. Dunque, per ogni $P \in \mathfrak{B}$, $|v(P)| \geq |v_{||}|$, ed il valor minimo verrà assunto da quei punti P per cui $v(P)_{\perp} = 0$, che sono proprio i punti che giacciono sull'asse istantaneo di moto. Bisogna però ricordare che i vettori ω e $v(O)$ sono variabili nel tempo, e quindi anche la posizione dell'asse di moto varia nel tempo

e i punti (solidali) che si trovano sull'asse ad un istante t_2 sono diversi da quelli che si trovavano all'istante t_1 . Ciò giustifica l'aggettivo "istantaneo".

Fissato t , la retta che descrive l'asse istantaneo di moto nel SdR Σ , può essere data in forma parametrica

$$\begin{cases} \xi = \xi(\mu, t), \\ \eta = \eta(\mu, t), \\ \zeta = \zeta(\mu, t), \end{cases} \quad (6.63)$$

dove $\mu \in \mathbb{R}$ è il parametro che descrive la retta a t fissato, e t è il tempo. Al variare di t si ottengono, come dicevamo, rette diverse. Se poi consideriamo il sistema (6.63) con t e μ entrambe variabili in \mathbb{R} , si ottiene l'equazione parametrica di una superficie (si rimanda alla sezione 4.1.3 per le superfici date in forma parametrica). Tale superficie è una **rigata**, in quanto costituita da rette, e viene detta **rigata fissa**. L'equazione della stessa superficie (6.63) nel SdR solidale S si ottiene riscrivendo l'equazione (6.62) nel SdR S . Quindi, se

$$\xi_Q = \begin{pmatrix} \xi_Q \\ \eta_Q \\ \zeta_Q \end{pmatrix}$$

denota la terna delle componenti di $(Q - O)$ nel SdR Σ , la terna $\mathbf{x}_Q = \begin{pmatrix} x_Q \\ y_Q \\ z_Q \end{pmatrix}$,

delle componenti del medesimo vettore nel SdR solidale S è data da

$$\mathbf{x}_Q = \mathbb{A}^T \xi_Q.$$

Applicando la matrice \mathbb{A}^T alla (6.62) otteniamo

$$0 = \mathbb{A}^T \mathbf{v}(O)_\perp + \mathbb{A}^T (\boldsymbol{\omega} \wedge \xi_Q) \stackrel{(6.38)}{=} \mathbb{A}^T \mathbf{v}(O)_\perp + \underbrace{\boldsymbol{\omega}_s}_{\mathbb{A}^T \boldsymbol{\omega}} \wedge \underbrace{\mathbf{x}_Q}_{\mathbb{A}^T \xi_Q}, \quad (6.64)$$

che è appunto l'equazione dell'asse istantaneo di moto nel SdR S . Evidentemente tale retta varia nel tempo e descrive una superficie che viene detta **rigata mobile**.

Esempio 6.5.1 Riprendiamo l'esempio 4.1.2 considerando un disco di raggio R che rotola senza strisciare lungo l'asse ξ , come mostrato in figura 6.9.

La velocità angolare è $\boldsymbol{\omega} = \omega e_3$, con $\omega = \dot{\varphi}$. Di conseguenza, siccome il moto avviene nel piano (ξ, η) , $\mathbf{v}_\parallel = 0$, e $\mathbf{v} = \mathbf{v}_\perp$. L'asse istantaneo del moto è dunque la retta dove $\mathbf{v}_\perp = 0$. Se Q è un generico punto del disco, le componenti del vettore $(Q - \Omega)$ rispetto al SdR Σ sono $(Q - \Omega) = \xi e_1 + \eta e_2$. Ricordando il vincolo di rotolamento puro (4.14), abbiamo

$$(O - \Omega) = -R\varphi e_1 + R e_2, \quad (6.65)$$

e pertanto, siccome $(Q - O) = (Q - \Omega) - (O - \Omega)$, si ha

$$\begin{pmatrix} \xi_Q \\ \eta_Q \\ \zeta_Q \end{pmatrix} = \begin{pmatrix} \xi - R\varphi \\ \eta - R \\ 0 \end{pmatrix}.$$

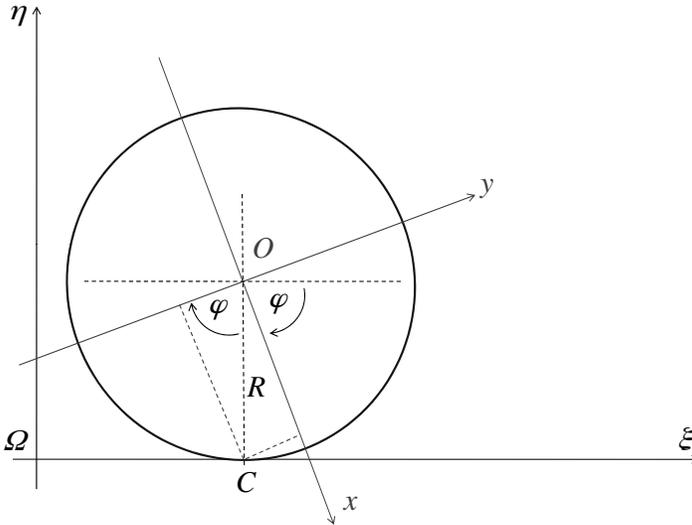


Figura 6.9: Disco che rotola di rotolamento puro. (O, x, y) è il SdR solido S . Le coordinate nel SdR S , del punto di contatto C sono $(-R \sin \varphi, -R \cos \varphi)$. Si osservi che in questo caso l'angolo di rotolamento φ è negativo.

Sempre la (6.65) ci consente di determinare le componenti del vettore $\mathbf{v}_\perp(O)$ nel SdR Σ . Infatti²⁹

$$\mathbf{v}(O) = \frac{d(O - \Omega)}{dt} = -R\omega \mathbf{e}_1.$$

Scrivendo la (6.62) otteniamo

$$\begin{aligned} 0 &= -R\omega \mathbf{e}_1 + \omega \mathbf{e}_3 \wedge [(\xi - R\varphi) \mathbf{e}_1 + (\eta - R) \mathbf{e}_2] \\ &= (-\omega\eta) \mathbf{e}_1 + [\omega(\xi + R\varphi)] \mathbf{e}_2. \end{aligned}$$

L'equazione parametrica, nel SdR Σ , dell'asse istantaneo di moto è dunque

$$\begin{cases} \eta = 0, \\ \xi = -R\varphi, \\ \zeta = \mu, \end{cases} \quad \text{con } \mu \in \mathbb{R}.$$

L'asse di moto è pertanto la retta ortogonale al piano ξ, η , che passa per il punto di contatto. Questo risultato è coerente con la "fisica" del problema, dal momento che il punto di contatto è quel punto che, istante per istante, ha velocità nulla (e quindi la minore in modulo). La rigata fissa è quindi il piano ξ, ζ .

²⁹Si ricordi che, in questo caso particolare, $\mathbf{v}(O)_\perp = \mathbf{v}(O)$.

Se vogliamo determinare l'equazione dell'asse di moto nel SdR S , dobbiamo lavorare con la (6.64). Abbiamo

$$\mathbb{A}^T \mathbf{v}(O) = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -R\omega \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -R\omega \cos \varphi \\ R\omega \sin \varphi \\ 0 \end{pmatrix},$$

mentre

$$\boldsymbol{\omega}_S \wedge \mathbf{x} = \omega \mathbf{k} \wedge (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = -\omega y\mathbf{i} + \omega x\mathbf{j}.$$

Scrivendo quindi la (6.64) otteniamo la seguente equazione parametrica dell'asse di moto nel SdR S

$$\begin{cases} x = -R \sin \varphi, \\ y = -R \cos \varphi \\ z = \mu, \end{cases} \quad \text{con } \mu \in \mathbb{R},$$

ovvero una retta perpendicolare al piano x, y . Tale retta tocca la circonferenza esterna del disco nel punto di coordinate $(R \sin(-\varphi), -R \cos \varphi)$, che sono proprio le coordinate, rispetto al SdR S , del punto di contatto (v. figura 6.9). In questo caso la rigata mobile è il cilindro di raggio R il cui asse è l'asse z .

Di particolare interesse il caso in cui l'invariante \mathbf{v}_{\parallel} è nullo (come nel precedente esempio 6.5.1). In questo caso i punti dell'asse di moto sono "istantaneamente fermi", cioè hanno velocità nulla. Ne segue che la rigata mobile *rotola senza strisciare* sulla rigata fissa e possiamo pensare quindi al moto rigido come "generato" dal rotolamento di queste due superfici, una sull'altra. Questa osservazione è alla base della teoria su cui si basa la costruzione degli *ingranaggi*. In questo caso l'asse viene anche detto *asse di istantanea rotazione*, in quanto la distribuzione (istantanea) delle velocità è la stessa di quella che ci sarebbe in un moto di rotazione (in cui l'asse del moto resta fisso).

Nel caso generale, al variare di P , si "sommano" il moto di rotazione dato dalla $\mathbf{v}(P)_{\perp}$, con il moto di "traslazione" dato da \mathbf{v}_{\parallel} (uguale per tutti i punti di \mathfrak{B}). Se ω e $\mathbf{v}(O)$, restano costanti allora le velocità hanno sempre questa distribuzione e, in particolare, l'asse di moto è sempre lo stesso. Il moto che ne risulta è quindi quello in cui i punti dell'asse si muovono lungo l'asse, mentre gli altri punti li "seguono girando attorno all'asse". Questo moto è detto "elicoidale", ed è il moto che ha un bullone mentre viene avvitato, o un cavatappi mentre entra nel tappo. Nel caso generale questa è la forma della distribuzione di velocità ad ogni istante, ma da un istante all'altro la posizione dell'asse muta. Per questa si parla, per un moto rigido generico, di **atto di moto elicoidale** o **roto-traslatorio**.

6.6 Cinematica relativa: composizione delle velocità

Vogliamo adesso determinare come due SdR Σ e S , in moto tra loro, giudichino la velocità di un punto P in moto rispetto ad entrambi gli osservatori (e quindi non più fermo rispetto al SdR S). Anche se i termini "fisso" e "mobile" sono arbitrari continueremo a riferirci a Σ come al SdR fisso e a S come al sistema mobile.

Supporremo inoltre di conoscere le caratteristiche del moto di S rispetto a Σ , ovvero i vettori $\mathbf{v}(O)$ e $\boldsymbol{\omega}$ (funzioni del tempo) e chiameremo **velocità assoluta**, \mathbf{v}_A , e

velocità relativa, \mathbf{v}_R , la velocità del punto P rispetto a Σ e a S . Riferendoci quindi alla figura 6.4

$$\mathbf{v}_A = \frac{d(P - \Omega)}{dt}, \quad \text{mentre} \quad \mathbf{v}_R = \frac{d(P - O)}{dt}.$$

Per ricavare il legame tra \mathbf{v}_A e \mathbf{v}_R partiamo dalla formula (6.16) che lega fra loro i vettori $(P - \Omega)$ e $(P - O)$. Passando alle terne, consideriamo la (6.17) che deriviamo rispetto al tempo. Questa volta però teniamo conto del fatto che *anche x , le coordinate del vettore $(P - O)$ in S , variano nel tempo*

$$\frac{d}{dt} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \frac{d}{dt} \begin{pmatrix} \xi_O \\ \eta_O \\ \zeta_O \end{pmatrix} + \left(\frac{d}{dt} \mathbb{A}(t) \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \mathbb{A}(t) \frac{d}{dt} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (6.66)$$

Il termine a primo membro nella (6.66) è la terna delle componenti della velocità assoluta del punto P nel sistema Σ . Il terzo addendo a secondo membro rappresenta le componenti, in S , del vettore velocità relativa moltiplicate per la matrice di cambiamento di riferimento: queste sono quindi le componenti in Σ del vettore \mathbf{v}_R , cioè

$$\mathbf{v}_R \rightarrow \begin{cases} \dot{\mathbf{x}}, & \text{nel SdR } S, \\ \mathbb{A}\dot{\mathbf{x}}, & \text{nel SdR } \Sigma. \end{cases}$$

I primi due addendi del secondo membro sono gli stessi che comparivano nella formula fondamentale del moto rigido, quindi possiamo riscrivere la (6.66) nella forma “vettoriale”

$$\mathbf{v}_A(P) = \mathbf{v}_A(O) + \boldsymbol{\omega} \wedge (P - O) + \mathbf{v}_R(P). \quad (6.67)$$

La somma $\mathbf{v}_A(O) + \boldsymbol{\omega} \wedge (P - O)$, rappresenta la velocità che il punto P avrebbe *se fosse solidale*, ovvero se fosse fermo (anche solo “istantaneamente”) rispetto all’osservatore S . Per questa ragione si dà a questo termine il nome di **velocità di trascinamento**, \mathbf{v}_T , e si riscrive la (6.67) nella forma

$$\mathbf{v}_A(P) = \mathbf{v}_T(P) + \mathbf{v}_R(P), \quad (6.68)$$

che si legge dicendo che la velocità assoluta è uguale alla velocità di trascinamento più la velocità relativa.

Nota 6.6.1 Con il “senno di poi” ci possiamo rendere conto che nella derivazione della formula (6.67) si utilizza un’ipotesi non esplicitata: ad ogni istante i due sistemi di riferimento devono essere in grado di confrontare le variazioni rispetto al tempo di una quantità scalare (le componenti del vettore che si sta derivando) nei due sistemi. Questo è possibile solo se i due sistemi possono “sincronizzare” in ogni momento e istantaneamente i loro orologi. Ma ciò è fisicamente impossibile, come messo in luce da Einstein nel suo lavoro (del 1905) sulla Relatività Ristretta.

Il motivo per cui questa ipotesi è rimasta nascosta per secoli dipende dal fatto che la massima velocità con cui due osservatori possono scambiarsi delle informazioni è la velocità della luce, che è molto maggiore di qualsiasi velocità in gioco nei fenomeni meccanici macroscopici. L’errore che quindi si compie tenendo per buona l’ipotesi di sincronizzazione istantanea degli orologi non è quindi rilevabile nell’ambito delle applicazioni meccaniche ordinarie.

6.7 Formula di Poisson

Notiamo ora che le considerazioni fatte per la velocità valgono per **qualsiasi quantità vettoriale**, in quanto il termine di differenza tra $v_A(P)$ e $v_R(P)$ dipende dalla *variazione dell'orientazione degli assi S al variare del tempo*. Useremo i simboli $\frac{d}{dt}\Big|_A$ e $\frac{d}{dt}\Big|_R$ per indicare la derivata rispetto al tempo di una qualsiasi quantità vettoriale come giudicata rispettivamente dagli osservatori Σ e S , e le chiameremo ancora **derivata assoluta e derivata relativa**.

Consideriamo un vettore libero \mathbf{U} definito nel sistema di riferimento solidale col corpo rigido \mathcal{B} . Nel generico SdR S solidale il vettore \mathbf{U} sarà espresso dalla terna \mathbf{u}_S , mentre sarà espresso dalla terna \mathbf{u}_Σ , nel SdR Σ , ovvero

$$\mathbf{U} \rightarrow \begin{cases} \mathbf{u}_S, & \text{nel SdR } S, \\ \mathbf{u}_\Sigma = \mathbb{A}\mathbf{u}_S, & \text{nel SdR } \Sigma. \end{cases}$$

Quindi rispetto all'osservatore Σ , \mathbf{U} varia sia a causa del moto di S , sia a causa della "variazione propria" di \mathbf{U} (ovvero quella "misurata" dall'osservatore solidale S), che dunque sarà

$$\frac{d\mathbf{U}}{dt}\Big|_R \rightarrow \begin{cases} \dot{\mathbf{u}}_S, & \text{nel SdR } S, \\ \dot{\mathbf{u}}_\Sigma|_R = \mathbb{A}\dot{\mathbf{u}}_S, & \text{nel SdR } \Sigma. \end{cases}$$

In accordo con le notazioni introdotte, avremo quindi per un *qualsiasi vettore* \mathbf{U} ,

$$\frac{d\mathbf{U}}{dt}\Big|_A = \frac{d\mathbf{U}}{dt}\Big|_R + \boldsymbol{\omega} \wedge \mathbf{U}, \quad (6.69)$$

che è nota con il nome di **formula di Poisson**³⁰.

Nota 6.7.1 Confrontando la (6.69) con la (6.67) ci si potrebbe chiedere perché nella (6.69) non compare $v_A(O)$. Osserviamo che la (6.67) è la derivata assoluta del vettore $(P - O)$ che si scompone come nella (6.16). Calcolando la derivata assoluta del vettore $(O - O)$ otteniamo $v_A(O)$, mentre per calcolare la derivata assoluta del vettore $(P - O)$ si applica la (6.69) dove $\mathbf{U} = (P - O)$. Si otteniamo così la (6.67) dal momento che $\frac{d\mathbf{U}}{dt}\Big|_R$ è proprio la velocità relativa v_R del punto P .

Per dimostrare la (6.69) consideriamo la derivata di \mathbf{u}_Σ rispetto al SdR Σ , che denoteremo con

$$\dot{\mathbf{u}}_\Sigma|_A = \frac{d}{dt}(\mathbb{A}\mathbf{u}_S).$$

³⁰Siméon-Denis Poisson (Pithiviers, 1781 – Parigi, 1840), matematico, fisico, astronomo e statistico francese.

Questa ci darà appunto le componenti della derivata di \mathbf{U} rispetto al SdR fisso Σ . Abbiamo

$$\begin{aligned}\dot{\mathbf{u}}_{\Sigma}|_A &= \frac{d\mathbb{A}}{dt}\mathbf{u}_S + \mathbb{A}\dot{\mathbf{u}}_S = \left(\frac{d\mathbb{A}}{dt}\mathbb{A}^T\right)\mathbb{A}\mathbf{u}_S + \mathbb{A}\dot{\mathbf{u}}_S \\ &= \boldsymbol{\omega} \wedge \underbrace{\mathbb{A}\mathbf{u}_S}_{\mathbf{u}_{\Sigma}} + \underbrace{\mathbb{A}\dot{\mathbf{u}}_S}_{\dot{\mathbf{u}}_{\Sigma}|_R} \\ &= \boldsymbol{\omega} \wedge \mathbf{u}_{\Sigma} + \dot{\mathbf{u}}_{\Sigma}|_R,\end{aligned}\tag{6.70}$$

dove $\boldsymbol{\omega}$ è la terna delle componenti della velocità angolare nel SdR Σ . Ora, come ampiamente illustrato nella sezione 6.3, si può far vedere che (6.70) è indipendente dagli specifici sistemi di riferimento considerati, e quindi concludere che la (6.70) è proprio l'espressione nel SdR Σ dell'identità vettoriale (6.69).

In particolare, l'espressione del vettore $\left.\frac{d\mathbf{U}}{dt}\right|_A$ rispetto al SdR S è data da

$$\begin{aligned}\mathbb{A}^T \dot{\mathbf{u}}_{\Sigma}|_A &= \mathbb{A}^T \frac{d\mathbb{A}}{dt} \underbrace{(\mathbb{A}^T \mathbb{A})}_{\mathbb{I}} \mathbf{u}_S + \underbrace{(\mathbb{A}^T \mathbb{A})}_{\mathbb{I}} \dot{\mathbf{u}}_S \\ &= \mathbb{A}^T \frac{d\mathbb{A}}{dt} \mathbf{u}_S + \dot{\mathbf{u}}_S \\ &= \boldsymbol{\omega}_S \wedge \mathbf{u}_S + \dot{\mathbf{u}}_S,\end{aligned}\tag{6.71}$$

dove, ricordando la (6.51), $\boldsymbol{\omega}_S$ è la terna delle componenti della velocità angolare nel SdR S .

Nota 6.7.2 Mostriamo come ottenere la (6.71) partendo da $\mathbf{u}_S = \mathbb{A}^T \mathbf{u}_{\Sigma}$ e derivandola rispetto al tempo. Abbiamo infatti

$$\dot{\mathbf{u}}_S = \left(\frac{d\mathbb{A}^T}{dt}\right)\mathbf{u}_{\Sigma} + \mathbb{A}_{\Sigma}^T \dot{\mathbf{u}}_{\Sigma} = \left(\frac{d\mathbb{A}^T}{dt}\right)\mathbb{A}\mathbb{A}^T \mathbf{u}_{\Sigma} + \mathbb{A}_{\Sigma}^T \dot{\mathbf{u}}_{\Sigma},$$

dove $\dot{\mathbf{u}}_{\Sigma}$ va ovviamente intesa come $\dot{\mathbf{u}}_{\Sigma}|_A$, essendo la derivata della terna \mathbf{u}_{Σ} costituita dalle componenti del vettore \mathbf{U} nel SdR Σ . Adesso si sfrutta la (6.5) che, come sappiamo, derivata rispetto al tempo fornisce

$$\left(\frac{d\mathbb{A}^T}{dt}\right)\mathbb{A} + \mathbb{A}^T \left(\frac{d\mathbb{A}}{dt}\right) = 0,$$

e quindi

$$\dot{\mathbf{u}}_S = -\mathbb{A}^T \frac{d\mathbb{A}}{dt} \underbrace{(\mathbb{A}^T \mathbf{u}_{\Sigma})}_{\mathbf{u}_S} + \mathbb{A}_{\Sigma}^T \dot{\mathbf{u}}_{\Sigma} \stackrel{(6.52)}{=} -\boldsymbol{\omega}_S \wedge \mathbf{u}_S + \mathbb{A}^T \dot{\mathbf{u}}_{\Sigma}|_A,$$

da cui la (6.71) discende banalmente.

Nota 6.7.3 Calcoliamo la derivata rispetto a Σ della velocità angolare. Se $\boldsymbol{\omega}$ è la terna delle componenti della velocità angolare rispetto a Σ , abbiamo $\dot{\boldsymbol{\omega}}|_A = \dot{\boldsymbol{\omega}}$. Se

invece consideriamo le componenti rispetto al SdR S , cioè la terna ω_S , sfruttando la (6.51), abbiamo

$$\begin{aligned} \dot{\omega}_S|_A &= \frac{d}{dt} (\mathbb{A}^T \omega) = \underbrace{\mathbb{A}^T \mathbb{A}}_I \left(\frac{d\mathbb{A}^T}{dt} \right) \omega + \mathbb{A}^T \dot{\omega} \\ &\stackrel{(6.25)}{=} \mathbb{A}^T \underbrace{\left(\mathbb{A} \left(\frac{d\mathbb{A}^T}{dt} \right) \right)}_{-\frac{d\mathbb{A}}{dt} \mathbb{A}^T} \omega + \mathbb{A}^T \dot{\omega}|_A \stackrel{(6.52)}{=} \underbrace{-\mathbb{A}^T \frac{d\mathbb{A}}{dt}}_{\omega_S} \underbrace{\mathbb{A}^T \omega}_{\omega_S} + \mathbb{A}^T \dot{\omega}|_A \\ &= \mathbb{A}^T \dot{\omega}|_A, \end{aligned}$$

e quindi $\dot{\omega}|_A = \mathbb{A} \dot{\omega}_S|_A$.

6.8 Composizione di moti rigidi

Siano dati adesso tre osservatori Σ , S e S' , con S' in moto sia rispetto a Σ che a S , e siano note le caratteristiche del moto di S rispetto a Σ , $v(O)$ e ω , e di S' rispetto³¹ a S , $v_R(O')$ e ω' . Vogliamo determinare il moto di S' rispetto a Σ .

Se P è un punto solidale con S' , la velocità di P rispetto a S è

$$v_R(P) = v_R(O') + \omega' \wedge (P - O').$$

Indichiamo con $v_A(P)$ la velocità di P rispetto all'osservatore Σ . Dalla (6.67) otteniamo

$$\begin{aligned} v_A(P) &= v_A(O) + \omega \wedge (P - O) + v_R(P) \\ &= v_A(O) + \omega \wedge (P - O) + v_R(O') + \omega' \wedge (P - O'), \end{aligned}$$

che riscriviamo

$$\begin{aligned} v_A(P) &= v_A(O) + \omega \wedge (P - O' + O' - O) + v_R(O') + \omega' \wedge (P - O') \\ &= \underbrace{v_A(O) + \omega \wedge (O - O') + v_R(O')}_{v_A(O')} + \\ &\quad \omega \wedge (P - O') + \omega' \wedge (P - O'). \end{aligned} \tag{6.72}$$

Il trinomio $v_A(O) + \omega \wedge (O - O') + v_R(O')$ rappresenta la velocità del punto O' rispetto al SdR Σ , che, secondo la nostra convenzione, abbiamo denotato con $v_A(O')$, velocità di O' (centro del SdR S') rispetto all'osservatore fisso Σ . Possiamo pertanto riscrivere la (6.72)

$$v_A(P) = v_A(O') + (\omega' + \omega) \wedge (P - O'),$$

³¹Con O' si è ovviamente indicato l'origine di S' , e si è usata la notazione $v_R(O')$ per evidenziare che questa è la velocità di O' valutata dall'osservatore S , cioè $v_R(O') = (O - O')$.

che ci dice che il sistema S' compie rispetto a Σ un moto rigido **di velocità angolare** ($\omega' + \omega$). Quindi **nella composizione di due moti rigidi, le velocità angolari si sommano**. Lo stesso vale se si compongono tre o più moti rigidi.

Un'applicazione immediata della formula di addizione delle velocità angolari si ha in un metodo per la determinazione di ω a partire dagli angoli di Eulero (metodo diverso da quello illustrato nella parte conclusiva del paragrafo 6.4). Da come abbiamo definito gli angoli di Eulero (v. sezione 6.4) risulta chiaro che la generica trasformazione che porta Σ' a coincidere con S è ottenuta tramite le tre rotazioni consecutive: (i) rotazione di ψ attorno all'asse e_3 (cioè l'asse ζ del SdR Σ'); (ii) rotazione di θ attorno alla linea dei nodi il cui versore è \mathbf{n} ; (iii) rotazione di ϕ attorno all'asse z (asse \mathbf{k} del SdR S). Ricordando la forma della velocità angolare in una rotazione, ovvero la (6.49), e utilizzando l'addizione delle velocità angolari, si ha la semplice espressione

$$\omega = \dot{\psi} e_3 + \dot{\theta} \mathbf{n} + \dot{\phi} \mathbf{k}. \quad (6.73)$$

La (6.73) è però un'espressione "ibrida" perché non ci dà le componenti di ω né nel sistema Σ , né nel sistema S . La (6.73) ha un altro difetto: mentre \mathbf{n} è perpendicolare³² sia a e_3 che a \mathbf{k} , questi ultimi due versori non sono in genere perpendicolari tra loro, quindi il modulo di ω non può essere ottenuto da (6.73) semplicemente sommando i quadrati di $\dot{\psi}$, $\dot{\theta}$, e $\dot{\phi}$. E' tuttavia facile ottenere l'espressione di ω in entrambi i riferimenti Σ e S , osservando che

$$\begin{aligned} \mathbf{n} &= \cos \psi e_1 + \sin \psi e_2, \\ \mathbf{n} &= \cos \phi \mathbf{i} - \sin \phi \mathbf{j}, \\ \mathbf{k} &= \sin \theta (\sin \psi e_1 - \cos \psi e_2) + \cos \theta e_3, \\ e_3 &= \sin \theta (\sin \phi \mathbf{i} + \cos \phi \mathbf{j}) + \cos \theta \mathbf{k}. \end{aligned}$$

Si ottengono così le formule (6.55) e (6.56).

Esempio 6.8.1 *La formula della composizione delle velocità angolare è estremamente utile quando si deve calcolare la velocità angolare di moti rigidi tridimensionali. Riprendendo l'esempio 4.1.3, dimostriamo la formula (4.19) facendo uso del metodo della composizione della velocità angolare.*

Riferendoci alla figura 6.10, denotiamo con $\Sigma : \{\Omega, x, y, z\}$ il sistema di riferimento fisso. Consideriamo il SdR $S : \{O, X, Y, Z\}$ solidale con il vettore $(Q - O)$. Tale SdR ruota attorno all'asse $z \equiv Z$, e quindi la velocità angolare di S rispetto a Σ è $\omega_1 = \dot{\theta} e_z$. Consideriamo poi il SdR solidale con il disco che rotola, cioè $S' : \{Q, X', Y', Z'\}$. Quest'ultimo ruota attorno all'asse $X \equiv X'$ e quindi la sua velocità angolare rispetto ad S è $\omega_2 = \dot{\varphi} e_{X'}$. La velocità angolare del SdR solidale S' rispetto al SdR fisso Σ (e dunque la velocità angolare del disco rigido rispetto a Σ) è quindi

$$\omega = \omega_1 + \omega_2 = \dot{\varphi} e_{X'} + \dot{\theta} e_z,$$

che possiamo esprimere sia nel SdR Σ che nel SdR solidale S' o, eventualmente, nel SdR "semisolidale" S . Se decidiamo di esprimere ω in Σ , allora otteniamo la (4.19) siccome

$$e_{X'} \equiv e_X = \cos \theta e_x + \sin \theta e_y.$$

³²Ricordiamo che la linea dei nodi è l'intersezione dei piani (ξ, η) e (x, y) , e quindi è ortogonale sia all'asse z , che all'asse ζ .

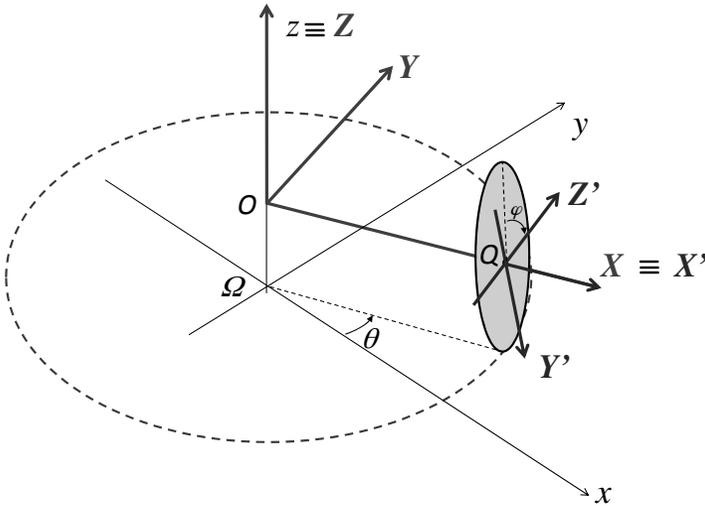


Figura 6.10: $\{\Omega, x, y, z\}$ è il SdR fisso Σ . Il sistema di riferimento solidale con il disco che rotola è $\{Q, X', Y', Z'\}$.

Viceversa se decidiamo di esprimere ω nel SdR solidale S' , allora dobbiamo esprimere e_z nel SdR solidale S' . Guardando la figura 6.10 è facile rendersi conto che

$$e_z = \cos \varphi e_{Z'} - \sin \varphi e_{Y'} ,$$

per cui

$$\omega = \dot{\varphi} e_{X'} - \dot{\theta} \sin \varphi e_{Y'} + \dot{\theta} \cos \varphi e_{Z'} . \tag{6.74}$$

6.9 Cinematica relativa: l'accelerazione

Vediamo infine come viene giudicata da due osservatori diversi l'accelerazione di un punto in movimento rispetto a entrambi. Per trovare la relazione che lega le due accelerazioni, basta derivare la formula (6.67) che lega la velocità relativa alla velocità assoluta tenendo presente che l'accelerazione giudicata da Σ (che chiameremo assoluta) è la derivata in Σ della velocità giudicata da Σ , mentre l'accelerazione giudicata da S (che chiameremo relativa) è la derivata in S della velocità giudicata da S

Avremo quindi che l'accelerazione del punto P "misurata" dall'osservatore Σ è

$$a_A(P) = \left. \frac{d}{dt} v_A(P) \right|_A \stackrel{(6.67)}{=} \left. \frac{d}{dt} [v_A(O) + \omega \wedge (P - O) + v_R(P)] \right|_A . \tag{6.75}$$

Calcoliamo la derivata a secondo membro

$$a_A(P) = \left. \frac{d}{dt} v_A(O) \right|_A + \left. \frac{d}{dt} \omega \wedge (P - O) \right|_A + \left. \frac{d}{dt} v_R(P) \right|_A ,$$

utilizzando la formula di Poisson (6.69) per il secondo e terzo addendo. Il primo addendo è chiaramente l'accelerazione del punto O giudicata dall'osservatore Σ . Otteniamo dunque:

- $\left. \frac{d}{dt} \mathbf{v}_A(O) \right|_A = \mathbf{a}_A(O)$, accelerazione del punto O rispetto al SdR Σ ;
- $\left. \frac{d}{dt} \mathbf{v}_R(P) \right|_A = \boldsymbol{\omega} \wedge \mathbf{v}_R(P) + \left. \frac{d}{dt} \mathbf{v}_R(P) \right|_R = \boldsymbol{\omega} \wedge \mathbf{v}_R(P) + \mathbf{a}_R(P)$, dove $\mathbf{a}_R(P)$ è l'accelerazione di P valutata dall'osservatore S ;
- $\left. \frac{d}{dt} \boldsymbol{\omega} \wedge (P - O) \right|_A = \left. \frac{d}{dt} \boldsymbol{\omega} \right|_A \wedge (P - O) + \boldsymbol{\omega} \wedge \left. \frac{d}{dt} (P - O) \right|_A$. La derivata di $\boldsymbol{\omega}$, valutata dall'osservatore Σ , viene denotata con $\dot{\boldsymbol{\omega}}$, cioè $\dot{\boldsymbol{\omega}} = \left. \frac{d}{dt} \boldsymbol{\omega} \right|_A$. Abbiamo quindi

$$\begin{aligned} \left. \frac{d}{dt} \boldsymbol{\omega} \wedge (P - O) \right|_A &= \dot{\boldsymbol{\omega}} \wedge (P - O) + \boldsymbol{\omega} \wedge \left. \frac{d}{dt} (P - O) \right|_A \\ &= \dot{\boldsymbol{\omega}} \wedge (P - O) + \boldsymbol{\omega} \wedge [\boldsymbol{\omega} \wedge (P - O) + \mathbf{v}_R(P)]. \end{aligned}$$

Possiamo riunire tutti i “pezzi” e otteniamo il legame cercato tra le due accelerazioni

$$\mathbf{a}_A(P) = \mathbf{a}_R(P) + \mathbf{a}_A(O) + \dot{\boldsymbol{\omega}} \wedge (P - O) + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge (P - O)) + 2 \boldsymbol{\omega} \wedge \mathbf{v}_R(P). \quad (6.76)$$

La (6.76) va letta osservando che il secondo, terzo e quarto termine della somma sono presenti *anche quando il punto P è fermo rispetto ad S* ; per questa ragione la loro somma prende il nome di accelerazione di trascinamento, $\mathbf{a}_T(P)$. Infine il termine $2 \boldsymbol{\omega} \wedge \mathbf{v}_R(P)$, indicato generalmente con il simbolo $\mathbf{a}_C(P)$, è dovuto alla correzione necessaria per il diverso modo che hanno i due osservatori di giudicare sia la variazione della velocità relativa sia la variazione della velocità di trascinamento. Questo termine prende il nome di **accelerazione complementare** o **accelerazione di Coriolis**³³. La (6.76) si può quindi scrivere nella forma compatta

$$\mathbf{a}_A = \mathbf{a}_R + \mathbf{a}_T + \mathbf{a}_C. \quad (6.77)$$

L'importanza di questa relazione è dovuta al fatto che molto spesso i fenomeni meccanici sono osservati da sistemi di riferimento non inerziali. Esempi noti sono la deviazione dalla verticale nella caduta di un grave³⁴, la deviazione dalla direzione dei meridiani dei venti alisei e il moto del pendolo di Foucault.

³³Gaspard-Gustave de Coriolis (Parigi, 1792 – Parigi, 1843) matematico, fisico e ingegnere francese.

³⁴Si noti che nella direzione della verticale, che è quella del filo a piombo, è compresa oltre alla forza di gravità, un termine centrifugo dovuto all'accelerazione di trascinamento della terra. Quando un grave cade, il termine di accelerazione complementare fa deviare la sua traiettoria dalla verticale. Se si guarda al fenomeno dal punto di vista di un osservatore posto al di fuori della terra, questa deviazione è l'ovvia conseguenza del fatto che nel tempo impiegato nella caduta, l'osservatore (la terra) ha ruotato spostando il piede della verticale del punto di partenza verso oriente. L'osservatore solidale con la terra vedrà quindi deviare il grave verso occidente. Tuttavia questo effetto è molto piccolo, e praticamente inosservabile, nella caduta di un grave da altezza ordinarie.

Capitolo 7

Dinamica dei Sistemi Rigidi

Abbiamo già visto nel capitolo 6 che un sistema di punti (P_k, m_k) , è detto **rigido** se è tenuto a rispettare (oltre a eventuali altri vincoli) il vincoli di rigidità, cioè

$$|P_i - P_j|^2 = d_{ij}^2, \quad \forall i \neq j, \quad \text{con } d_{ij} \text{ costante nel tempo.} \quad (7.1)$$

Sempre nel capitolo dedicato alla cinematica abbiamo osservato che, se il sistema possiede almeno tre punti non allineati, allora le sue configurazioni sono in corrispondenza uno a uno con le posizioni che un sistema di riferimento ortonormale (solidale) può assumere rispetto a un altro sistema di riferimento ortogonale (fisso).

Un'importante conseguenza è che, almeno dal punto di vista cinematico, non ci sono differenze nella descrizioni del moto di un sistema rigido “discreto”, cioè formato da un numero finito di punti materiali, e un **corpo rigido continuo**¹. In quello che segue tratteremo il problema della dinamica di un sistema rigido discreto. Le equazioni di moto che ricaveremo saranno però facilmente “adattabili” al caso continuo. Esse saranno quindi adottate come modello per la dinamica dei corpi rigidi continui.

Il vincolo di rigidità, che, come visto nell'esempio 5.2.1, è liscio, è mantenuto da opportune forze (o reazioni) vincolari che risultano incognite. In questo capitolo ricaveremo una coppia equazioni vettoriali, dette **prima equazione cardinale** e **seconda equazione cardinale**, che non contengono le forze vincolari interne relative al vincolo di rigidità. Per la precisione le equazioni cardinali sono valide anche per sistemi non rigidi. La loro caratteristica è proprio quella che in esse non compaiono le forze interne al sistema ma soltanto le forze esterne (daremo più avanti la definizione di forze interne ed esterne).

Vedremo poi che, nel caso di sistemi rigidi, la seconda equazione cardinale acquista una struttura particolare a causa della particolare forma della velocità dei punti di un corpo rigido (si veda il paragrafo 6.3 del capitolo 6).

Concludiamo sottolineando una proprietà fondamentale delle equazioni cardinali che verrà dimostrata in questo capitolo: proveremo che, nel caso di un sistema rigido,

¹Anche se è abbastanza chiaro cosa si debba, intuitivamente, intendere per corpo rigido continuo, è bene dare una definizione più formale. Diremo che \mathfrak{B} un corpo continuo se esso occupa una regione $\mathcal{C}(t)$, ed esiste una funzione continua, non negativa $\rho(x, y, z, t)$ di cui $\mathcal{C}(t)$ è il supporto al tempo t , ovvero se $\rho(x, y, z, t) > 0$ per $(x, y, z) \in \mathcal{C}(t)$ e nulla altrimenti: la funzione ρ è detta **densità di massa**, e rappresenta la massa per unità di volume. Il corpo si dice rigido se esiste un sistema di riferimento in cui la *funzione densità ρ non dipende dal tempo*.

le due equazioni cardinali formano un sistema di 6 equazioni scalari indipendenti cioè tante quanti sono i gradi di libertà de sistema (sezione 6.2.3 del capitolo 6). Le due equazioni cardinali sono quindi sufficienti a determinare la dinamica del corpo rigido (se, ovviamente, sono note tutte le forze esterne agenti su di esso).

7.1 Il centro di massa

Sia dato in uno spazio affine euclideo A un sistema \mathfrak{B} di N punti materiali (P_k, m_k) , $k = 1, \dots, N$, (anche non rigido). Si fissa poi un punto O nello spazio A .

Definizione 7.1.1 Si dice **centro di massa**, o **baricentro**, del sistema \mathfrak{B} il punto P_o dato da

$$(P_o - O) = \frac{\sum_{k=1}^N m_k (P_k - O)}{M}, \quad (7.2)$$

dove $M = \sum_{k=1}^N m_k$, è **la massa totale del sistema**.

Il centro di massa è un'entità puramente concettuale: in generale P_o non coincide con un punto materiale P_k di \mathfrak{B} . Tuttavia P_o , pur non coincidendo in generale con alcun punto materiale, è una quantità intrinseca di \mathfrak{B} : dipende cioè soltanto dalla struttura geometrico-materiale del sistema. Vale infatti la seguente

Proposizione 7.1.1 Il centro di massa P_o non dipende dalla scelta del punto O .

Dim. Per dimostrare la proposizione dobbiamo fare vedere in pratica che, scegliendo un punto $O' \neq O$, il vettore che si ottiene applicando la definizione (7.2), cioè

$$(P'_o - O') = \frac{\sum_{k=1}^N m_k (P_k - O')}{M},$$

identifica lo stesso punto P_o .

Riferendoci alla figura 7.1, per provare che $P'_o \equiv P_o$, dobbiamo mostrare che

$$(P'_o - O') + (O' - O) = (P_o - O). \quad (7.3)$$

Partendo dalla (7.2) abbiamo

$$\begin{aligned} (P_o - O) &= \frac{1}{M} \sum_{k=1}^N m_k (P_k - O) = \frac{1}{M} \sum_{k=1}^N m_k [(P_k - O') + (O' - O)] \\ &= \underbrace{\frac{1}{M} \sum_{k=1}^N m_k (P_k - O')}_{(P'_o - O')} + \underbrace{\frac{1}{M} \sum_{k=1}^N m_k (O' - O)}_{\frac{1}{M} (\sum_{k=1}^N m_k) (O' - O)} \\ &= (P'_o - O') + (O' - O), \end{aligned}$$

cioè la (7.3).

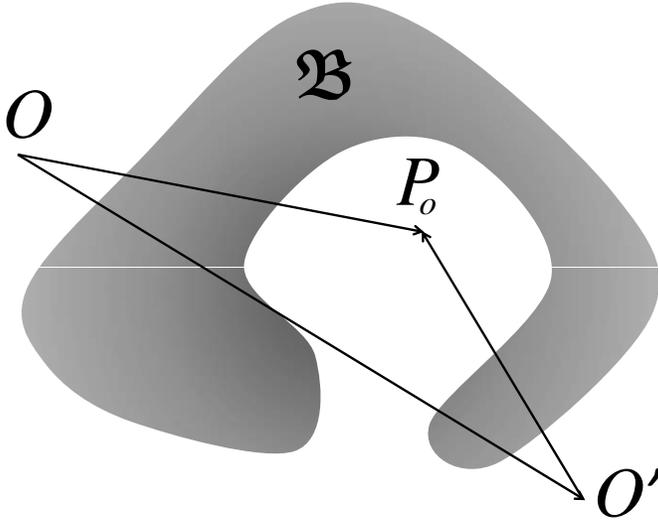


Figura 7.1: Il sistema materiale \mathfrak{B} , ed il centro di massa P_o , che, in questo caso, non coincide con alcun punto di \mathfrak{B} .

□

Una diretta conseguenza di tale proposizione è che, ai fini del calcolo di P_o , conviene porsi nel sistema di riferimento dove i calcoli sono i più semplici possibili. Quindi, se \mathfrak{B} è un corpo rigido conviene sempre considerare un sistema di riferimento solidale con il corpo.

Un'altra proprietà utile ai fini del calcolo è l'additività rispetto alle masse, ovvero la proprietà distributiva. Supponiamo di dividere \mathfrak{B} in due parti $\mathfrak{B}_1, \mathfrak{B}_2$ disgiunte, cioè $\mathfrak{B} = \mathfrak{B}_1 \cup \mathfrak{B}_2$, con $\mathfrak{B}_1 \cap \mathfrak{B}_2 = \emptyset$. Quindi se $\mathfrak{B}_1 = \{(P_k, m_k) : k = 1, \dots, N_1\}$, tale che $M_1 = \sum_{k=1}^{N_1} m_k$ è la massa parziale di \mathfrak{B}_1 , e $\mathfrak{B}_2 = \{(P_k, m_k) : k = N_1 + 1, \dots, N\}$, con $M_2 = \sum_{k=N_1+1}^N m_k$, abbiamo, con evidente significato dei simboli

$$\begin{aligned}
 (P_o - O) &= \frac{1}{M} \sum_{k=1}^N m_k (P_k - O) \\
 &= \frac{1}{M} \left[\sum_{k=1}^{N_1} m_k (P_k - O) + \sum_{k=N_1+1}^N m_k (P_k - O) \right] \\
 &= \frac{1}{M} [M_1 (P_o, \mathfrak{B}_1 - O) + M_2 (P_o, \mathfrak{B}_2 - O)] .
 \end{aligned}$$

Il centro di massa dell'intero sistema \mathfrak{B} coincide quindi col centro di massa dei due punti materiali P_o, \mathfrak{B}_1 e P_o, \mathfrak{B}_2 , aventi masse M_1 e M_2 , che, a loro volta, sono i centri

di massa “parziali” dei sottosistemi \mathfrak{B}_1 e \mathfrak{B}_2 . Evidentemente tale procedura si può generalizzare, supponendo di suddividere \mathfrak{B} in L sottosistemi \mathfrak{B}_i , $i = 1, \dots, L$, disgiunti di masse M_i . Avremo quindi

$$(P_o - O) = \frac{1}{M} \sum_{i=1}^L M_i (P_o, \mathfrak{B}_i - O).$$

Nota 7.1.1 Se \mathfrak{B} è un corpo rigido continuo e $\mathcal{C}(t)$ è la regione da esso occupata, il centro di massa di \mathfrak{B} è così definito

$$\mathbf{x}_{P_o} = \frac{1}{M} \int_{\mathcal{C}(t)} \mathbf{x} dm = \frac{1}{M} \int_{\mathcal{C}(t)} \mathbf{x} \rho(\mathbf{x}) dV,$$

dove:

- \mathbf{x} è il vettore posizione del generico punto di $\mathcal{C}(t)$;
- dm è la misura di massa che viene espressa come ρdV , con ρ massa per unità di volume (densità di massa) e dV misura di volume.

La massa totale del corpo è

$$M = \int_{\mathcal{C}(t)} dm = \int_{\mathcal{C}(t)} \rho(\mathbf{x}) dV.$$

Introduciamo la seguente

Definizione 7.1.2 Siano \mathbf{v}_k , $k = 1, \dots, N$, le velocità dei punti materiali di \mathfrak{B} rispetto ad un dato SdR S . Si dice **velocità del centro di massa** rispetto ad S , il vettore

$$\mathbf{v}(P_o) = \frac{1}{M} \sum_{k=1}^N m_k \mathbf{v}_k. \quad (7.4)$$

Talvolta $\mathbf{v}(P_o)$, viene anche denotato con $(P_o - O)$. Inoltre il vettore

$$\mathbf{K} = M \mathbf{v}(P_o) = \sum_{k=1}^N m_k \mathbf{v}_k, \quad (7.5)$$

viene detto **quantità di moto**, o **momento lineare**, o anche semplicemente **momento**, del sistema di \mathfrak{B} rispetto al SdR S .

Concludiamo osservando che, nel caso in cui le masse degli N punti materiali del sistema siano costanti nel tempo, abbiamo

$$\frac{d\mathbf{K}}{dt} = \sum_{k=1}^N m_k \frac{d\mathbf{v}_k}{dt} = \sum_{k=1}^N m_k \mathbf{a}_k, \quad (7.6)$$

dove \mathbf{a}_k denota l'accelerazione del k -esimo punto di \mathfrak{B} rispetto al SdR S . In particolare, se definiamo l'**accelerazione del centro di massa** rispetto al SdR S , come

$\mathbf{a}(P_o) = (P_o - O) = \frac{1}{M} \sum_{k=1}^N m_k \mathbf{a}_k$, la (7.6) si può scrivere come

$$\frac{d\mathbf{K}}{dt} = M \mathbf{a}(P_o).$$

7.2 Il momento angolare

Consideriamo un sistema \mathfrak{B} di N punti materiali (non necessariamente rigido) ed un generico punto Q (non necessariamente coincidente con un punto materiale del sistema). Adesso, seguendo la procedura della sezione 1.5.2, consideriamo il sistema di vettori applicati

$$\mathcal{S} = \{(P_k, \mathbf{v}_k) \in A \times V : k = 1, \dots, N\},$$

dove $\mathbf{v}_k, k = 1, \dots, N$, è la velocità del punto P_k , si introduce il momento del sistema rispetto al polo, o centro di riduzione, Q , che, in questo specifico caso viene detto **momento angolare**. Abbiamo cioè la seguente

Definizione 7.2.1 Si dice **momento angolare**, o **momento della quantità di moto**, del sistema \mathfrak{B} rispetto al centro di riduzione Q , il vettore

$$\mathbf{L}(Q) = \sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{v}_k. \tag{7.7}$$

Il punto Q viene talvolta anche detto polo.

A differenza del centro di massa, il momento angolare $\mathbf{L}(Q)$ dipende dalla struttura geometrico-materiale del sistema \mathfrak{B} , ma anche dal centro di riduzione. Se infatti consideriamo $D \neq Q$, abbiamo

$$\mathbf{L}(D) = \sum_{k=1}^N m_k (P_k - D) \wedge \mathbf{v}_k,$$

che, in generale, è diverso da $\mathbf{L}(Q)$. Infatti

$$\begin{aligned} \mathbf{L}(D) &= \sum_{k=1}^N m_k (P_k - D) \wedge \mathbf{v}_k = \sum_{k=1}^N m_k [(P_k - Q) + (Q - D)] \wedge \mathbf{v}_k \\ &= \underbrace{\sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{v}_k}_{\mathbf{L}(Q)} + \underbrace{\sum_{k=1}^N m_k (Q - D) \wedge \mathbf{v}_k}_{(Q-D) \wedge \sum_{k=1}^N m_k \mathbf{v}_k}. \end{aligned}$$

Quindi, ricordando la (7.5), abbiamo

$$\mathbf{L}(D) = \mathbf{L}(Q) + (Q - D) \wedge \mathbf{K},$$

che altro non è che la (1.20), già incontrata nella sezione 1.5.2. Quindi $\mathbf{L}(D) = \mathbf{L}(Q)$, soltanto se $(Q - D)$ e \mathbf{K} sono paralleli oppure la velocità del centro di massa è nulla, sì che $\mathbf{K} = 0$.

7.3 Geometria delle masse

In questa sezione verrà introdotto il momento d'inerzia ed altri concetti ad esso collegati. Verranno poi presentati alcuni esempi riguardanti il calcolo di momenti d'inerzia,

la determinazione della terna principale d'inerzia e proprietà della matrice d'inerzia. L'applicazione di queste grandezze alla dinamica dei sistemi rigidi verrà presentata nelle sezioni 7.6 ed in quelle successive.

7.3.1 Momenti d'inerzia

Sia \mathfrak{B} un sistema di N punti materiali (P_i, m_i) , $i = 1, 2, \dots, N$.

Definizione 7.3.1 *Data una retta r , il cui versore è e_r , passante per il punto Q , la grandezza*

$$I_r(Q) = \sum_{i=1}^N m_i [(P_i - Q) \wedge e_r]^2,$$

si dice **momento d'inerzia del sistema \mathfrak{B} rispetto ad r** .

Riferendosi alla figura 7.2, $[(P_i - Q) \wedge e_r]^2$, rappresenta la distanza al quadrato di P_i dalla retta r passante per Q (il segmento $\overline{P_i R_i}$). Pertanto, la definizione è indipendente dalla scelta del punto $Q \in r$.

E' facile vedere che $I_r(Q) > 0$, e $I_r(Q) = 0$ soltanto nel caso particolare in cui tutti i punti P_i giacciono sulla retta r .

Nel caso di un sistema continuo che occupa un dominio \mathcal{D} , la somma si sostituisce con l'integrale, ovvero

$$I_r(Q) = \int_{\mathcal{D}} [(P - Q) \wedge e_r]^2 dm = \int_{\mathcal{D}} [(P - Q) \wedge e_r]^2 \rho dV,$$

dove, come nell'osservazione 7.1.1, dm è la misura di massa che viene espressa come ρdV , con ρ densità di massa e dV misura di volume.

Definizione 7.3.2 *Siano dati due piani π_1 e π_2 non paralleli, i cui versori normali sono e_{π_1} e e_{π_2} . Sia r la retta tale che $r = \pi_1 \cap \pi_2$, e sia Q un punto di r . Si definisce il **momento d'inerzia di \mathfrak{B} rispetto ai piani π_1 e π_2** come*

$$I_{\pi_1 \pi_2}(Q) = - \sum_{i=1}^N m_i [(P_i - Q) \cdot e_{\pi_1}] [(P_i - Q) \cdot e_{\pi_2}].$$

Analogamente alla definizione di $I_r(Q)$, è facile dimostrare che $I_{\pi_1 \pi_2}(Q)$ non dipende dalla scelta del punto Q su r . Infatti, $[(P_i - Q) \cdot e_{\pi_1}]$ (e allo stesso modo anche $[(P_i - Q) \cdot e_{\pi_2}]$) rappresenta la "quota" (che può essere quindi positiva, negativa o nulla) del punto P_i rispetto a π_1 .

Come nel caso precedente, se il sistema è continuo la somma andrà sostituita con un integrale (definito sul dominio \mathcal{D} occupato da \mathfrak{B}) che, al solito, farà intervenire ρdV .

Teorema 7.3.1 (Teorema di Huygens) *Sia dato un sistema di punti N materiali \mathfrak{B} di cui P_o è il centro di massa. Si ha*

$$I_r(Q) = I_{r_{P_o}}(P_o) + M [(P_o - Q) \wedge e_r]^2, \quad (7.8)$$

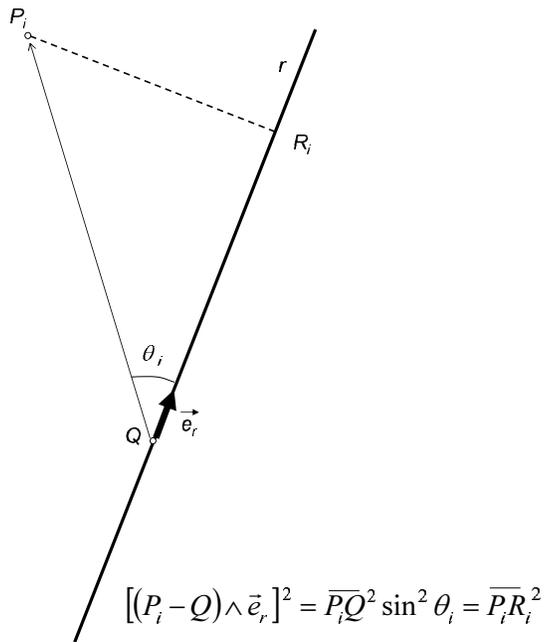


Figura 7.2: Retta r e distanza di P_i da r .

e

$$I_{\pi_1 \pi_2}(Q) = I_{\pi_1, P_o, \pi_2, P_o}(P_o) - M[(P_o - Q) \cdot e_{\pi_1}][(P_o - Q) \cdot e_{\pi_2}], \quad (7.9)$$

dove con r_{P_o} si intende una retta parallela ad r , passante per P_o , e con $\pi_{1, P_o}, \pi_{2, P_o}$ i due piani, paralleli rispettivamente a π_1 e π_2 , la cui retta d'intersezione passa per P_o .

Dim. Si nota che (7.8) si può anche scrivere come

$$I_r(Q) = I_{r_{P_o}}(P_o) + Md^2,$$

essendo d la distanza fra le retta r e la retta r_{P_o} , ovvero $d = |(P_o - Q) \wedge e_r|$.

Ricordando la definizione di centro di massa (7.2), abbiamo

$$\begin{aligned}
 I_r(Q) &= \sum_{i=1}^N m_i [(P_i - Q) \wedge \mathbf{e}_r]^2 = \sum_{i=1}^N m_i [((P_i - P_o) + (P_o - Q)) \wedge \mathbf{e}_r]^2 \\
 &= \underbrace{\sum_{i=1}^N m_i [(P_i - P_o) \wedge \mathbf{e}_r]^2}_{I_{rP_o}(P_o)} + \underbrace{\left(\sum_{i=1}^N m_i \right)}_M \underbrace{[(P_o - Q) \wedge \mathbf{e}_r]^2}_{d^2} + \\
 &\quad + 2 \left\{ \sum_{i=1}^N m_i [(P_i - P_o) \wedge \mathbf{e}_r] \right\} \cdot [(P_o - Q) \wedge \mathbf{e}_r].
 \end{aligned}$$

L'ultimo termine si annulla per definizione di centro di massa,

$$\begin{aligned}
 \sum_{i=1}^N m_i [(P_i - P_o) \wedge \mathbf{e}_r] &= \left\{ \sum_{i=1}^N m_i [(P_i - O) - (P_o - O)] \right\} \wedge \mathbf{e}_r \\
 &= \underbrace{\left[\sum_{i=1}^N m_i (P_i - O) \right]}_{M(P_o - O) \wedge \mathbf{e}_r} \wedge \mathbf{e}_r - \underbrace{\left(\sum_{i=1}^N m_i \right)}_M (P_o - O) \wedge \mathbf{e}_r = 0.
 \end{aligned}$$

La prima parte del teorema è così dimostrata.

La (7.9) si dimostra con la stessa tecnica, tenendo cioè conto del fatto che

$$\sum_{i=1}^N m_i (P_i - P_o) = 0.$$

Abbiamo infatti

$$\begin{aligned}
 I_{\pi_1 \pi_2}(Q) &= - \sum_{i=1}^N m_i [(P_i - Q) \cdot \mathbf{e}_{\pi_1}] [(P_i - Q) \cdot \mathbf{e}_{\pi_2}] \\
 &= - \sum_{i=1}^N m_i [((P_i - P_o) + (P_o - Q)) \cdot \mathbf{e}_{\pi_1}] [(P_i - P_o) + (P_o - Q)) \cdot \mathbf{e}_{\pi_2}] \\
 &= I_{\pi_1, P_o, \pi_2, P_o}(P_o) - M [(P_o - Q) \cdot \mathbf{e}_{\pi_1}] [(P_o - Q) \cdot \mathbf{e}_{\pi_2}].
 \end{aligned}$$

□

Vediamo adesso alcune proprietà dei momenti d'inerzia.

Proposizione 7.3.1 *Sia \mathfrak{B} un sistema materiale di N punti e sia $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ una terna cartesiana ortogonale centrata in Q . Se $I_i(Q)$, $i = 1, 2, 3$, denota il momento d'inerzia rispetto all'asse \mathbf{e}_i (passante per Q), allora*

$$I_1(Q) + I_2(Q) + I_3(Q) = 2 \sum_{i=1}^N m_i |P_i - Q|^2.$$

Dim. Il generico punto P_i è individuato dal vettore $(P_i - Q)$, che nella terna $\{e_1, e_2, e_3\}$ è rappresentato da

$$P_i - Q = x_i e_1 + y_i e_2 + z_i e_3.$$

Applicando la definizione 7.3.1, abbiamo

$$\underbrace{\sum_{i=1}^N m_i (y_i^2 + z_i^2)}_{I_1(Q)} + \underbrace{\sum_{i=1}^N m_i (x_i^2 + z_i^2)}_{I_2(Q)} + \underbrace{\sum_{i=1}^N m_i (x_i^2 + y_i^2)}_{I_3(Q)} = 2 \underbrace{\sum_{i=1}^N m_i (x_i^2 + y_i^2 + z_i^2)}_{\sum_{i=1}^N m_i |P_i - Q|^2}. \tag{7.10}$$

□

Nota 7.3.1 Osserviamo che, anche considerando un'altra terna $\{e'_1, e'_2, e'_3\}$, sempre centrata in Q ma diversa da $\{e_1, e_2, e_3\}$ (sì che il vettore $(P_i - Q)$ non è più rappresentato da (x_i, y_i, z_i) ma, per esempio, da (x'_i, y'_i, z'_i)), la relazione (7.10) continuerà a valere. Infatti, il secondo membro della (7.10) sarà sempre dato da $\sum_{i=1}^N m_i |P_i - Q|^2$.

Corollario 7.3.1 Se \mathfrak{B} giace su un piano π , e Q sta nel medesimo piano, allora, data una generica terna ortonormale con, e_3 ortogonale a π , si ha

$$I_3(Q) = I_1(Q) + I_2(Q)$$

Dim. Si procede come nel caso precedente, supponendo, senza perdere di generalità, che $z_i = 0$, per ogni $i = 1, 2, \dots, N$.

□

Nota 7.3.2 Vogliamo rimarcare una proprietà importante dei momenti d'inerzia: la proprietà distributiva. Come già visto per il centro di massa, se il sistema \mathfrak{B} è composto da due sottosistemi \mathfrak{B}_1 e \mathfrak{B}_2 , tali che $\mathfrak{B}_1 \cap \mathfrak{B}_2 = \emptyset$, e $\mathfrak{B}_1 \cup \mathfrak{B}_2 = \mathfrak{B}$, allora, definendo con $I_r^{\mathfrak{B}_i}(Q)$, $I_{\pi_1, \pi_2}^{\mathfrak{B}_i}(Q)$, $i = 1, 2$, i momenti d'inerzia dei sottosistemi \mathfrak{B}_i , abbiamo

$$I_r(Q) = I_r^{\mathfrak{B}_1}(Q) + I_r^{\mathfrak{B}_2}(Q), \tag{7.11}$$

$$I_{\pi_1, \pi_2}(Q) = I_{\pi_1, \pi_2}^{\mathfrak{B}_1}(Q) + I_{\pi_1, \pi_2}^{\mathfrak{B}_2}(Q). \tag{7.12}$$

Evidentemente la (7.11) e la (7.12) si generalizzano banalmente al caso in cui \mathfrak{B} sia composto non solo da 2 ma $n > 2$ sottosistemi disgiunti e tali che la loro unione dia \mathfrak{B} .

7.3.2 Omografia d'inerzia, matrice d'inerzia e terna principale d'inerzia

In questa sezione vogliamo analizzare come varia il momento d'inerzia al variare del versore e_r di una retta uscente dal punto Q . Mostreremo che esiste un'applicazione lineare che consente di calcolare $I_r(Q)$ come forma bilineare.

Definizione 7.3.3 Sia dato un sistema di N punti materiali e sia dato un punto Q . Sia poi V lo spazio vettoriale definito sullo spazio affine. Su V si definisce il seguente operatore lineare, detto **omografia d'inerzia**,

$$\sigma(Q) : V \rightarrow V$$

che opera così,

$$\mathbf{v} \in V \xrightarrow{\sigma(Q)} \sigma(Q) \mathbf{v} = \sum_{i=1}^N m_i (P_i - Q) \wedge [\mathbf{v} \wedge (P_i - Q)]. \quad (7.13)$$

Si osserva che in base alla nota formula del calcolo vettoriale

$$\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}),$$

$\sigma(Q)$ può anche essere così definito

$$\sigma(Q) \mathbf{v} = \sum_{i=1}^N m_i \left[|P_i - Q|^2 \mathbf{v} - (\mathbf{v} \cdot (P_i - Q)) (P_i - Q) \right]. \quad (7.14)$$

Proposizione 7.3.2 L'operatore $\sigma(Q)$ gode delle seguenti proprietà:

1. $\mathbf{v} \cdot \sigma(Q) \mathbf{v} = I_{\mathbf{v}}(Q) |\mathbf{v}|^2$, dove $I_{\mathbf{v}}(Q)$ è il momento d'inerzia rispetto alla retta passante per Q parallela a \mathbf{v} .
2. Se \mathbf{u} e \mathbf{v} sono due vettori non paralleli

$$\mathbf{u} \cdot \sigma(Q) \mathbf{v} = I_{\pi_{\mathbf{v}}\pi_{\mathbf{u}}}(Q) |\mathbf{v}| |\mathbf{u}| + \sum_i m_i |P_i - Q|^2 \mathbf{u} \cdot \mathbf{v},$$

dove $I_{\pi_{\mathbf{v}}\pi_{\mathbf{u}}}(Q)$ è il momento d'inerzia rispetto ai due piani $\pi_{\mathbf{v}}$, e $\pi_{\mathbf{u}}$, rispettivamente ortogonali a \mathbf{u} e \mathbf{v} , la cui retta d'intersezione passa per Q .

Dim. La dimostrazione dei punti 1 e 2 si effettua facilmente sfruttando la formula (7.14). Infatti, per quanto riguarda il punto 1, ricordando che $\mathbf{v} \cdot (P_i - Q) = |\mathbf{v}| |P_i - Q| \cos \theta_i$, essendo θ_i l'angolo fra $(P_i - Q)$ e \mathbf{v} , si ha

$$\begin{aligned} \mathbf{v} \cdot \sigma(Q) \mathbf{v} &= \sum_{i=1}^N m_i |P_i - Q|^2 |\mathbf{v}|^2 \underbrace{(1 - \cos^2 \theta_i)}_{\sin^2 \theta_i} \\ &= \underbrace{\left(\sum_{i=1}^N m_i [(P_i - Q) \wedge \mathbf{e}_{\mathbf{v}}]^2 \right)}_{I_{\mathbf{e}_{\mathbf{v}}}(Q)} |\mathbf{v}|^2, \end{aligned}$$

dove $\mathbf{e}_{\mathbf{v}}$ vettore unitario parallelo a \mathbf{v} .

Analogamente, partendo dalla (7.14) e ricordando la definizione 2, il punto 2 è facilmente dimostrabile.

□

La conseguenza della proposizione 7.3.2 è la seguente: $\sigma(Q)$ definisce una forma bilineare simmetrica definita positiva. La simmetria è banale da dimostrare come il fatto che $\mathbf{v} \cdot \sigma(Q) \mathbf{v} > 0$, se $\mathbf{v} \neq 0$, e che $\mathbf{v} \cdot \sigma(Q) \mathbf{v} = 0$, solo se $\mathbf{v} = 0$ (si esclude qui il particolarissimo caso in cui i punti di \mathfrak{B} giacciono sulla retta passante per Q e parallela al vettore \mathbf{v}). Inoltre, il significato della forma è chiaro: se \mathbf{e}_r è il vettore della retta passante per Q , la quantità $\mathbf{e}_r \cdot \sigma(Q) \mathbf{e}_r$ è il momento d'inerzia del sistema \mathfrak{B} rispetto a tale retta.

Data adesso una terna ortonormale $S : \{e_1, e_3, e_3\}$, centrata in Q , possiamo rappresentare $\sigma(Q)$ tramite la matrice simmetrica \mathbb{I}_Q

$$\mathbb{I}_Q = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{33} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix}$$

i cui elementi si determinano così

$$I_{jk} = \mathbf{e}_j \cdot \sigma(Q) \mathbf{e}_k, \quad j, k = 1, 2, 3.$$

Evidentemente, gli elementi I_{ij} della matrice, detta **matrice d'inerzia**, o **tensore d'inerzia**, dipendono dalla particolare terna scelta. Pertanto, se il generico vettore $(P_i - Q)$ si rappresenta nel SdR S come

$$P_i - Q = x_i e_1 + y_i e_2 + z_i e_3, \tag{7.15}$$

la matrice \mathbb{I}_Q assumerà (rispetto ad S) questa forma

$$\mathbb{I}_Q = \begin{pmatrix} \sum_i m_i (y_i^2 + z_i^2) & -\sum_i m_i x_i y_i & -\sum_i m_i x_i z_i \\ -\sum_i m_i x_i y_i & \sum_i m_i (x_i^2 + z_i^2) & -\sum_i m_i z_i y_i \\ -\sum_i m_i x_i z_i & -\sum_i m_i z_i y_i & \sum_i m_i (x_i^2 + y_i^2) \end{pmatrix}. \tag{7.16}$$

Gli elementi diagonali I_{ii} sono i momenti di inerzia rispetto agli assi x, y e z rispettivamente, e gli elementi extra-diagonali sono detti **momenti deviatori** rispetto alle coppie di piani $\{x = 0, y = 0\}$, $\{x = 0, z = 0\}$ e $\{y = 0, z = 0\}$.

Notiamo che se invece di S avessimo considerato un'altra terna ortonormale $S' : \{e'_1, e'_3, e'_3\}$, sempre centrata in Q , avremmo ottenuto un'altra matrice \mathbb{I}'_Q , che rappresenta l'operatore $\sigma(Q)$ rispetto S' . In particolare, le matrici \mathbb{I}'_Q e \mathbb{I}_Q sono legate da

$$\mathbb{I}_Q = \mathbb{B} \mathbb{I}'_Q \mathbb{B}^T, \tag{7.17}$$

dove \mathbb{B} è la matrice ortogonale (cioè $\mathbb{B}^T = \mathbb{B}^{-1}$) del cambio di base da S' in S , definita nella sezione 1.3. Infatti, ricordando la (1.10) abbiamo

$$\mathbf{e}_i = \sum_{j=1}^3 b_{ij} \mathbf{e}'_j, \quad i = 1, 2, 3,$$

dove² b_{ij} , è l'elemento i, j della matrice \mathbb{B} la cui forma generale è data dalla (1.12). Otteniamo dunque

$$\begin{aligned} \underbrace{e_i \cdot \sigma(Q) e_j}_{(\mathbb{I}_Q)_{ij}} &= \left[\sum_{k=1}^3 b_{ik} e'_k \right] \cdot \sigma(Q) \left[\sum_{s=1}^3 b_{js} e'_s \right] \\ &= \sum_{k=1}^3 \sum_{s=1}^3 b_{ik} \underbrace{(e'_k \cdot \sigma(Q) e'_s)}_{(\mathbb{I}'_Q)_{ks}} b_{js} \\ &= \sum_{k=1}^3 \sum_{s=1}^3 b_{ik} (\mathbb{I}'_Q)_{ks} b_{js}^T, \end{aligned}$$

ovvero la (7.17).

Data la proprietà di simmetria di \mathbb{I}_Q (che ovviamente discende dalla simmetria di $\sigma(Q)$), è possibile determinare una particolare terna ortonormale, detta **terna principale d'inerzia**, rispetto alla quale la matrice d'inerzia è diagonale, ovvero

$$\mathbb{I}_Q = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}.$$

I versori che costituiscono la terna principale d'inerzia (centrata in Q) sono detti **assi principali d'inerzia**, ed i momenti I_1, I_2, I_3 , **momenti principali d'inerzia**.

I vettori unitari $e_i, i = 1, 2, 3$, della terna principale d'inerzia, sono gli autovettori dell'operatore $\sigma(Q)$. Infatti, se un generico vettore v è parallelo ad un autovettore di σ , allora viene "trasportato" da $\sigma(Q)$ in un altro vettore, $\sigma(Q)v$, parallelo a v stesso, ovvero

$$\sigma(Q)v = I_i v,$$

con l'evidente significato del momento I_i : l'autovalore corrispondente all'autovettore e_i .

La determinazione della terna principale d'inerzia si fa diagonalizzando \mathbb{I}_Q , ovvero ricercando gli autovettori di \mathbb{I}_Q stessa. Tale procedimento può essere talvolta lungo. Si preferisce allora seguire (quando possibile) alcune "scorciatoie" che permettono di individuare subito una terna principale d'inerzia a partire dalle proprietà geometrico-materiali del sistema di punti materiali \mathfrak{B} .

Nota 7.3.3 *È interessante osservare che si può giungere alla matrice (7.16) partendo direttamente dalla definizione (7.13). Selezioniamo infatti un SdR dove il vettore $(P_i - Q)$ è dato dalla (7.15) e dove $v = v_1 e_1 + v_2 e_2 + v_3 e_3$. Utilizziamo poi l'espressione matriciale del prodotto vettoriale vista nella sezione 6.3, cioè*

$$v \wedge (P_i - Q) = - (P_i - Q) \wedge v = - \begin{pmatrix} 0 & z_i & -y_i \\ -z_i & 0 & x_i \\ y_i & -x_i & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

²Dalla (1.11) si ha $b_{ij} = e_i \cdot e'_j, i, j = 1, 2, 3$.

Scrivendo quindi esplicitamente la (7.16) si ottiene

$$\begin{aligned} \sigma(Q) \mathbf{v} &= - \sum_{i=1}^N m_i \begin{pmatrix} 0 & z_i & -y_i \\ -z_i & 0 & x_i \\ y_i & -x_i & 0 \end{pmatrix} \begin{pmatrix} 0 & z_i & -y_i \\ -z_i & 0 & x_i \\ y_i & -x_i & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \\ &= \underbrace{\sum_{i=1}^N m_i \begin{pmatrix} y_i^2 + z_i^2 & -x_i y_i & -x_i z_i \\ -x_i y_i & x_i^2 + z_i^2 & -y_i z_i \\ -x_i z_i & -y_i z_i & x_i^2 + y_i^2 \end{pmatrix}}_{\mathbb{I}_Q} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}. \end{aligned}$$

Nota 7.3.4 Sia data la terna ortonormale $S : \{ \mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_3 \}$, tale che \mathbf{e}_3 , per esempio, è asse principale d'inerzia, allora $I_{13} = I_{23} = 0$. Viceversa, se \mathbb{I}_Q è la matrice d'inerzia rispetto a S , ed è tale che $I_{13} = I_{23} = 0$, allora l'asse \mathbf{e}_3 è asse principale d'inerzia. La dimostrazione di tali affermazioni è molto semplice. Infatti se \mathbf{e}_3 è asse principale d'inerzia, allora $\mathbb{I}_Q \mathbf{e}_3 = I_3 \mathbf{e}_3$ ovvero

$$\mathbb{I}_Q \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ I_3 \end{pmatrix}.$$

Del resto, in generale, $\mathbb{I}_Q = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{33} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix}$. Avremo quindi

$$\begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{33} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} I_{13} \\ I_{23} \\ I_{33} \end{pmatrix},$$

da cui, per banale confronto, discende $I_{13} = I_{23} = 0$. Se invece \mathbb{I}_Q ha questa forma

$$\mathbb{I}_Q = \begin{pmatrix} I_{11} & I_{12} & 0 \\ I_{12} & I_{33} & 0 \\ 0 & 0 & I_{33} \end{pmatrix},$$

applicando \mathbb{I}_Q ad ogni vettore del tipo $\beta \mathbf{e}_3$ (parallelo all'asse \mathbf{e}_3)

$$\begin{pmatrix} I_{11} & I_{12} & 0 \\ I_{12} & I_{33} & 0 \\ 0 & 0 & I_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ I_{33} \beta \end{pmatrix} = I_{33} \beta \mathbf{e}_3,$$

si ottiene un vettore ancora parallelo ad \mathbf{e}_3 . L'asse \mathbf{e}_3 è dunque asse principale d'inerzia. Evidentemente, tale ragionamento resta valido anche se, invece dell'asse \mathbf{e}_3 , avessimo scelto l'asse \mathbf{e}_1 , oppure l'asse \mathbf{e}_2 .

Definizione 7.3.4 Dato un sistema \mathfrak{B} di N punti materiali, un piano π si dice piano di simmetria materiale se per ogni punto materiale P_i , ne esiste un altro, avente la stessa massa, disposto in posizione simmetrica rispetto a π .

Proposizione 7.3.3 *Sia \mathfrak{B} un sistema di punti materiali e sia π un piano di simmetria materiale. Sia inoltre Q un qualsiasi punto di π . Allora una retta passante per Q e perpendicolare a π , è asse principale d'inerzia*

Dim. Consideriamo una generica terna ortonormale $S : \{e_1, e_2, e_3\}$, centrata in Q , con e_3 , perpendicolare al piano che, senza perdere in generalità, sarà dato da $z = 0$. Rifacendosi alla nota 7.3.4, per provare che e_3 è asse principale d'inerzia, è sufficiente mostrare che $I_{23} = I_{13} = 0$. Se $P_i - Q = x_i e_1 + y_i e_2 + z_i e_3$, possiamo raggruppare i primi $N/2$ punti come quelli caratterizzati da $z_i \leq 0$, e i rimanenti $N/2$ come i loro simmetrici (la cui coordinata rispetto ad e_3 è $-z_i$). Quindi, ricordando la (7.16), abbiamo

$$I_{13} = \sum_{i=1}^{N/2} m_i x_i (z_i - z_i) = 0, \quad \text{e} \quad I_{23} = \sum_{i=1}^{N/2} m_i y_i (z_i - z_i) = 0.$$

□

Corollario 7.3.2 *Se \mathfrak{B} è un sistema piano (tutti i punti giacciono su un piano) allora una retta perpendicolare al piano è asse principale d'inerzia.*

Dim. Possiamo considerare, senza perdere di generalità, che $z_i = 0, \forall i = 1, 2, \dots, N$ (il sistema giace sul piano $z = 0$). Procedendo come nella dimostrazione della proposizione 7.3.3, si ottiene $I_{13} = I_{23} = 0$. La retta e_3 (retta perpendicolare al piano su cui giace il sistema) è asse principale d'inerzia.

□

Proposizione 7.3.4 *Sia $S_{P_o} : \{e_1, e_2, e_3\}$ una terna principale d'inerzia centrata in P_o , centro di massa di \mathfrak{B} . Dato un qualsiasi punto A appartenente ad uno degli assi di S (ovvero $(A - P_o)$ parallelo a e_1 , oppure ad e_2 , oppure ad e_3), allora la terna $\{e_1, e_2, e_3\}$ centrata in A , che si denota con S_A , è terna principale d'inerzia.*

Dim. Incominciamo col considerare un punto A qualsiasi

$$(A - P_o) = x_A e_1 + y_A e_2 + z_A e_3.$$

Il sistema S_{P_o} è terna principale d'inerzia, per cui

$$\mathbb{I}_{P_o} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}.$$

Consideriamo adesso la terna S_A centrata in A , e applichiamo il teorema di Huygens

$$\mathbb{I}_A = \begin{pmatrix} I_1 + M(y_A^2 + z_A^2) & -Mx_A y_A & -Mx_A z_A \\ -Mx_A y_A & I_2 + M(x_A^2 + z_A^2) & -My_A z_A \\ -Mx_A z_A & -My_A z_A & I_3 + M(x_A^2 + y_A^2) \end{pmatrix} \quad (7.18)$$

Quindi, se, per esempio, A giace sull'asse e_1 (sì che $y_A = z_A = 0$), avremo \mathbb{I}_A diagonale, come del resto se A giace sull'asse e_2 o sull'asse e_3 .

□

7.3.3 Ellissoide d'inerzia

Sia dato il sistema di punti materiali \mathfrak{B} . Fissata una generica direzione individuata dal versore \mathbf{w} , $|\mathbf{w}| = 1$, il momento d'inerzia di S rispetto alla retta avente direzione \mathbf{w} , passante per Q , è dato da (proposizione 7.3.2, punto 1)

$$I_{\mathbf{w}}(Q) = \mathbf{w} \cdot \boldsymbol{\sigma} \mathbf{w}. \tag{7.19}$$

Consideriamo adesso il vettore \mathbf{W} parallelo a \mathbf{w} , dato da $\mathbf{W} = \frac{1}{\sqrt{I_{\mathbf{w}}(Q)}} \mathbf{w}$. Evidentemente $|\mathbf{W}| = \frac{1}{\sqrt{I_{\mathbf{w}}(Q)}}$. Dalla (7.19) abbiamo

$$\mathbf{W} \cdot \boldsymbol{\sigma} \mathbf{W} = 1.$$

Fissiamo una generica terna ortonormale $S_Q : \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \}$, centrata in Q . Rispetto a S_Q , il vettore \mathbf{W} avrà componenti

$$\mathbf{W} = \xi \mathbf{e}_1 + \eta \mathbf{e}_2 + \zeta \mathbf{e}_3, \tag{7.20}$$

tali che $\sqrt{\xi^2 + \eta^2 + \zeta^2} = \frac{1}{\sqrt{I_{\mathbf{w}}(Q)}}$. Inoltre, sempre rispetto a S_Q , l'omografia d'inerzia $\boldsymbol{\sigma}$, sarà rappresentata dalla matrice

$$\mathbb{I}_Q = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix}.$$

Di conseguenza

$$1 = \mathbf{W} \cdot \boldsymbol{\sigma} \mathbf{W} = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \cdot \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix},$$

ovvero

$$I_{11}\xi^2 + I_{22}\eta^2 + I_{33}\zeta^2 + 2I_{12}\xi\eta + 2I_{13}\xi\zeta + 2I_{23}\eta\zeta = 1. \tag{7.21}$$

L'equazione (7.21) è l'equazione di una superficie (ellissoide), detto **ellissoide d'inerzia**, centrata in Q . Quindi, data comunque una direzione \mathbf{w} , il vettore $\mathbf{W} = \frac{\mathbf{w}}{\sqrt{I_{\mathbf{w}}(Q)}}$, le cui generiche componenti sono specificate dalla (7.20), individua un ben preciso punto geometrico che giace sulla superficie (7.21). E' ovviamente vero anche il viceversa. Ad ogni punto A che giace sulla superficie (7.21), e che è individuato da una generica terna (ξ, η, ζ) , possiamo associare il vettore $\mathbf{W} = (A - Q) = \xi \mathbf{e}_1 + \eta \mathbf{e}_2 + \zeta \mathbf{e}_3$. La norma di tale vettore ha un ben preciso significato fisico

$$|A - Q| = |\mathbf{W}| = \sqrt{\xi^2 + \eta^2 + \zeta^2} = \frac{1}{\sqrt{I_{\mathbf{w}}(Q)}}. \tag{7.22}$$

Ora, i punti che stanno la superficie (7.21), si trovano a varie distanze³ dal centro Q . In particolare, i due punti che si trovano in corrispondenza dell'asse maggiore

³I punti si troverebbero tutti alla stessa distanza da Q nel caso in cui l'ellissoide fosse una sfera.

sono quelli che si trovano a maggior distanza da Q , ossia quelli per cui $|A - Q| = \sqrt{\xi^2 + \eta^2 + \zeta^2}$ è la più grande possibile. Allo stesso modo i due punti che stanno sull'asse minore sono quelli che si trovano a minor distanza da Q , ossia quelli per cui $|A - Q| = \sqrt{\xi^2 + \eta^2 + \zeta^2}$ è la più piccola possibile. Detti dunque A_{\max} e A_{\min} i punti dell'ellissoide che si trovano in corrispondenza dell'asse maggiore e dell'asse minore, per ogni punto A che giace sull'ellissoide, varrà

$$|A_{\min} - Q| \leq |A - Q| \leq |A_{\max} - Q|.$$

Ma allora, dalla (7.22), data una generica direzione individuata da \mathbf{w} , abbiamo che

$$|A_{\min} - Q| \leq \frac{1}{\sqrt{I_{\mathbf{w}}(Q)}} \leq |A_{\max} - Q|,$$

ossia

$$\frac{1}{|A_{\max} - Q|} \leq \sqrt{I_{\mathbf{w}}(Q)} \leq \frac{1}{|A_{\min} - Q|}.$$

Questo vuol dire che il momento d'inerzia di un sistema \mathfrak{B} rispetto alla retta per Q avente direzione \mathbf{w} , è limitato dall'alto e dal basso. Non solo, ma, fissato Q , al variare della direzione \mathbf{w} , il momento d'inerzia ha un massimo ed un minimo. Il massimo è ottenuto quando la direzione \mathbf{w} corrisponde alla direzione dell'asse minore dell'ellissoide mentre il minimo quando la direzione \mathbf{w} è parallela all'asse maggiore dell'ellissoide.

7.3.4 Determinazione della terna principale d'inerzia nel caso di sistemi piani

Supponiamo che tutti i punti materiali di \mathfrak{B} giacciono su un piano che, senza perdere in generalità, può essere il piano $z = 0$. Fissiamo poi un punto Q (giacente su $z = 0$) ed una terna $S_Q : \{e_1, e_2, e_3\}$, tale che il versore e_3 sia perpendicolare al piano $z = 0$ (e pertanto, corollario 7.3.2, è asse principale d'inerzia). I punti P_i di \mathfrak{B} avranno coordinate $(P_i - Q) = x_i e_1 + y_i e_2$, $i = 1, 2, \dots, N$. Riferendoci sempre a S_Q e limitandoci solo alle due componenti "piane", la matrice d'inerzia sarà

$$\mathbb{I}_Q = \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix}.$$

Il problema che ci poniamo è il seguente: determinare un nuovo sistema $S'_Q : \{e'_1, e'_2, e_3\}$, con e'_1 e e'_2 ruotati di un angolo α rispetto a e_1 ed e_2 , in modo che S'_Q sia terna principale d'inerzia. Evidentemente, $-\frac{\pi}{2} \leq \alpha \leq \frac{\pi}{2}$. Nel seguito presentiamo alcuni metodi per individuare l'angolo α , in modo tale che S'_Q risulti terna principale d'inerzia.

METODO DEL CERCHIO DI MOHR

Riferendoci alla figura 7.3, S_Q indica il sistema $\{x, y\}$, mentre S'_Q il sistema $\{X, Y\}$. Nel sistema S'_Q la matrice d'inerzia è data da

$$\mathbb{J}_Q = \begin{pmatrix} J_{11} & J_{12} \\ J_{12} & J_{22} \end{pmatrix}.$$

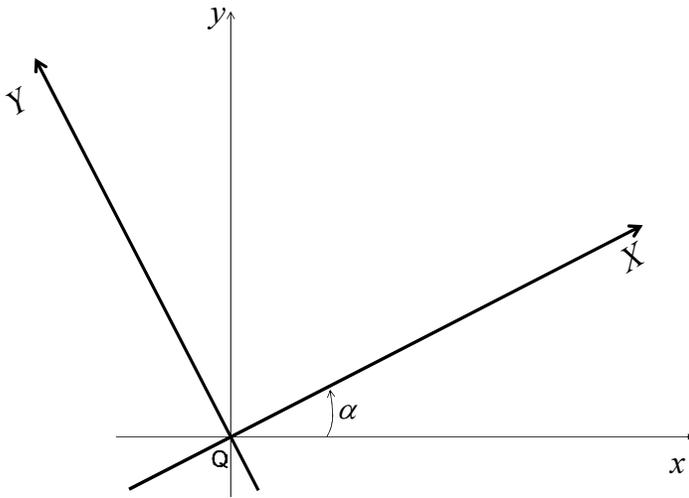


Figura 7.3: Sistema \$(X; Y)\$ ruotato di \$\alpha\$ rispetto a \$(x, y)\$.

Ora, è facile vedere che

$$\begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}}_{\mathbb{B}} \begin{pmatrix} X \\ Y \end{pmatrix},$$

per cui, ricordando la (7.17), la relazione che lega \$\mathbb{I}_Q\$ a \$\mathbb{J}_Q\$ è la seguente

$$\mathbb{I}_Q = \mathbb{B} \mathbb{J}_Q \mathbb{B}^T, \quad \Leftrightarrow \quad \mathbb{J}_Q = \mathbb{B}^T \mathbb{I}_Q \mathbb{B}. \tag{7.23}$$

Se supponiamo nota \$\mathbb{I}_Q\$, gli elementi della matrice \$\mathbb{J}_Q\$ si ottengono dalla seconda della (7.23), cioè

$$\left\{ \begin{array}{l} J_{11} = \frac{I_{11} + I_{22}}{2} - \frac{I_{22} - I_{11}}{2} \cos 2\alpha + I_{12} \sin 2\alpha, \\ J_{22} = \frac{I_{11} + I_{22}}{2} + \frac{I_{22} - I_{11}}{2} \cos 2\alpha - I_{12} \sin 2\alpha, \\ J_{12} = \frac{I_{22} - I_{11}}{2} \sin 2\alpha + I_{12} \cos 2\alpha, \end{array} \right. \tag{7.24}$$

dove abbiamo sfruttato le formule di duplicazione: \$\sin^2 \alpha = \frac{1 - \cos 2\alpha}{2}\$, \$\cos^2 \alpha = \frac{1 + \cos 2\alpha}{2}\$.

Se selezioniamo la (7.24)₁ e la (7.24)₃, abbiamo, nel piano (J_{11}, J_{12}) , una curva definita tramite il parametro 2α , $-\pi \leq 2\alpha \leq \pi$,

$$\begin{cases} J_{11} = \frac{I_{11} + I_{22}}{2} - \frac{I_{22} - I_{11}}{2} \cos 2\alpha + I_{12} \sin \alpha, \\ J_{12} = \frac{I_{22} - I_{11}}{2} \sin 2\alpha + I_{12} \cos \alpha. \end{cases}$$

Tale curva rappresenta una circonferenza, detta **cerchio di Mohr** (v. figura 7.4), di raggio

$$R = \sqrt{\left(\frac{I_{22} - I_{11}}{2}\right)^2 + I_{12}^2},$$

ed il cui centro nel piano (J_{11}, J_{12}) ha coordinate $\left(\frac{I_{22} + I_{11}}{2}, 0\right)$.

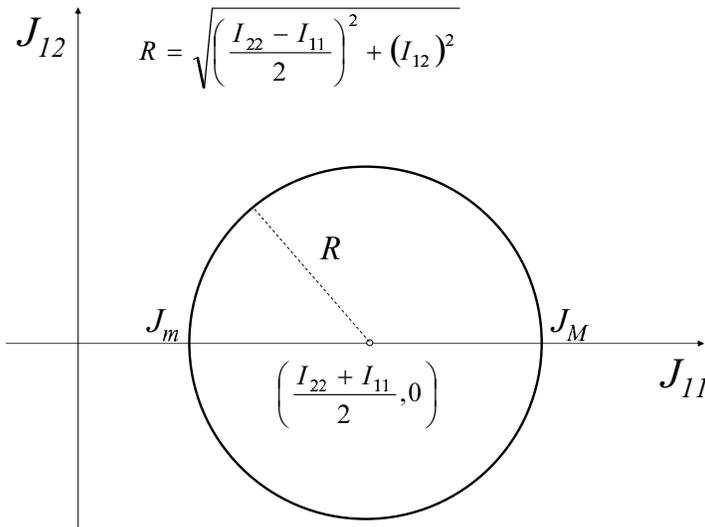


Figura 7.4: Cerchio di Mohr.

Quindi, l'angolo α per cui S'_Q è principale d'inerzia, è quello per cui $J_{12} = 0$, ovvero

$$\frac{I_{22} - I_{11}}{2} \sin 2\alpha + I_{12} \cos 2\alpha = 0, \quad \Leftrightarrow \quad \tan 2\alpha = -\frac{2I_{12}}{I_{22} - I_{11}}. \quad (7.25)$$

In corrispondenza di tali angoli⁴ J_{11} assume valore massimo

$$J_M = \frac{I_{11} + I_{22}}{2} + \sqrt{\left(\frac{I_{22} - I_{11}}{2}\right)^2 + I_{12}^2}, \quad (7.26)$$

⁴Gli angoli α che soddisfano la (7.25) sono due e sono separati da $\pi/2$.

o minimo

$$J_m = \frac{I_{11} + I_{22}}{2} - \sqrt{\left(\frac{I_{22} - I_{11}}{2}\right)^2 + I_{12}^2}. \quad (7.27)$$

Notiamo che gli autovalori di \mathbb{I}_Q sono ottenuti dall'equazione

$$\det(\mathbb{I}_Q - \lambda \mathbb{I}) = 0, \quad \Rightarrow \quad (I_{11} - \lambda)(I_{22} - \lambda) - I_{12}^2 = 0,$$

le cui soluzioni sono proprio le (7.26), (7.27).

METODO DEL PRODOTTO VETTORIALE

Se il versore \mathbf{u} è parallelo ad un autovettore di $\sigma(Q)$, allora $\sigma(Q)\mathbf{u}$ è parallelo ad \mathbf{u} stesso, per cui

$$\mathbf{u} \wedge \sigma(Q)\mathbf{u} = 0. \quad (7.28)$$

Quindi, fissata una generica terna S_Q , rispetto alla quale \mathbf{u} ha la seguente espressione

$$\mathbf{u} = \cos \alpha \mathbf{e}_1 + \sin \alpha \mathbf{e}_2, \quad (7.29)$$

avremo

$$\begin{aligned} \sigma(Q)\mathbf{u} &= \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix} \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \\ &= (I_{11} \cos \alpha + I_{12} \sin \alpha) \mathbf{e}_1 + (I_{12} \cos \alpha + I_{22} \sin \alpha) \mathbf{e}_2. \end{aligned}$$

Pertanto

$$\mathbf{u} \wedge \sigma(Q)\mathbf{u} = \left(\frac{I_{22} - I_{11}}{2} \sin 2\alpha + I_{12} \cos 2\alpha \right) \mathbf{e}_3,$$

da cui, imponendo la (7.28), si ottiene nuovamente la (7.25).

METODO DELLA DERIVAZIONE

Il metodo consiste nello sfruttare il fatto (messo in luce nella sezione 7.3.3) che, fissato Q , una direzione \mathbf{u} è asse principale d'inerzia se il momento d'inerzia rispetto ad essa è massimo o minimo, o comunque stazionario. Quindi, se \mathbf{u} è una generica direzione data da (7.29), avremo

$$\begin{aligned} I_{\mathbf{u}}(Q) &= \mathbf{u} \cdot \sigma(Q)\mathbf{u} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \cdot \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix} \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \\ &= \frac{I_{11} + I_{22}}{2} - \frac{(I_{22} - I_{11})}{2} \cos 2\alpha + I_{12} \sin 2\alpha. \end{aligned}$$

Quindi, $I_{\mathbf{u}}(Q)$ risulta una essere una funzione di α . Ricercando gli angoli α che rendono stazionaria tale funzione, imponiamo $\frac{dI_{\mathbf{u}}(Q)}{d\alpha} = 0$, da cui si ottiene, ancora una volta, la (7.25).

7.3.5 Esempi e complementi

In questa sezione daremo alcuni esempi sul calcolo dei momenti d'inerzia e illustreremo alcune proprietà della matrice d'inerzia dovute alla simmetria della stessa (proprietà note dal corso di algebra lineare).

Proposizione 7.3.5 *Gli autovalori della matrice d'inerzia sono tutti reali.*

Dim. La dimostrazione di questa proposizione è un caso particolare della dimostrazione del lemma 5.10.1 al quale rimandiamo.

□

Mostriamo adesso che gli autovalori sono invarianti, cioè non dipendono dal particolare SdR ortonormale scelto.

Proposizione 7.3.6 *Gli autovalori della matrice d'inerzia sono invarianti per cambiamenti di base.*

Dim. Sia $S_Q : \{e_1, e_3, e_3\}$, una base ortonormale centrata in Q . In tale base $\sigma(Q)$ è rappresentata dalla matrice simmetrica \mathbb{I}_Q . Se $S'_Q : \{e'_1, e'_3, e'_3\}$, è un'altra base ortonormale centrata in Q . La matrice che rappresenta l'operatore $\sigma(Q)$ rispetto S'_Q è \mathbb{I}'_Q . Sappiamo però che le matrici \mathbb{I}'_Q e \mathbb{I}_Q sono legate dalla (7.17), con \mathbb{B} matrice ortogonale.

E' facile provare che $\det(\mathbb{I}'_Q - \lambda\mathbb{I}) = \det(\mathbb{I}_Q - \lambda\mathbb{I})$. Infatti

$$\begin{aligned} \det(\mathbb{I}_Q - \lambda\mathbb{I}) &= \det(\mathbb{B}\mathbb{I}'_Q\mathbb{B}^{-1} - \lambda\mathbb{B}\mathbb{I}\mathbb{B}^{-1}) = \det(\mathbb{B}(\mathbb{I}'_Q - \lambda\mathbb{I})\mathbb{B}^{-1}) \\ &= \det\mathbb{B} \det(\mathbb{I}'_Q - \lambda\mathbb{I}) \det\mathbb{B}^{-1} = \det(\mathbb{I}'_Q - \lambda\mathbb{I}). \end{aligned}$$

Vediamo però anche un'altra dimostrazione. Sviluppando $\det(\mathbb{I}'_Q - \lambda\mathbb{I}) = 0$, otteniamo il seguente polinomio in λ (detto polinomio caratteristico)

$$\lambda^3 - \mathcal{I}_1\lambda^2 + \mathcal{I}_2\lambda - \mathcal{I}_3 = 0,$$

con

$$\begin{cases} \mathcal{I}_1 = \text{tr } \mathbb{I}_Q, \\ \mathcal{I}_2 = \frac{1}{2} \left[(\text{tr } \mathbb{I}_Q)^2 - \text{tr}(\mathbb{I}_Q^2) \right], \\ \mathcal{I}_3 = \det \mathbb{I}_Q, \end{cases}$$

dove $\text{tr } \mathbb{I}_Q = \sum_{i=1}^3 I_{ii}$, è la somma degli elementi diagonali. Le tre quantità \mathcal{I}_1 , \mathcal{I}_2 e \mathcal{I}_3 sono detti invarianti della matrice \mathbb{I}_Q , proprio perché non dipendono dalla particolare base. Infatti, ricordando che $\text{tr}\mathbb{A}\mathbb{B} = \text{tr}\mathbb{B}\mathbb{A}$, si ha

$$\begin{aligned} \text{tr } \mathbb{I}_Q &= \text{tr}(\mathbb{B}\mathbb{I}'_Q\mathbb{B}^{-1}) = \text{tr}(\mathbb{B}^{-1}\mathbb{B}\mathbb{I}'_Q) = \text{tr}(\mathbb{I}'_Q), \\ \text{tr}(\mathbb{I}_Q^2) &= \text{tr}(\underbrace{\mathbb{B}\mathbb{I}'_Q\mathbb{B}^{-1}}_{\mathbb{I}_Q} \underbrace{\mathbb{B}\mathbb{I}'_Q\mathbb{B}^{-1}}_{\mathbb{I}_Q}) = \text{tr}(\mathbb{B}^{-1}\mathbb{B}\mathbb{I}'_Q{}^2) = \text{tr}(\mathbb{I}'_Q{}^2). \end{aligned}$$

Quindi, $\det(\mathbb{I}_Q - \lambda \mathbb{I}) = 0$ darà luogo sempre allo stesso polinomio indipendentemente dalla base su cui è stata rappresentata $\sigma(Q)$. Di conseguenza gli autovalori di $\sigma(Q)$ non dipendono dalla particolare base.

□

Come ultima proprietà degli assi principali d'inerzia (che, come sappiamo, rappresentano gli autovettori della matrice d'inerzia) dimostriamo la seguente

Proposizione 7.3.7 *Esiste una base ortonormale di \mathbb{R}^3 formata da autovettori di $\sigma(Q)$.*

Dim. Siano \mathbf{u}_1 , e \mathbf{u}_2 due autovettori relativi agli autovalori distinti I_1 e I_2 , ovvero

$$\sigma(Q) \mathbf{u}_i = I_i \mathbf{u}_i, \quad i = 1, 2.$$

Siccome $\sigma(Q)$ è simmetrica avremo $\mathbf{u}_1 \cdot \sigma(Q) \mathbf{u}_2 = \sigma(Q) \mathbf{u}_1 \cdot \mathbf{u}_2$, e quindi

$$0 = \mathbf{u}_1 \cdot \sigma(Q) \mathbf{u}_2 - \sigma(Q) \mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_1 \cdot I_2 \mathbf{u}_2 - I_1 \mathbf{u}_1 \cdot \mathbf{u}_2 = (I_2 - I_1) \mathbf{u}_1 \cdot \mathbf{u}_2,$$

che comporta $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, dal momento che $I_1 \neq I_2$.

Evidentemente tale procedura non si applica nel caso in cui $I_1 = I_2$, ovvero se l'autovalore I_2 ha molteplicità algebrica doppia. Consideriamo quindi il caso in cui i tre autovalori di $\sigma(Q)$ sono i seguenti $I_1 = I_2 \neq I_3$. Consideriamo gli autovettori relativi a I_3 ed a I_2 , ovvero

$$\sigma(Q) \mathbf{u}_i = I_i \mathbf{u}_i, \quad i = 2, 3.$$

Sia poi \mathbf{v} un vettore ortogonale sia a \mathbf{u}_2 che a \mathbf{u}_3 . Mostriamo che \mathbf{v} è autovettore di $\sigma(Q)$. Infatti le componenti di $\sigma(Q) \mathbf{v}$ lungo \mathbf{u}_2 e \mathbf{u}_3 sono nulle

$$\begin{aligned} \mathbf{u}_2 \cdot \sigma(Q) \mathbf{v} &= \sigma(Q) \mathbf{u}_2 \cdot \mathbf{v} = I_2 \mathbf{u}_2 \cdot \mathbf{v} = 0, \\ \mathbf{u}_3 \cdot \sigma(Q) \mathbf{v} &= \sigma(Q) \mathbf{u}_3 \cdot \mathbf{v} = I_3 \mathbf{u}_3 \cdot \mathbf{v} = 0. \end{aligned}$$

Dunque $\sigma(Q) \mathbf{v}$ è parallelo a \mathbf{v} , ovvero $\sigma(Q) \mathbf{v} = \lambda \mathbf{v}$. Se adesso supponiamo che \mathbf{v} , \mathbf{u}_2 e \mathbf{u}_3 , siano vettori unitari, si deduce che essi costituiscono una base ortonormale. Calcolando la traccia di $\sigma(Q)$ su tale base abbiamo $\text{tr} \sigma(Q) = \lambda + I_2 + I_3$. Ma la traccia è un invariante, per cui abbiamo anche $\text{tr} \sigma(Q) = I_2 + I_2 + I_3$. Di conseguenza $\lambda = I_2$. Si conclude che, anche nel caso in cui un autovalore abbia molteplicità doppia, si trovano comunque tre autovettori ortogonali. Non solo, ma si dimostra che un qualsiasi vettore ortogonale a \mathbf{u}_3 è autovettore relativo all'autovalore doppio I_2 . Infatti, consideriamo un generico vettore $\mathbf{w} = \alpha \mathbf{v} + \beta \mathbf{u}_2$, ortogonale a \mathbf{u}_3 . Abbiamo

$$\sigma(Q) \mathbf{w} = \sigma(Q) (\alpha \mathbf{v} + \beta \mathbf{u}_2) = I_2 \alpha \mathbf{v} + I_2 \beta \mathbf{u}_2 = I_2 \mathbf{w}.$$

In termini di assi principali d'inerzia, questo fatto vuol dire che ogni retta passante per Q , e che giace sul piano ortogonale a \mathbf{u}_3 , è asse principale d'inerzia.

Vediamo infine l'ultimo caso: i tre autovalori sono coincidenti $I_1 = I_2 = I_3 = I$. Sia \mathbf{u} un autovettore, $\sigma(Q) \mathbf{u} = I \mathbf{u}$, che supporremo unitario. Consideriamo adesso altri due vettori unitari \mathbf{v} e \mathbf{w} , tra loro ortogonali e ortogonali rispetto a \mathbf{u} . Il sistema

\mathbf{u} , \mathbf{v} e \mathbf{w} , costituisce una base ortonormale, su cui possiamo rappresentare $\sigma(Q)$. In particolare, siccome

$$\mathbf{u} \cdot \sigma(Q) \mathbf{v} = \sigma(Q) \mathbf{u} \cdot \mathbf{v} = I \mathbf{u} \cdot \mathbf{v} = 0,$$

e analogamente $\mathbf{u} \cdot \sigma(Q) \mathbf{w} = 0$, si deduce che $\sigma(Q) \mathbf{v}$ e $\sigma(Q) \mathbf{w}$ giacciono nel piano π , piano generato da \mathbf{v} e \mathbf{w} (e che è ortogonale a \mathbf{u}). Rappresentando dunque $\sigma(Q)$ nella base $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ otterremo la seguente matrice

$$\mathbb{I}_Q = \begin{pmatrix} \sigma(Q) \mathbf{v} \cdot \mathbf{v} & \sigma(Q) \mathbf{v} \cdot \mathbf{w} & \sigma(Q) \mathbf{v} \cdot \mathbf{u} \\ \sigma(Q) \mathbf{w} \cdot \mathbf{v} & \sigma(Q) \mathbf{w} \cdot \mathbf{w} & \sigma(Q) \mathbf{w} \cdot \mathbf{u} \\ \sigma(Q) \mathbf{u} \cdot \mathbf{v} & \sigma(Q) \mathbf{u} \cdot \mathbf{w} & \sigma(Q) \mathbf{u} \cdot \mathbf{u} \end{pmatrix} = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \delta & 0 \\ 0 & 0 & I \end{pmatrix}$$

Se adesso consideriamo la matrice simmetrica 2×2 , $\begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix}$, ovvero la rappresentazione di $\sigma(Q)$ sul piano π (piano ortogonale a \mathbf{u}), si deduce che la stessa ammette una coppia di autovettori ortogonali (entrambi giacenti su π), che denoteremo con \mathbf{v}' e \mathbf{w}' , rispetto ai quali assume la forma diagonale $\begin{pmatrix} \gamma & 0 \\ 0 & \theta \end{pmatrix}$. Ma allora, nella base ortonormale $\{\mathbf{u}, \mathbf{v}', \mathbf{w}'\}$, $\sigma(Q)$ viene rappresentata dalla seguente matrice

$$\widehat{\mathbb{I}}_Q = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \theta & 0 \\ 0 & 0 & I \end{pmatrix}$$

Calcolando adesso gli invarianti di $\widehat{\mathbb{I}}_Q$, otteniamo il seguente sistema

$$\begin{cases} \gamma + \theta + I = 3I, \\ I\gamma\theta = I^3, \end{cases} \quad \Rightarrow \quad \begin{cases} \gamma = I, \\ \theta = I. \end{cases}$$

Abbiamo quindi determinato tre vettori ortogonali, \mathbf{u} , \mathbf{v}' e \mathbf{w}' , che sono autovettori di $\sigma(Q)$. Non solo, ma nella base $\{\mathbf{u}, \mathbf{v}', \mathbf{w}'\}$, la matrice $\widehat{\mathbb{I}}_Q$ è multipla della matrice identità. Quindi ogni matrice d'inerzia sarà un multiplo della matrice identità. Di conseguenza ogni vettore è autovettore di $\sigma(Q)$, ovvero ogni direzione è asse principale d'inerzia.

□

Concludiamo la sezione con alcuni esempi.

Esempio 7.3.1 *Determinare il momento d'inerzia di un anello omogeneo di massa M , raggio R e spessore trascurabile, rispetto ad una retta passante per il centro e ortogonale al piano dell'anello.*

Applicando semplicemente la definizione 7.3.1 si ha $I = MR^2$.

Esempio 7.3.2 *Determinare la matrice d'inerzia di un disco omogeneo di massa M e raggio R , rispetto ad una terna centrata nel centro del disco, il cui asse z è perpendicolare al piano del disco (v. figura 7.5). L'asse z è asse principale d'inerzia così*

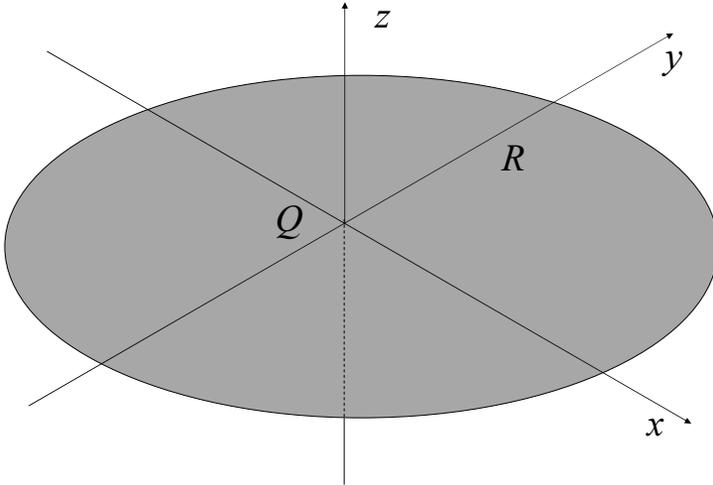


Figura 7.5: Disco omogeneo di massa M e raggio R .

come lo sono l'asse x e l'asse y (v. proposizione 7.3.3). Per determinare I_3 (momento rispetto a z) si ricorre all'osservazione 7.3.2, immaginando di suddividere il disco in un'infinità di "anellini" di raggio r , $0 < r < R$, e spessore infinitesimo dr . La massa infinitesima di ciascun "anellino" sarà

$$dm = \varkappa 2\pi r dr$$

essendo \varkappa la massa per unità di superficie (densità superficiale di massa) $\varkappa = \frac{M}{\pi R^2}$. Avremo⁵

$$I_3 = \int_0^R r^2 dm = 2\pi \varkappa \int_0^R r^3 dr = \frac{MR^2}{2}.$$

Inoltre, dal momento che $I_x = I_y$, per ovvie ragioni di simmetria, sfruttando il corollario 1, avremo

$$2I_x = \frac{MR^2}{2}, \Rightarrow I_x = I_y = \frac{MR^2}{4}.$$

⁵Osserviamo che questa procedura è equivalente a calcolare, in coordinate polari, l'integrale $\int_D (x^2 + y^2) \varkappa dx dy$, dove $\varkappa = M/\pi R^2$ e

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < R^2\}.$$

La matrice d'inerzia sarà dunque

$$\mathbb{I}_Q = \begin{pmatrix} \frac{MR^2}{4} & 0 & 0 \\ 0 & \frac{MR^2}{4} & 0 \\ 0 & 0 & \frac{MR^2}{2} \end{pmatrix}. \quad (7.30)$$

Esempio 7.3.3 Determinare il momento d'inerzia di un segmento di massa M e lunghezza L , rispetto ad una retta ortogonale al segmento stesso passante per il suo centro.

Anche in questo caso si ricorre alla nota 7.3.2, considerando il segmento come l'unione di "massetine infinitesime" dm , la cui posizione sulla retta individuata dal segmento, è identificata con la coordinata x , $-L/2 < x < L/2$, essendo $x = 0$ il centro del segmento. Introducendo la massa per unità di lunghezza $\lambda = M/L$, possiamo scrivere

$$dm = \lambda dx,$$

e quindi, il momento d'inerzia rispetto ad una qualunque retta passante per il centro del segmento e ad esso ortogonale sarà

$$I = \int_{-L/2}^{L/2} x^2 \lambda dx = \frac{ML^2}{12}.$$

Esempio 7.3.4 Determinare la matrice d'inerzia di una lamina rettangolare omogenea di massa M e lati a e b , rispetto ad una terna $\{x, y, z\}$, centrata in Q , come in figura 7.6. L'asse z è asse principale d'inerzia, mentre gli assi x e y lo sono perché ortogonali a piani di simmetria materiale (proposizione 7.3.3). Iniziamo col calcolare il momento d'inerzia rispetto all'asse y , I_y . Seguendo ancora l'osservazione 7.3.2 si suddivide il rettangolo in un'infinità di sbarrette di lunghezza b e spessore infinitesimo dy . La massa di ciascuna sbarretta infinitesima sarà

$$dm = \varkappa b dy,$$

con $\varkappa = \frac{M}{ab}$, densità superficiale di massa. Il momento d'inerzia infinitesimo del segmento che si trova a coordinata y , $-a/2 < y < a/2$, sarà $\frac{b^2}{12} dm = \varkappa \frac{b^3}{12} dy = \frac{Mb^2}{12} \frac{dy}{a}$, (v. esempio 7.3.3) per cui, sfruttando l'additività del momento d'inerzia (si veda la nota 7.3.2),

$$I_y = \frac{Mb^2}{12} \int_{-a/2}^{a/2} \frac{dy}{a} = \frac{Mb^2}{12}.$$

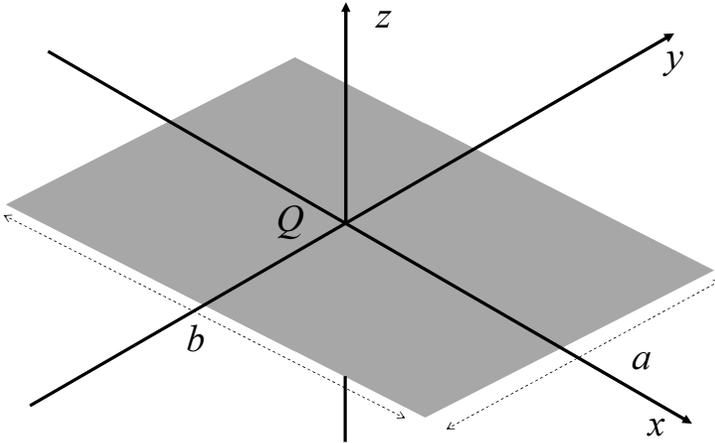


Figura 7.6: Lamina rettangolare omogenea di massa M .

Con lo stesso procedimento si prova che $I_x = \frac{Ma^2}{12}$, per cui, applicando il corollario 7.3.1, si ha

$$\mathbb{I}_Q = \begin{pmatrix} \frac{Ma^2}{12} & 0 & 0 \\ 0 & \frac{Mb^2}{12} & 0 \\ 0 & 0 & \frac{M}{12}(a^2 + b^2) \end{pmatrix}.$$

Esempio 7.3.5 Determinare la matrice d'inerzia di una semidisco omogeneo di massa M e raggio R , rispetto ad una terna $\{x, y, z\}$, centrata in Q , come in figura 7.7. L'asse z è asse principale d'inerzia, così come l'asse x (perchè ortogonale ad un piano di simmetria materiale). Il terzo asse sarà di conseguenza asse principale d'inerzia. Per calcolare I_z , si ricorre ancora all'osservazione 7.3.2, notando che il momento d'inerzia rispetto all'asse z di un disco di massa $2M$ è il doppio di I_z . Avremo così

$$2I_z = \frac{2MR^2}{2}, \Rightarrow I_z = \frac{MR^2}{2}.$$

Allo stesso modo, $2I_x$ è il momento d'inerzia di un disco di massa $2M$ rispetto all'asse x ,

$$2I_x = \frac{2MR^2}{4}, \Rightarrow I_x = \frac{MR^2}{4}.$$

La matrice \mathbb{I}_Q è quindi identica alla (7.30).

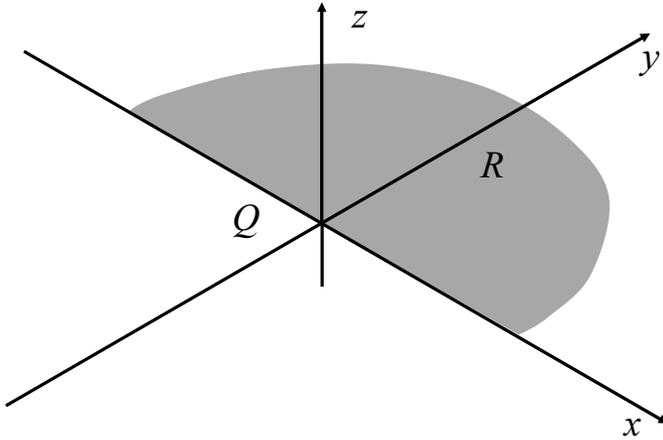


Figura 7.7: Semidisco di massa M .

Esempio 7.3.6 *Data una lamina piana omogenea, a forma di triangolo isoscele, avente massa M , base a ed altezza h , determinare il momento d'inerzia rispetto ad una retta ortogonale alla base e passante per il vertice del triangolo (retta r in figura 7.8 (I)). Si indica con I il momento d'inerzia che vogliamo determinare. La figura 7.8 (II) rappresenta la metà del triangolo isoscele, ovvero un triangolo rettangolo di base $a/2$ ed altezza h , la cui massa è $M/2$, ed è indicato con A . Il momento della lamina A , rispetto alla retta (1) (v. ancora figura 7.8 (II)) si indica con $I_A^{(1)}$. Ricordando l'osservazione 7.3.2, avremo*

$$I = 2I_A^{(1)}. \quad (7.31)$$

Ora consideriamo il rettangolo (che denotiamo con $A \cup B$) di figura 7.8 (II), dato dal triangolo A , e dal triangolo "virtuale" B . Il momento d'inerzia del rettangolo $A \cup B$ (avente massa M e lati h e $a/2$) rispetto alla retta (1) sarà

$$I_{A \cup B}^{(1)} = I_A^{(1)} + I_B^{(1)}, \quad (7.32)$$

dove $I_B^{(1)}$ rappresenta il momento d'inerzia del triangolo "virtuale" B rispetto alla retta (1). Quest'ultimo sarà uguale al momento d'inerzia di A rispetto alla retta (2), $I_A^{(2)}$, ovvero $I_B^{(1)} = I_A^{(2)}$. Quindi tornando alla (7.32) si ha

$$I_{A \cup B}^{(1)} = I_A^{(1)} + I_A^{(2)}, \quad (7.33)$$

dove sappiamo quanto vale $I_{A \cup B}^{(1)}$ (è sufficiente ricordare l'esempio 7.3.4 ed il teorema di Huygens)

$$I_{A \cup B}^{(1)} = \frac{1}{12}M(a/2)^2 + M\left(\frac{a/2}{2}\right)^2 = \frac{M}{3}(a/2)^2. \quad (7.34)$$

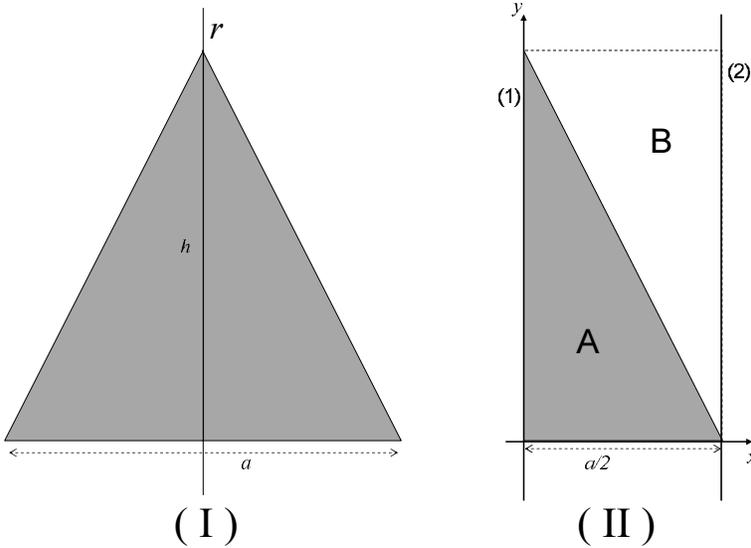


Figura 7.8: Lamina piana omogenea a forma di triangolo isoscele.

Dobbiamo quindi valutare sia $I_A^{(1)}$, che $I_A^{(2)}$. Anche qui si ricorre al teorema di Huygens sfruttando il fatto che, riferendoci al sistema di riferimento $\{x, y\}$ indicato in figura 7.8 (II), il centro di massa del triangolo A ha le coordinate $\left(\frac{a/2}{3}, \frac{h}{3}\right)$. Pertanto, indicando con I_A^{CM} il momento d'inerzia di A rispetto alla retta passante per il centro di massa e parallela alla retta (1), abbiamo

$$I_A^{(1)} = I_A^{CM} + \frac{M}{2} \left(\frac{a/2}{3}\right)^2.$$

e

$$\begin{aligned} I_A^{(2)} &= I_A^{CM} + \frac{M}{2} \left(2\frac{a/2}{3}\right)^2 = I_A^{(1)} - \frac{M}{2} \left(\frac{a/2}{3}\right)^2 + \frac{M}{2} \left(2\frac{a/2}{3}\right)^2 = \\ &= I_A^{(1)} + \frac{M}{2} (a/2)^2 \left(-\frac{1}{9} + \frac{4}{9}\right) = I_A^{(1)} + \frac{M}{6} (a/2)^2. \end{aligned}$$

Inserendo queste ultime due formule nella (7.33) si ha

$$I_{A \cup B}^{(1)} = 2I_A^{(1)} + \frac{M}{6} (a/2)^2.$$

da cui, ricordando anche la (7.34),

$$2I_A^{(1)} = I_{A \cup B}^{(1)} - \frac{M}{6} (a/2)^2 = \frac{M}{6} (a/2)^2.$$

E infine, dalla (7.31),

$$I = \frac{Ma^2}{24}.$$

Ovviamente il procedimento qui illustrato non è l'unico per determinare I , ma ha il pregio di far spesso ricorso al teorema di Huygens e all'osservazione 7.3.2.

Esempio 7.3.7 Data una lamina omogenea quadrata di massa M e lato a , dimostrare, facendo uso del metodo del cerchio di Mohr che, dato un qualunque punto A sulle diagonali del quadrato, una base $\{e_1, e_2\}$ è principale d'inerzia se $\{e_1, e_2\}$ sono paralleli alle diagonali del quadrato.

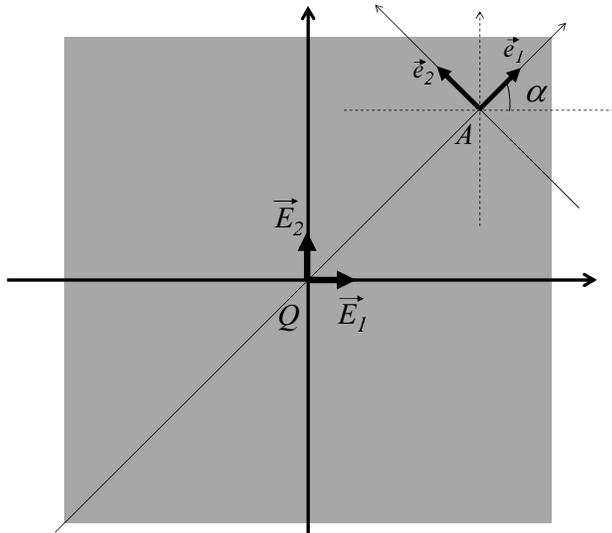


Figura 7.9: Quadrato omogeneo.

Riferendoci alla figura 7.9, si indica con S_Q il sistema $\{\vec{E}_1, \vec{E}_2\}$ centrato in Q , mentre con S_A sistema $\{\vec{e}_1, \vec{e}_2\}$ centrato in A . S_Q è principale d'inerzia e, rispetto ad esso, la matrice (2×2) d'inerzia è

$$\mathbb{I}_Q = \begin{pmatrix} \frac{Ma^2}{12} & 0 \\ 0 & \frac{Ma^2}{12} \end{pmatrix}.$$

In generale il SdR S_A non sarà principale d'inerzia. Infatti la matrice d'inerzia è data da (v. formula (7.18))

$$\mathbb{I}_A = \begin{pmatrix} \frac{Ma^2}{12} + My_A^2 & -Mx_Ay_A \\ -Mx_Ay_A & \frac{Ma^2}{12} + Mx_A^2 \end{pmatrix}.$$

Dobbiamo adesso individuare l'angolo α di cui ruotare S_A al fine di ottenere un nuovo sistema $S'_A : \{ \mathbf{e}_1, \mathbf{e}_2 \}$, che sia principale d'inerzia. Ricordando la (7.18) abbiamo

$$M \frac{x_A^2 - y_A^2}{2} \sin 2\alpha - M x_A y_A \cos 2\alpha = 0.$$

Ora se A giace su una diagonale del quadrato vale $x_A = \pm y_A$, per cui la formula precedente diventa

$$\pm M y_A^2 \cos 2\alpha = 0,$$

le cui due soluzioni sono: $\alpha = \frac{\pi}{4}$, e $\alpha = \frac{\pi}{4} + \frac{\pi}{2}$. Quindi S'_A è principale d'inerzia se $\{ \mathbf{e}_1, \mathbf{e}_2 \}$ sono paralleli alle diagonali del quadrato.

7.4 Le equazioni cardinali

Iniziamo con delle considerazioni generali sulle equazioni di moto di un sistema di punti vere per sistemi qualsiasi, anche non rigidi. Sia (P_k, m_k) , $k = 1, \dots, N$, un sistema di N punti materiali. Se adesso per ogni punto scriviamo la seconda legge di Newton otteniamo N equazioni del tipo

$$m_k \mathbf{a}_k = \mathbf{F}_k^{ext} + \mathbf{F}_k^{int}, \quad k = 1, \dots, N, \quad (7.35)$$

dove abbiamo indicato con \mathbf{a}_k è l'accelerazione del k -esimo punto, con \mathbf{F}_k^{int} e con \mathbf{F}_k^{ext} le forze risultanti di tutte le forze applicate a P_k e distinte a secondo che la loro reazione (in accordo al principio di azione e reazione) risulti applicata ad un altro punto del sistema o a un punto esterno al sistema. Nel primo caso parleremo di **forze interne**, \mathbf{F}_k^{int} , mentre nel secondo caso di **forze esterne**, \mathbf{F}_k^{ext} .

Sommando le N equazioni (7.35) otteniamo

$$\sum_{k=1}^N m_k \mathbf{a}_k = \sum_{k=1}^N \mathbf{F}_k^{ext} + \sum_{k=1}^N \mathbf{F}_k^{int}. \quad (7.36)$$

Ricordando la (7.6), il primo termine della (7.36) è proprio la derivata rispetto a t della quantità di moto⁶ \mathbf{K} .

Osserviamo ora che la somma $\sum_{k=1}^N \mathbf{F}_k^{int}$ è nulla in virtù del principio di azione-reazione; infatti possiamo scrivere

$$\mathbf{F}_k^{int} = \sum_{j=1}^N \mathbf{F}_{j, k}^{int}, \quad (7.37)$$

dove $\mathbf{F}_{j, k}^{int}$ è la forza che si esercita sul punto P_k e la cui reazione è applicata al punto P_j . Si ha quindi $\mathbf{F}_{j, k}^{int} = -\mathbf{F}_{k, j}^{int}$, e i due termini si cancellano nella somma in (7.36). Quindi, ponendo $\mathbf{F}^{ext} = \sum_{k=1}^N \mathbf{F}_k^{ext}$, la (7.36) diventa

$$\sum_{k=1}^N m_k \mathbf{a}_k = \mathbf{F}^{ext}, \quad (7.38)$$

⁶Ovviamente stiamo assumendo che le masse degli N punti materiali rimangano costanti nel tempo.

cioè

$$\frac{d\mathbf{K}}{dt} = M \frac{d^2(P_o - O)}{dt^2} = \mathbf{F}^{ext}, \quad (7.39)$$

dove abbiamo sfruttato la (7.6). Il vettore \mathbf{F}^{ext} prende il nome di **risultante delle forze esterne**.

Possiamo infine “leggere” l’equazione (7.39) come: *la variazione della quantità di moto totale del sistema è uguale alla risultante delle forze esterne applicate al sistema*. Inoltre, in virtù dell’uguaglianza tra \mathbf{K} e $M\mathbf{a}(P_o)$, abbiamo che: *il centro di massa del sistema si muove come un punto materiale, avente come massa la massa totale del sistema, sotto l’azione della risultante delle forze esterne*. L’equazione (7.39) è detta **prima equazione cardinale**.

Rimarchiamo che la (7.39) è stata ottenuta supponendo che le masse m_k , $k = 1, \dots, N$, siano costanti nel tempo e che nel SdR S valga la seconda legge di Newton, ovvero che S sia un SdR inerziale. Osserviamo poi che, come anticipato nell’introduzione del capitolo, l’equazione (7.39) è particolarmente adatta a descrivere la dinamica dei sistemi rigidi in quanto in essa non compaiono le forze vincolari interne che mantengono il vincolo di rigidità.

Riprendiamo ora le equazioni (7.35) e, fissato un centro di riduzione Q , moltiplichiamo ognuna vettorialmente per $(P_k - Q)$ e sommiamo,

$$\sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{a}_k = \sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{ext} + \sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{int}. \quad (7.40)$$

Sempre in virtù del principio di azione-reazione abbiamo che

$$\sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{int} = 0. \quad (7.41)$$

Ricordando la (7.37) abbiamo

$$\sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{int} = \sum_{k=1}^N (P_k - Q) \wedge \sum_{j=1}^N \mathbf{F}_{j,k}^{int},$$

dove $\mathbf{F}_{j,k}^{int}$ è la forza che si esercita sul punto P_k con reazione applicata nel punto P_j . Riscriviamo la somma come

$$\begin{aligned} & \sum_{k=1}^N (P_k - Q) \wedge \sum_{j=1}^N \mathbf{F}_{j,k}^{int} = \sum_{k=1}^N \sum_{j=1}^N (P_k - Q) \wedge \mathbf{F}_{j,k}^{int} \\ &= \sum_{k=1}^N \sum_{j=1}^N [(P_k - P_j) + (P_j - Q)] \wedge \mathbf{F}_{j,k}^{int} \\ &= \sum_{k=1}^N \sum_{j=1}^N (P_k - P_j) \wedge \mathbf{F}_{j,k}^{int} + \sum_{k=1}^N \sum_{j=1}^N (P_j - Q) \wedge \mathbf{F}_{j,k}^{int}. \end{aligned}$$

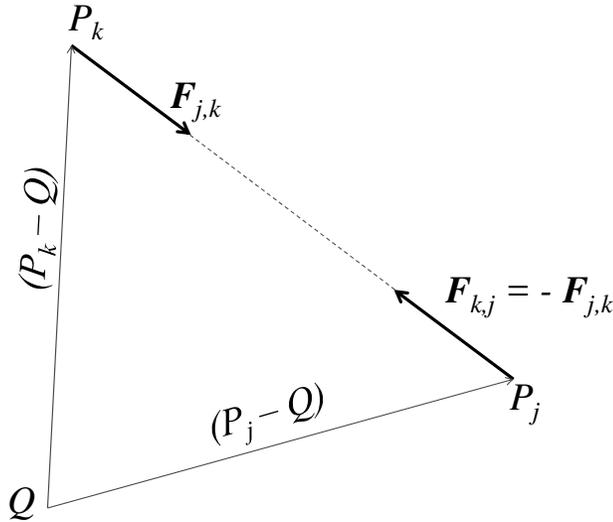


Figura 7.10: La forza $\mathbf{F}_{j,k}^{int}$, è la forza interna agente sul punto materiale P_k . La reazione ad $\mathbf{F}_{j,k}^{int}$ è la forza $\mathbf{F}_{k,j}^{int}$, che agisce sul punto materiale P_j . Dal principio di azione e reazione abbiamo: (i) $\mathbf{F}_{j,k}^{int} = -\mathbf{F}_{k,j}^{int}$; (ii) $\mathbf{F}_{j,k}^{int}$ è parallela al vettore $P_k - P_j$.

Riferendoci ora alla figura 7.10, $(P_k - P_j) \wedge \mathbf{F}_{j,k}^{int} = 0$, in quanto l'azione e la reazione hanno la stessa retta di azione, determinata dalla congiungente dei due punti di applicazione. Quindi

$$\sum_{k=1}^N \sum_{j=1}^N (P_k - Q) \wedge \mathbf{F}_{j,k}^{int} = \sum_{k=1}^N \sum_{j=1}^N (P_j - Q) \wedge \mathbf{F}_{j,k}^{int} = - \sum_{k=1}^N \sum_{j=1}^N (P_j - Q) \wedge \mathbf{F}_{k,j}^{int}.$$

Le due somme nel primo e l'ultimo termine dell'uguaglianza sono la stessa somma: si sono soltanto scambiati gli indici di somma. Ne segue che $\sum_{k=1}^N \sum_{j=1}^N (P_k - Q) \wedge \mathbf{F}_{j,k}^{int}$, è uguale al suo opposto e quindi è nullo e questo prova la (7.41).

Definizione 7.4.1 Il vettore⁷

$$\boldsymbol{\tau}^{ext}(Q) = \sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{ext}, \tag{7.42}$$

è detto *momento risultante delle forze esterne rispetto al punto Q*.

⁷Rimandiamo alla sezione 1.5.2 per le proprietà del momento.

Deriviamo ora rispetto al tempo il momento angolare $\mathbf{L}(Q)$ dato dalla (7.7)

$$\begin{aligned}
 \frac{d}{dt}\mathbf{L}(Q) &= \frac{d}{dt} \left[\sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{v}_k \right] \\
 &= \frac{d}{dt} \left\{ \sum_{k=1}^N m_k [(P_k - O) - (Q - O)] \wedge \mathbf{v}_k \right\} \\
 &= \sum_{k=1}^N m_k \left[\frac{d}{dt} (P_k - O) - \frac{d}{dt} (Q - O) \right] \wedge \mathbf{v}_k + \sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{a}_k \\
 &= \sum_{k=1}^N (\mathbf{v}_k - \mathbf{v}_Q) \wedge m_k \mathbf{v}_k + \sum_{k=1}^N (P_k - Q) \wedge \underbrace{m_k \mathbf{a}_k}_{\mathbf{F}_k^{ext} + \mathbf{F}_k^{int}} \\
 &= -\mathbf{v}_Q \wedge \underbrace{\sum_{k=1}^N m_k \mathbf{v}_k}_{\mathbf{K} \quad (7.5)} + \underbrace{\sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{ext}}_{\boldsymbol{\tau}^{ext}(Q) \quad (7.42)} + \underbrace{\sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{int}}_{=0 \quad (7.41)},
 \end{aligned}$$

da cui abbiamo

$$\frac{d}{dt}\mathbf{L}(Q) = \boldsymbol{\tau}^{ext}(Q) - \mathbf{v}_Q \wedge \mathbf{K}. \quad (7.43)$$

La (7.43) è detta **seconda equazione cardinale**. Il termine $\mathbf{v}_Q \wedge \mathbf{K}$ può essere reso nullo con una opportuna scelta del centro di riduzione Q . Infatti se il punto Q ha velocità parallela alla velocità del centro di massa questo termine si annulla (si ricordi la (7.5)). Le scelte più utilizzate sono Q fisso oppure $Q \equiv P_o$; con queste scelte abbiamo $\frac{d}{dt}\mathbf{L}(Q) = \boldsymbol{\tau}^{ext}(Q)$, ovvero *la variazione del momento della quantità di moto uguaglia il momento risultante delle forze esterne*.

7.5 Le equazioni cardinali sono sufficienti per determinare il moto dei rigidi

Le equazioni cardinali formano un sistema di sei equazioni scalari, valide per qualsiasi sistema materiale (sia discreto che continuo). In generale esse risultano insufficienti per determinare il moto del sistema, non solo nel caso in cui i gradi di libertà siano più di sei, ma anche in casi semplici come quello di due punti isolati tra loro interagenti con una forza qualsiasi (problema a due corpi): in questo caso il sistema ha 6 gradi di libertà ma il moto è solo in parte determinato dalle equazioni cardinali (che in questo caso ci dicono che il centro di massa del sistema si muove di moto rettilineo uniforme e che si conserva il momento della quantità di moto rispetto al centro di massa); non possiamo tuttavia ricavare dalle equazioni cardinali la distanza, variabile nel tempo, tra i due punti in quanto la sua variazione dipende dalla forza di interazione tra i due punti, che è una forza interna.

Nel caso dei sistemi rigidi liberi le equazioni cardinali sono anche sufficienti per determinare il moto del sistema. Questo è dovuto al fatto che esse sono equivalenti all'equazione simbolica della dinamica (5.22), che, evitando la scomposizione (5.20), si scrive come⁸

$$\sum_{k=1}^N (\mathbf{F}_k - m_k \mathbf{a}_k) \cdot \delta \boldsymbol{\xi}_k = 0, \quad \forall (\delta \boldsymbol{\xi}_1, \dots, \delta \boldsymbol{\xi}_N), \quad (7.44)$$

dove, riprendendo la notazione del capitolo 6, sezione 6.2, $\delta \boldsymbol{\xi}_k$, $k = 1, \dots, N$, è lo spostamento virtuale del k -esimo punto materiale del sistema rigido compatibile con il vincolo di rigidità. Supponendo il sistema rigido libero (cioè non soggetto a vincoli se non a quello di rigidità), possiamo scrivere la (6.17) come

$$\boldsymbol{\xi}_k = \boldsymbol{\xi}_Q(q_1, q_2, q_3) + \mathbb{A}(\varphi_1, \varphi_2, \varphi_3) \mathbf{x}_k, \quad k = 1, \dots, N,$$

dove Q è un **punto solidale con il corpo rigido** e:

- (q_1, q_2, q_3) sono i tre gradi di libertà che identificano la posizione del punto Q (per esempio le tre coordinate cartesiane di Q rispetto al SdR fisso);
- $(\varphi_1, \varphi_2, \varphi_3)$ sono gli altri tre parametri lagrangiani che individuano la rotazione (per esempio i tre angoli di Eulero);
- \mathbf{x}_k , rappresenta il vettore $(P_k - Q)$, vettore posizione del k -esimo punto materiale nel SdR solidale⁹.

Nel seguito di questo paragrafo, dopo aver trovato la forma degli spostamenti “virtuali rigidi”, mostreremo, sfruttando l'arbitrarietà degli spostamenti virtuali, che dalle (7.44) si ottengono proprio le equazioni cardinali.

Cominciamo col determinare lo spostamento virtuale $\delta \boldsymbol{\xi}_k$. Applicando la (4.26) abbiamo

$$\begin{aligned} \delta \boldsymbol{\xi}_k &= \sum_{i=1}^3 \frac{\partial \boldsymbol{\xi}_Q}{\partial q_i} \delta q_i + \sum_{i=1}^3 \frac{\partial \mathbb{A}}{\partial \varphi_i} \mathbf{x}_k \delta \varphi_i = \sum_{i=1}^3 \frac{\partial \boldsymbol{\xi}_Q}{\partial q_i} \delta q_i + \sum_{i=1}^3 \frac{\partial \mathbb{A}}{\partial \varphi_i} \underbrace{(\mathbb{A}^T \mathbb{A})}_{\mathbb{I}} \mathbf{x}_k \delta \varphi_i \\ &= \sum_{i=1}^3 \frac{\partial \boldsymbol{\xi}_Q}{\partial q_i} \delta q_i + \left[\sum_{i=1}^3 \frac{\partial \mathbb{A}}{\partial \varphi_i} \mathbb{A}^T \delta \varphi_i \right] \mathbb{A} \mathbf{x}_k. \end{aligned} \quad (7.45)$$

Ricordando adesso la procedura della sezione 6.3, abbiamo

$$\left[\sum_{i=1}^3 \frac{\partial \mathbb{A}}{\partial \varphi_i} \mathbb{A}^T \delta \varphi_i \right] \mathbb{A} \mathbf{x}_k = \delta \boldsymbol{\omega} \wedge \mathbb{A} \mathbf{x}_k,$$

⁸Si ricorda che nel vettore \mathbf{F}_k non devono essere considerate le eventuali reazioni vincolari esterne dovute a vincoli lisci, né le forze vincolari interne dovute al vincolo di rigidità. Quest'ultimo è infatti un vincolo liscio, come mostrato nell'esempio 5.2.1.

⁹Ovviamente si conoscono i vettori posizione, rispetto al SdR solidale, di tutti i punti che costituiscono il sistema rigido, dal momento che la struttura geometrico-materiale del corpo è data.

dove $\delta\boldsymbol{\omega} = \sum_{i=1}^3 \frac{\partial \mathbb{A}}{\partial \varphi_i} \mathbb{A}^T \delta\varphi_i$, è lo spostamento virtuale associato alla rotazione.

Ponendo¹⁰ $(P_k - Q) = \mathbb{A}\mathbf{x}_k$, e $\delta\xi_Q = \sum_{i=1}^3 \frac{\partial \xi_Q}{\partial q_i} \delta q_i$, la (7.45) diventa

$$\delta\xi_k = \delta\xi_Q + \delta\boldsymbol{\omega} \wedge (P_k - Q),$$

che, utilizzando una notazione vettoriale, scriveremo anche come

$$\delta(P_k - O) = \delta(Q - O) + \delta\boldsymbol{\omega} \wedge (P_k - Q), \quad (7.46)$$

dove O è il centro del SdR fisso mentre Q è quello del SdR solidale (avevamo già detto che Q è un punto solidale col corpo rigido).

L'equazione simbolica della dinamica (7.44) diventa quindi

$$\sum_{k=1}^N (\mathbf{F}_k - m_k \mathbf{a}_k) \cdot (\delta\xi_Q + \delta\boldsymbol{\omega} \wedge (P_k - Q)) = 0, \quad \forall \delta\xi_Q, \delta\boldsymbol{\omega}. \quad (7.47)$$

Operando la scomposizione $\mathbf{F}_k = \mathbf{F}_k^{ext} + \mathbf{F}_k^{int}$, e ricordando la formula (1.16), la (7.47) può essere così riscritta

$$\begin{aligned} & \sum_{k=1}^N (\mathbf{F}_k - m_k \mathbf{a}_k) \cdot (\delta\xi_Q + \delta\boldsymbol{\omega} \wedge (P_k - Q)) \\ &= \left[\sum_{k=1}^N (\mathbf{F}_k^{ext} + \mathbf{F}_k^{int} - m_k \mathbf{a}_k) \right] \cdot \delta\xi_Q + \\ & \quad \sum_{k=1}^N [\delta\boldsymbol{\omega} \wedge (P_k - Q)] \cdot (\mathbf{F}_k^{ext} + \mathbf{F}_k^{int} - m_k \mathbf{a}_k) \\ &= \left[\sum_{k=1}^N (\mathbf{F}_k^{ext} - m_k \mathbf{a}_k) \right] \cdot \delta\xi_Q + \underbrace{\left(\sum_{k=1}^N \mathbf{F}_k^{int} \right)}_{=0} \cdot \delta\xi_Q + \\ & \quad \left[\sum_{k=1}^N (P_k - Q) \wedge (\mathbf{F}_k^{ext} - m_k \mathbf{a}_k) \right] \cdot \delta\boldsymbol{\omega} + \underbrace{\left(\sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{int} \right)}_{=0} \cdot \delta\boldsymbol{\omega}. \end{aligned}$$

A questo punto sfruttiamo l'arbitrarietà degli spostamenti virtuali, $\delta\xi_Q$ e $\delta\boldsymbol{\omega}$, considerando dapprima $\delta\boldsymbol{\omega} = 0$, e poi $\delta\xi_Q = 0$. Nel primo caso, $\delta\boldsymbol{\omega} = 0$, la (7.44) si riduce a

$$\sum_{k=1}^N (\mathbf{F}_k^{ext} - m_k \mathbf{a}_k) \cdot \delta\xi_Q = 0,$$

¹⁰Volendo utilizzare la stessa notazione della sezione 6.2, avremmo dovuto indicare $\mathbb{A}\mathbf{x}_k$ con ξ_{p_k} .

che, dovendo valere per ogni $\delta \xi_Q$, equivale alla (7.38), ovvero alla (7.39). Se adesso consideriamo $\delta \xi_Q = 0$, otteniamo

$$\left[\sum_{k=1}^N (P_k - Q) \wedge (\mathbf{F}_k^{ext} - m_k \mathbf{a}_k) \right] \cdot \delta \boldsymbol{\omega} = 0,$$

che, per l'arbitrarietà di $\delta \boldsymbol{\omega}$, equivale a

$$\sum_{k=1}^N (P_k - Q) \wedge (\mathbf{F}_k^{ext} - m_k \mathbf{a}_k) = 0. \tag{7.48}$$

Ora, riprendendo la (7.7) e derivandola rispetto al tempo abbiamo

$$\begin{aligned} \frac{d}{dt} \mathbf{L}(Q) &= \sum_{k=1}^N m_k \frac{d}{dt} [(P_k - O) - (Q - O)] \wedge \mathbf{v}_k + \\ &\quad + \sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{a}_k \\ &= \sum_{k=1}^N m_k \underbrace{\mathbf{v}_k \wedge \mathbf{v}_k}_{=0} - \mathbf{v}_Q \wedge \underbrace{\sum_{k=1}^N m_k \mathbf{v}_k}_{\mathbf{K}} + \\ &\quad + \sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{a}_k, \end{aligned}$$

ovvero

$$\sum_{k=1}^N m_k (P_k - Q) \wedge \mathbf{a}_k = \frac{d}{dt} \mathbf{L}(Q) + \mathbf{v}_Q \wedge \mathbf{K},$$

che, sostituita nella (7.48), dà luogo a

$$\underbrace{\sum_{k=1}^N (P_k - Q) \wedge \mathbf{F}_k^{ext}}_{\boldsymbol{\tau}^{ext}(Q)} - \frac{d}{dt} \mathbf{L}(Q) - \mathbf{v}_Q \wedge \mathbf{K} = 0,$$

cioè alla (7.43).

7.6 Momento angolare, energia cinetica e seconda equazione cardinale per i sistemi rigidi

Affrontiamo adesso tre argomenti fondamentali di questo capitolo: la scrittura del momento angolare per un corpo rigido (rispetto ad un centro di riduzione solidale col corpo stesso), la scrittura della seconda equazione cardinale e la scrittura dell'energia

cinetica di un corpo rigido (il cosiddetto **Teorema di König**¹¹). Illustreremo poi un argomento strettamente collegato alla seconda equazione cardinale: il momento delle reazioni vincolari nel caso di rotazioni uniformi.

7.6.1 Momento angolare per un sistema rigido

Torniamo all'espressione del momento angolare, o momento della quantità di moto, (7.7). Nel caso di un sistema rigido possiamo sostituire l'espressione di \mathbf{v}_k con la formula fondamentale dei moti rigidi (6.32), se il polo Q è un **punto solidale con il sistema rigido**. Nel caso in cui il centro di riduzione non sia solidale la formula è leggermente più complicata e si invita il lettore a ricavarla per esercizio. Otteniamo così

$$\begin{aligned} \mathbf{L}(Q) &= \sum_{k=1}^N m_k (P_k - Q) \wedge [\mathbf{v}(Q) + \boldsymbol{\omega} \wedge (P_k - Q)] \\ &= M(P_o - Q) \wedge \mathbf{v}(Q) + \sum_{k=1}^N m_k (P_k - Q) \wedge [\boldsymbol{\omega} \wedge (P_k - Q)]. \end{aligned} \quad (7.49)$$

Anche questa espressione può essere semplificata scegliendo opportunamente il polo Q . Se Q è fisso sì che $\mathbf{v}(Q) = 0$ (cioè abbiamo un moto rigido con un punto fisso, detto moto di precessione), oppure se $Q \equiv P_o$, la (7.49) si semplifica in

$$\mathbf{L}(Q) = \sum_{k=1}^N m_k (P_k - Q) \wedge [\boldsymbol{\omega} \wedge (P_k - Q)] \quad (7.50)$$

che, lo ricordiamo ancora, è valida soltanto se Q è solidale con il sistema rigido.

Concentrandoci sulla parte di destra della (7.50) e ricordando la definizione 7.3.3, abbiamo

$$\mathbf{L}(Q) = \boldsymbol{\sigma}(Q)\boldsymbol{\omega}. \quad (7.51)$$

Ora, come illustrato nella sezione 7.3.2, l'operatore $\boldsymbol{\sigma}(Q)$ è rappresentato da una matrice \mathbb{I}_Q una volta che viene fissato sistema di riferimento¹² centrato in Q . Se il SdR in cui si calcolano le componenti della matrice \mathbb{I}_Q ha gli assi fissi mentre il corpo si muove, allora gli elementi della matrice (7.16) dipendono dal tempo. *E' quindi evidente che conviene selezionare un SdR cartesiano ortogonale i cui assi siano solidali col corpo rigido, ovvero un sistema di riferimento solidale S centrato in Q .* Non solo, ma tra tutti i sistemi di riferimento solidali conviene scegliere quelli in cui \mathbb{I}_Q è diagonale, ovvero SdR che sono terne principali d'inerzia.

La selezione di un riferimento solidale (possibilmente una terna principale d'inerzia) porta, come detto, alla ovvia conseguenza che gli elementi della matrice \mathbb{I}_Q sono costanti, ma porta anche ad un'altra conseguenza: *il vettore $\mathbf{L}(Q)$ viene ad essere espresso in S , SdR solidale* (v. figura 7.11). Bisogna quindi di aver cura di esprimere anche $\boldsymbol{\omega}$ in S , e soprattutto, bisogna porre particolare attenzione quando si calcola la derivata rispetto al tempo di $\mathbf{L}(Q)$, come vedremo nella prossima sezione.

¹¹Johann Samuel König (Büdingen, 1712 – Zuilenstein, 1757), matematico tedesco.

¹²Ricordiamoci comunque che i sistemi di riferimento devono essere ortonormali.

7.6.2 Seconda equazione cardinale per i sistemi rigidi

La grande differenza fra la dinamica del punto e la dinamica dei sistemi rigidi risiede nella seconda equazione cardinale, che è deputata a descrivere la parte rotatoria del moto. Riferendoci alla figura 7.11, consideriamo un SdR cartesiano ortogonale fisso Σ ed uno, denotato da S , solidale col corpo rigido. Supponiamo inoltre che il centro di riduzione Q sia fisso o coincidente con centro di massa del corpo, sì che $\mathbf{L}(Q)$ assume la forma (7.51).

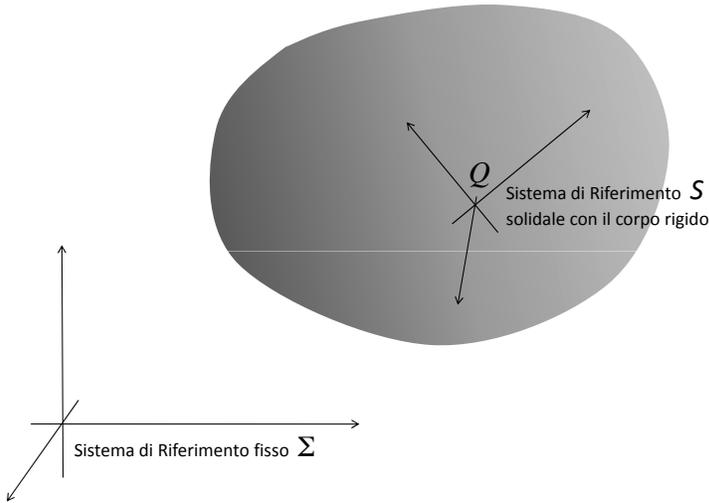


Figura 7.11: Sistema di riferimento fisso Σ e SdR S solidale con il corpo rigido.

Vogliamo calcolare la variazione rispetto al tempo (la derivata temporale) di $\mathbf{L}(Q)$ rispetto all'osservatore Σ . Per mantenere la notazione più chiara possibile indicheremo, ma soltanto in questa e nelle prossime sezioni, con:

- \mathbb{I}_Q^S , la matrice d'inerzia rispetto al SdR solidale S . Come già detto, le componenti di \mathbb{I}_Q^S , sono costanti.
- \mathbb{I}_Q^Σ , la matrice d'inerzia rispetto al SdR fisso Σ , o meglio, la matrice d'inerzia rispetto ad un SdR i cui assi sono paralleli a quelli di Σ ed è centrato in Q . Ovviamente le le componenti di \mathbb{I}_Q^Σ , non sono costanti nel tempo. Infatti, ricordando la (7.17), abbiamo

$$\mathbb{I}_Q^\Sigma = \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T, \tag{7.52}$$

dove $\mathbb{A} = \mathbb{A}(t)$, è la matrice di rotazione (6.7), introdotta nella sezione 6.2.

Riprendendo la notazione della sezione 6.3, se $\boldsymbol{\omega}$ è il vettore velocità angolare, denotiamo con $\boldsymbol{\omega}_\Sigma$ la terna delle componenti $\boldsymbol{\omega}$ rispetto al SdR fisso Σ e con $\boldsymbol{\omega}_S$ la terna

delle componenti ω rispetto al SdR solidale S . Quindi

$$\omega \longrightarrow \begin{cases} \omega_\Sigma, & \text{terna delle componenti rispetto a } \Sigma, \\ \omega_S, & \text{terna delle componenti rispetto a } S. \end{cases}$$

In particolare, ricordando la (6.52), $\omega_\Sigma = \mathbb{A}\omega_s$.

L'espressione di $L(Q)$ nel SdR fisso Σ oppure nel SdR solidale S è dunque (7.51)

$$L(Q) \stackrel{(7.51)}{=} \sigma(Q)\omega \longrightarrow \begin{cases} \mathbb{I}_Q^\Sigma \omega_\Sigma, & \text{nel SdR } \Sigma, \\ \mathbb{I}_Q^S \omega_S, & \text{nel SdR } S. \end{cases}$$

Sfruttando la (7.52), è facile ricavare la relazione che lega le due espressioni

$$\mathbb{I}_Q^\Sigma \omega_\Sigma = \mathbb{A} \mathbb{I}_Q^S \underbrace{\mathbb{A}^T \omega_\Sigma}_{\omega_s} = \mathbb{A} \mathbb{I}_Q^S \omega_s. \quad (7.53)$$

Per calcolare la variazione nel tempo di $L(Q)$ rispetto all'osservatore Σ , deriviamo rispetto al tempo $\mathbb{I}_Q^\Sigma \omega_\Sigma$. Tenendo presenti la (7.53) e la (6.26), otteniamo

$$\begin{aligned} \frac{d}{dt} (\mathbb{I}_Q^\Sigma \omega_\Sigma) &= \frac{d}{dt} (\mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \omega_\Sigma) \\ &= \frac{d\mathbb{A}}{dt} \mathbb{I}_Q^S \mathbb{A}^T \omega_\Sigma + \mathbb{A} \mathbb{I}_Q^S \frac{d\mathbb{A}^T}{dt} \omega_\Sigma + \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \frac{d\omega_\Sigma}{dt} \\ &= \frac{d\mathbb{A}}{dt} \underbrace{\mathbb{A}^T \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \omega_\Sigma}_{\mathbb{I}} + \mathbb{A} \mathbb{I}_Q^S \underbrace{\mathbb{A}^T \mathbb{A}}_{\mathbb{I}} \frac{d\mathbb{A}^T}{dt} \omega_\Sigma + \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \frac{d\omega_\Sigma}{dt} \\ &= \left(\frac{d\mathbb{A}}{dt} \mathbb{A}^T \right) (\mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \omega_\Sigma) + \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \underbrace{\left(\mathbb{A} \frac{d\mathbb{A}^T}{dt} \right)}_{-\frac{d\mathbb{A}}{dt} \mathbb{A}^T} \omega_\Sigma + \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \frac{d\omega_\Sigma}{dt}. \end{aligned}$$

Ora, rammentando la formula (6.30), cioè $\frac{d\mathbb{A}}{dt} \mathbb{A}^T \mathbf{v} = \omega \wedge \mathbf{v}$, abbiamo

$$\begin{aligned} \frac{d}{dt} (\mathbb{I}_Q^\Sigma \omega_\Sigma) &= \omega_\Sigma \wedge \underbrace{(\mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T)}_{\mathbb{I}_Q^\Sigma} \omega_\Sigma - \underbrace{\mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T \omega_\Sigma \wedge \omega_\Sigma}_{=0} + \underbrace{(\mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T)}_{\mathbb{I}_Q^\Sigma} \frac{d\omega_\Sigma}{dt} \\ &= \omega_\Sigma \wedge \mathbb{I}_Q^\Sigma \omega_\Sigma + \mathbb{I}_Q^\Sigma \frac{d\omega_\Sigma}{dt}. \end{aligned}$$

Notiamo che la stessa equazione si ottiene facendo uso della formula di Poisson (6.69). Infatti, sfruttando la seconda uguaglianza della (7.53), cioè $\mathbb{I}_Q^\Sigma \omega_\Sigma = \mathbb{A} \mathbb{I}_Q^S \omega_S$,

otteniamo

$$\begin{aligned} \left. \frac{d}{dt} (\mathbb{I}_Q^\Sigma \omega_\Sigma) \right|_A &= \frac{d}{dt} (\mathbb{A} \mathbb{I}_Q^S \omega_S) = \frac{d\mathbb{A}}{dt} \mathbb{I}_Q^S \omega_S + \mathbb{A} \mathbb{I}_Q^S \frac{d\omega_S}{dt} = \frac{d\mathbb{A}}{dt} \underbrace{\mathbb{A}^T \mathbb{A}}_I \mathbb{I}_Q^S \omega_S + \mathbb{A} \mathbb{I}_Q^S \frac{d\omega_S}{dt} \\ &= \omega_\Sigma \wedge \underbrace{\mathbb{A} \mathbb{I}_Q^S \omega_S}_{\mathbb{I}_Q^\Sigma \omega_\Sigma} + \underbrace{\mathbb{A} \mathbb{I}_Q^S \frac{d\omega_S}{dt}}_{\left. \frac{d(\mathbb{I}_Q^\Sigma \omega_\Sigma)}{dt} \right|_R} = \omega_\Sigma \wedge \mathbb{I}_Q^\Sigma \omega_\Sigma + \left. \frac{d(\mathbb{I}_Q^\Sigma \omega_\Sigma)}{dt} \right|_R \end{aligned}$$

Osserviamo che talvolta conviene esprimere la derivata rispetto al tempo di¹³ $L(Q)$ nel SdR solidale S . In accordo con la (6.52) e la (6.71), abbiamo

$$\begin{aligned} \mathbb{A}^T \frac{d}{dt} (\mathbb{I}_Q^\Sigma \omega_\Sigma) &= \mathbb{A}^T \frac{d}{dt} (\mathbb{A} \mathbb{I}_Q^S \omega_S) = \underbrace{\left(\mathbb{A}^T \frac{d\mathbb{A}}{dt} \right)}_{\omega_S \wedge} \mathbb{I}_Q^S \omega_S + \underbrace{\mathbb{A}^T \mathbb{A}}_I \mathbb{I}_Q^S \frac{d\omega_S}{dt} \\ &= \omega_S \wedge \mathbb{I}_Q^S \omega_S + \mathbb{I}_Q^S \frac{d\omega_S}{dt}. \end{aligned}$$

Quindi, **prescindendo dal SdR, la seconda equazione cardinale per un corpo rigido, nell'ipotesi che il polo Q sia fisso o coincidente col centro di massa¹⁴, si scrive come**

$$\omega \wedge \sigma(Q)\omega + \sigma(Q)\dot{\omega} = \tau^{ext}(Q), \tag{7.54}$$

intendendo che le quantità vettoriali ω , $\tau^{ext}(Q)$ e l'omografia d'inerzia $\sigma(Q)$ che ivi compaiono possono essere espresse sia nel SdR fisso Σ , sia nel SdR solidale S . La scelta di quale SdR utilizzare dipende essenzialmente dalla convenienza nel fare i calcoli.

Alcune osservazioni sono doverose:

1. Notiamo per prima cosa che $\dot{\omega} = 0$, **rotazione uniforme**, non comporta necessariamente $\tau^{ext}(Q) = 0$. Infatti, se $\dot{\omega} = 0$, avremo che $\tau^{ext}(Q) = 0$ **soltanto se ω è parallelo ad un asse principale d'inerzia**, cioè se $\sigma(Q)\omega$ è parallelo ad ω (ovvero ω è un autovettore di $\sigma(Q)$).
2. La (7.54) assume una forma particolarmente semplice nel caso di moti piani. Se infatti e_3 è il versore ortogonale al piano dove si svolge il moto, abbiamo

$$\omega_\Sigma = \omega_S = \dot{\varphi} e_3,$$

e

$$\mathbb{I}_Q = \begin{pmatrix} I_{11} & I_{12} & 0 \\ I_{12} & I_{22} & 0 \\ 0 & 0 & I_{33} \end{pmatrix} \Rightarrow \mathbb{I}_Q \omega = I_{33} \dot{\varphi} e_3,$$

¹³Attenzione a non fare confusione: la derivata rispetto al tempo di $L(Q)$ viene calcolata dall'osservatore Σ . Tale derivata però, che è un vettore, può essere espressa sia nel SdR fisso Σ , ma anche nel SdR solidale S .

¹⁴In ogni caso Q deve essere solidale con il corpo rigido.

per cui $\omega \wedge \mathbb{I}_Q \omega = \dot{\varphi} \mathbf{e}_3 \wedge I_{33} \dot{\varphi} \mathbf{e}_3 = 0$, e la (7.54) si riduce a

$$I_{33} \ddot{\varphi} = \boldsymbol{\tau}^{ext}(Q) \cdot \mathbf{e}_3. \quad (7.55)$$

La configurazione del sistema è determinata completamente dall'angolo φ , ed il suo stato cinematico da φ e $\dot{\varphi}$. Inoltre, se la componente $\boldsymbol{\tau}^{ext}(Q) \cdot \mathbf{e}_3$ del momento è funzione anch'essa solo di φ (oppure di φ e $\dot{\varphi}$ nel caso sia presente una qualche forma di attrito o resistenza al moto), ne segue che la (7.55) determina completamente il moto rotatorio del sistema. Da un punto di vista "matematico" l'equazione (7.55) è del tutto analoga all'equazione di moto di un punto materiale, dove ora il ruolo della forza è tenuto dal momento delle forze esterne e il ruolo della massa dal momento di inerzia. Possiamo quindi identificare il momento di inerzia di un corpo, relativo ad un dato asse, come la sua tendenza a permanere in uno stato di moto di rotazione uniforme attorno a quell'asse, così come la massa inerziale di un punto materiale misura la tendenza a permanere nel suo stato di moto rettilineo uniforme: far variare la velocità di rotazione sarà tanto più difficile quanto maggiore è il momento d'inerzia.

Esempio 7.6.1 Una lamina rettangolare è vincolata a ruotare uniformemente attorno alla propria diagonale maggiore, come mostrato nella figura 7.12. Si vuole determinare il momento $\boldsymbol{\tau}^{ext}$ che i cuscinetti che sostengono l'asse di rotazione esercitano sull'asse stesso.

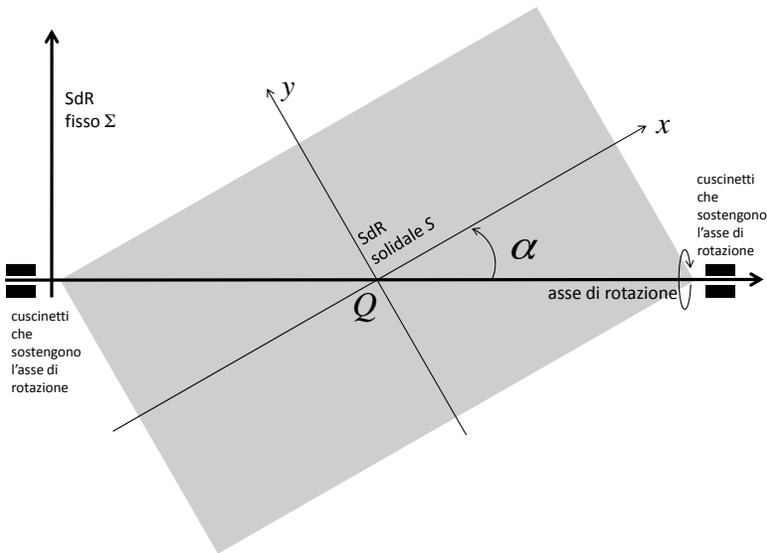


Figura 7.12: Lamina rettangolare in rotazione uniforme attorno alla diagonale principale. Si denota con α l'angolo fra la diagonale della lamina e l'asse x .

Applichiamo la (7.54), scrivendola nel SdR solidale. Il vettore ω , nel SdR fisso è dato dalla terna

$$\omega_{\Sigma} = \begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix}, \quad (7.56)$$

mentre nel SdR solidale è $\omega_S = \mathbb{A}^T \omega_{\Sigma}$, dove, al solito, \mathbb{A} è la matrice di rotazione. Evidentemente, la terna delle componenti della velocità angolare nel SdR solidale è

$$\omega_S = \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix}, \quad (7.57)$$

dal momento che la velocità angolare giace sempre nel piano (x, y) del SdR solidale. Tuttavia, mostriamo come si giunge ad ω_S passando tramite l'usuale relazione $\omega_S = \mathbb{A}^T \omega_{\Sigma}$, con ω_{Σ} data dalla (7.56). Per passare dal SdR fisso a quello solidale, dobbiamo prima ruotare attorno all'asse (diagonale maggiore della lamina) di un angolo $\phi = \omega t$, la cui matrice è

$$\mathbb{A}_{\phi}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix},$$

e poi ruotare attorno all'asse ortogonale al piano della lamina passante per Q , di un angolo α , la cui matrice è

$$\mathbb{A}_{\alpha}^T = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Quindi

$$\begin{aligned} \mathbb{A}^T &= \mathbb{A}_{\alpha}^T \mathbb{A}_{\phi}^T = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha & \cos \phi \sin \alpha & \sin \alpha \sin \phi \\ -\sin \alpha & \cos \alpha \cos \phi & \cos \alpha \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix}. \end{aligned}$$

Si ottiene così

$$\begin{aligned} \omega_S &= \mathbb{A}^T \omega_{\Sigma} = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix}, \end{aligned}$$

cioè la (7.57). Osserviamo adesso che il SdR solidale è una terna principale d'inerzia per cui

$$\mathbb{I}_Q^S = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix},$$

e quindi

$$\mathbb{I}_Q^S \boldsymbol{\omega}_S = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} \omega I_1 \cos \alpha \\ -\omega I_2 \sin \alpha \\ 0 \end{pmatrix},$$

da cui ricaviamo l'espressione del momento applicato dai cuscinetti all'asse nel SdR solidale

$$\begin{aligned} \boldsymbol{\tau}^{ext} &= \boldsymbol{\omega}_S \wedge \mathbb{I}_Q^S \boldsymbol{\omega}_S = \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix} \wedge \begin{pmatrix} \omega I_1 \cos \alpha \\ -\omega I_2 \sin \alpha \\ 0 \end{pmatrix} \\ &= \omega^2 \frac{\sin 2\alpha}{2} (I_1 - I_2) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned} \quad (7.58)$$

Si deduce che $\boldsymbol{\tau}^{ext} = 0$, se $I_1 = I_2$, cioè la lamina è un quadrato, oppure se $\sin 2\alpha = 0$, $\Rightarrow \alpha = 0, \pm \frac{\pi}{2}$, ovvero se l'asse di rotazione coincide con l'asse x o con l'asse y (che sono principali d'inerzia). Osserviamo che l'espressione (7.58) fornisce le componenti di $\boldsymbol{\tau}^{ext}$ nel SdR solidale (nello specifico $\boldsymbol{\tau}^{ext}$ è ortogonale al piano della lamina). L'espressione di $\boldsymbol{\tau}^{ext}$ nel SdR fisso è data da

$$\mathbb{A} \boldsymbol{\tau}^{ext} = \mathbb{A}_\phi \mathbb{A}_\alpha \boldsymbol{\tau}^{ext} = \omega^2 \frac{\sin 2\alpha}{2} (I_1 - I_2) \begin{pmatrix} 0 \\ -\sin \phi \\ \cos \phi \end{pmatrix}, \quad (7.59)$$

dove ϕ varia nel tempo.

E' interessante notare che si giunge esattamente allo stesso risultato se fin da principio si lavora nel SdR fisso. In quest'ottica dobbiamo esprimere la matrice d'inerzia nel SdR fisso, cioè

$$\begin{aligned} \mathbb{I}_Q^\Sigma &= \mathbb{A} \mathbb{I}_Q^S \mathbb{A}^T = \mathbb{A}_\phi \mathbb{A}_\alpha \mathbb{I}_Q^S \mathbb{A}_\alpha^T \mathbb{A}_\phi^T \\ &= \begin{pmatrix} I_1 \cos^2 \alpha + I_2 \sin^2 \alpha & \frac{\sin 2\alpha}{2} \cos \phi (I_1 - I_2) \\ \frac{\sin 2\alpha}{2} \cos \phi (I_1 - I_2) & \cos^2 \phi (I_2 \cos^2 \alpha + I_1 \sin^2 \alpha) + I_3 \sin^2 \phi \\ \frac{\sin 2\alpha}{2} \sin \phi (I_1 - I_2) & \cos \phi \sin \phi (I_2 \cos^2 \alpha + I_1 \sin^2 \alpha - I_3) \\ \frac{\sin 2\alpha}{2} \sin \phi (I_1 - I_2) & \cos \phi \sin \phi (I_2 \cos^2 \alpha + I_1 \sin^2 \alpha - I_3) \\ \cos \phi \sin \phi (I_2 \cos^2 \alpha + I_1 \sin^2 \alpha - I_3) & \sin^2 \phi (I_2 \cos^2 \alpha + I_1 \sin^2 \alpha) + I_3 \cos^2 \phi \end{pmatrix}. \end{aligned}$$

Ora, ricordando che $\phi = \omega t$, la precedente espressione di \mathbb{I}_Q^Σ mette in luce un fatto importante (già sottolineato alla fine della sezione 7.6.1): **in generale gli elementi della matrice d'inerzia espressa nel SdR fisso dipendono dal tempo**. Tornando alla seconda cardinale,

$$\mathbb{I}_Q^\Sigma \boldsymbol{\omega}_\Sigma = \mathbb{I}_Q^\Sigma \begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \omega (I_1 \cos^2 \alpha + I_2 \sin^2 \alpha) \\ \omega (I_1 - I_2) \frac{\sin 2\alpha}{2} \cos \phi \\ \omega (I_1 - I_2) \frac{\sin 2\alpha}{2} \sin \phi \end{pmatrix},$$

per cui le componenti di τ^{ext} nel SdR fisso sono

$$\begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix} \wedge \mathbb{I}_Q \begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix} = \omega^2 \frac{\sin 2\alpha}{2} (I_1 - I_2) \begin{pmatrix} 0 \\ -\sin \phi \\ \cos \phi \end{pmatrix},$$

che coincide con (7.59). Per concludere osserviamo che τ^{ext} è il momento imposto all'asse di rotazione dai cuscinetti i quali subiscono, per il principio di azione e reazione, una sollecitazione uguale ed opposta. Infatti, ponendoci nel SdR solidale scriveremo

$$\omega^2 \frac{\sin 2\alpha}{2} (I_1 - I_2) = 2 \left(\frac{d}{2} F_{cuscinetti} \right),$$

dove $d/2$ è la semidiagonale e $F_{cuscinetti}$ è la forza applicata dai cuscinetti all'asse. La forza che i cuscinetti subiscono è dunque

$$F_{cuscinetti} = \omega^2 \frac{\sin 2\alpha}{2d} (I_1 - I_2).$$

7.6.3 Reazioni vincolari applicate all'asse di rotazione

Alla luce dell'esempio 7.6.1, vogliamo sottolineare ancora il ruolo delle reazioni vincolari nel caso di rotazioni. Supponiamo che il sistema rigido sia vincolato a muoversi di moto rotatorio attorno ad un asse fisso, che, senza perdere in generalità, assumeremo coincidere con l'asse e_3 . Nello specifico vogliamo:

1. determinare le condizioni che devono valere affinché i vincoli che impongono la rotazione attorno ad un asse si possano dire lisci;
2. determinare le reazioni vincolari, o meglio il momento delle reazioni, nel caso di rotazioni uniformi.

Per prima cosa vediamo cosa si deve intendere per *vincolo di rotazione liscio*. Applicando la definizione 5.2.1, i vincoli che impongono al corpo di ruotare attorno all'asse e_3 sono lisci se le reazioni vincolari R_k^{ext} soddisfano la condizione

$$\sum_k R_k^{ext} \cdot \delta (P_k - O) = 0, \tag{7.60}$$

per ogni sistema di spostamenti virtuali $\delta (P_k - O)$ compatibile con i vincoli. La maniera più naturale per vincolare un rigido a ruotare attorno ad un asse fisso è quello di collocare una o più cerniere in altrettanti punti della retta il cui versore è e_3 , come mostrato nella figura 7.13

Poiché l'asse di rotazione è fisso, supponiamo che il centro di riduzione Q ed O coincidano e che Q giaccia sull'asse di rotazione (quindi $Q \equiv O$, ancorché solidale col sistema rigido, è fisso). La (7.46) comporta

$$\delta (P_k - O) = \delta \varphi e_3 \wedge (P_k - Q),$$

dove, come detto, e_3 è il versore dell'asse di rotazione. Usando l'arbitrarietà di $\delta \varphi$ nella (7.60) abbiamo

$$\sum_k R_k^{ext} \cdot [e_3 \wedge (P_k - Q)] \stackrel{(1.16)}{=} \sum_k [(P_k - Q) \wedge R_k^{ext}] \cdot e_3 = 0. \tag{7.61}$$



Figura 7.13: Corpo rigido in rotazione attorno ad un asse fisso.

Diremo quindi che **il corpo rigido ruota senza attrito attorno all'asse**, ovvero che il vincolo di **rotazione attorno all'asse è liscio**, se le forze di reazione dovute al vincolo di rotazione hanno momento con componente nulla lungo l'asse di rotazione¹⁵. Si osservi tuttavia che la (7.61) **non comporta assolutamente che le reazioni vincolari R_k^{ext} siano nulle**¹⁶.

Nota 7.6.1 *Se consideriamo un moto piano e supponiamo che il vincolo di rotazione sia liscio, nel termine di destra della (7.55), cioè in $\tau^{ext}(Q) \cdot e_3$, le reazioni vincolari non compaiono. In altri termini, $\tau^{ext}(Q) \cdot e_3$ è dovuto alle sole forze esterne direttamente applicate. Se poi $\tau^{ext}(Q) \cdot e_3 = 0$, cioè se il momento risultante delle forze esterne applicate è nullo, la rotazione risulterà uniforme, cioè $\dot{\varphi} = \text{costante}$.*

¹⁵ Allo stesso modo si dimostra che se si impone al sistema il vincolo di mantenere un punto fisso Q durante il moto, la condizione di vincolo liscio equivale a dire che il momento delle reazioni vincolari rispetto ad Q deve essere nullo.

¹⁶ Ricordiamo, a titolo di esempio, che nel caso di un punto materiale vincolato su una superficie, il vincolo (la superficie) si dice liscio non se la reazione vincolare è nulla, ma se questa è ortogonale al vincolo.

Concentriamoci adesso sul secondo punto, supponendo che la rotazione sia uniforme, cioè

$$\boldsymbol{\omega} = \dot{\varphi} \mathbf{e}_3, \quad \text{con } \dot{\varphi} = \text{costante},$$

e che non vi siano forze esterne direttamente applicate al corpo. La (7.54) diventa

$$\boldsymbol{\omega} \wedge \mathbb{I}_Q \boldsymbol{\omega} = \boldsymbol{\tau}^{ext}(Q),$$

dove $\boldsymbol{\tau}^{ext}(Q)$ è il momento delle sole reazioni vincolari. Mostriamo innanzitutto che il vincolo è liscio. Infatti

$$\boldsymbol{\tau}^{ext}(Q) \cdot \mathbf{e}_3 = (\dot{\varphi} \mathbf{e}_3 \wedge \mathbb{I}_Q \dot{\varphi} \mathbf{e}_3) \cdot \mathbf{e}_3 \stackrel{(1.16)}{=} \dot{\varphi}^2 \underbrace{(\mathbf{e}_3 \wedge \mathbf{e}_3)}_{=0} \cdot \mathbb{I}_Q \mathbf{e}_3.$$

Calcoliamo adesso $\boldsymbol{\tau}^{ext}(Q)$. Consideriamo un SdR solidale S in cui l'asse z coincida con l'asse \mathbf{e}_3 . Lavorando in S , avremo

$$\boldsymbol{\omega} = \begin{pmatrix} 0 \\ 0 \\ \dot{\varphi} \end{pmatrix},$$

e

$$\mathbb{I}_Q \boldsymbol{\omega} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \dot{\varphi} \end{pmatrix} = \dot{\varphi} \begin{pmatrix} I_{13} \\ I_{23} \\ I_{33} \end{pmatrix},$$

siccome, in generale, S non sarà una terna principale d'inerzia. Otteniamo così

$$\boldsymbol{\tau}^{ext}(Q) = \boldsymbol{\omega} \wedge \mathbb{I}_Q \boldsymbol{\omega} = \dot{\varphi}^2 \begin{pmatrix} -I_{23} \\ I_{13} \\ 0 \end{pmatrix} \tag{7.62}$$

Quindi, se l'asse di rotazione *non è principale di inerzia*, il termine $\boldsymbol{\omega} \wedge \mathbb{I}_Q \boldsymbol{\omega}$ non si annulla. Le forze vincolari devono fornire un momento per garantire il moto di rotazione attorno all'asse stesso. Si può calcolare il modulo del momento che le reazioni devono fornire (e quindi “quantificare” la sollecitazione che questa rotazione impone ai vincoli)

$$|\boldsymbol{\tau}^{ext}(Q)| = \dot{\varphi}^2 \sqrt{I_{13}^2 + I_{23}^2} \neq 0. \tag{7.63}$$

Si osservi che il momento $\boldsymbol{\tau}^{ext}(Q)$, che è costante nel sistema di riferimento solidale S , *ruota* nel sistema di riferimento fisso, ovvero quello in cui si trovano i dispositivi che realizzano fisicamente i vincoli (cuscinetti, bronzine, etc.). Quindi la presenza di un momento deviatore non nullo tende a “destabilizzare” il moto di rotazione, cercando di far “deviare” l'asse di rotazione dalla sua direzione “scardinando” (alla lunga) i vincoli. Non solo, ma il modulo del momento dipende dal *quadrato* della velocità angolare. Ciò significa che nei sistemi rigidi che ruotano ad altissime velocità, come per esempio le turbine, è di cruciale importanza posizionare l'asse di rotazione in corrispondenza di un asse principale d'inerzia.

Nota 7.6.2 Nonostante l'apparenza "esoterica", questo fenomeno dovrebbe essere abbastanza noto a chiunque guidi un'automobile: sono le vibrazioni che si avvertono sullo sterzo quando si guida un'auto che abbia subito una (lieve) deformazione di una delle ruote a seguito, per esempio, di un urto con un marciapiede o una buca. Il "rimedio" è la cosiddetta "equilibratura" della ruota: si aggiunge un peso in un punto opportuno sulla parte esterna del cerchio della ruota. L'effetto è di far tornare l'asse di rotazione della ruota un asse principale di inerzia, togliendo quindi le sollecitazioni sul semiasse a cui è attaccata la ruota.

Esempio 7.6.2 Prendendo spunto dalla nota precedente, calcoliamo il momento imposto dalle forze vincolari all'asse di rotazione di un disco che sta uniformemente ruotando attorno alla retta z , come mostrato in figura 7.14.

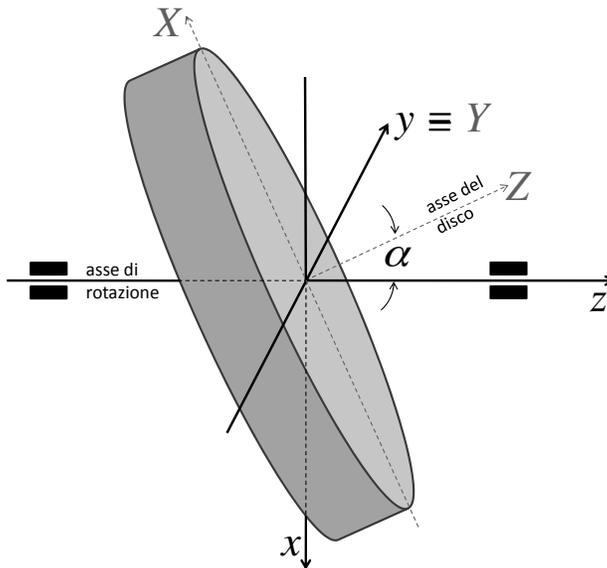


Figura 7.14: Disco in rotazione uniforme alla retta z , che forma un angolo α con l'asse del disco Z . Si individuano due SdR solidali: $S' : \{X, Y, Z\}$, che è una terna principale d'inerzia e $S : \{x, y, z\}$, che è non una terna principale d'inerzia, il cui asse y coincide con Y . Il SdR S è ottenuto da S' ruotando quest'ultimo di un angolo α attorno all'asse Y .

Nel SdR solidale $S' : \{X, Y, Z\}$, la cui origine è nel punto Q , centro del disco, la matrice d'inerzia ha la forma (7.30), cioè

$$\mathbb{I}'_Q = \frac{MR^2}{4} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Il SdR solidale $S : \{x, y, z\}$ si ottiene ruotando S' attorno all'asse Y di un angolo α (come evidenziato dalla figura 7.14). La matrice di rotazione da S' ad S è data da

$$\mathbb{B} = \begin{pmatrix} \cos \alpha & 0 & -\sin \alpha \\ 0 & 1 & 0 \\ \sin \alpha & 0 & \cos \alpha \end{pmatrix}.$$

L'espressione della matrice d'inerzia nel SdR S si ottiene sfruttando la (7.17), cioè

$$\mathbb{I}_Q = \mathbb{B} \mathbb{I}'_Q \mathbb{B}^T = \frac{MR^2}{4} \begin{pmatrix} \cos^2 \alpha + 2 \sin^2 \alpha & 0 & -\cos \alpha \sin \alpha \\ 0 & 1 & 0 \\ -\cos \alpha \sin \alpha & 0 & 2 \cos^2 \alpha + \sin^2 \alpha \end{pmatrix}.$$

Nel SdR S la velocità angolare è $\boldsymbol{\omega}_S = \dot{\varphi} \mathbf{e}_3$, per cui

$$\mathbb{I}_Q \boldsymbol{\omega}_S = \dot{\varphi} \frac{MR^2}{4} \begin{pmatrix} -\cos \alpha \sin \alpha \\ 0 \\ 2 \cos^2 \alpha + \sin^2 \alpha \end{pmatrix},$$

ed il momento applicato dalle forze vincolari all'asse di rotazione è

$$\boldsymbol{\tau}^{ext}(Q) = \boldsymbol{\omega}_S \wedge \mathbb{I}_Q \boldsymbol{\omega}_S = -\dot{\varphi}^2 \frac{MR^2}{8} \begin{pmatrix} 0 \\ \sin 2\alpha \\ 0 \end{pmatrix}.$$

Notiamo che il momento si annulla se $\alpha = 0$, $\pm \frac{\pi}{2}$, ovvero il disco ruota attorno all'asse Z , oppure attorno all'asse X , che sono assi principali d'inerzia.

7.6.4 L'energia cinetica

Il tensore d'inerzia ci fornisce anche una forma compatta per esprimere l'energia cinetica in di un sistema rigido. Selezioniamo un punto Q solidale con il corpo rigido, e sfruttiamo la formula fondamentale dei moti rigidi (6.32)

$$T = \frac{1}{2} \sum_{k=1}^N m_k \mathbf{v}_k^2 = \frac{1}{2} \sum_{k=1}^N m_k [\mathbf{v}(Q) + \boldsymbol{\omega} \wedge (P_k - Q)]^2. \quad (7.64)$$

Sviluppando il quadrato otteniamo

$$T = \frac{1}{2} M \mathbf{v}^2(Q) + \mathbf{v}(Q) \cdot \left[\boldsymbol{\omega} \wedge \sum_{k=1}^N m_k (P_k - Q) \right] + \frac{1}{2} \sum_{k=1}^N m_k [\boldsymbol{\omega} \wedge (P_k - Q)]^2. \quad (7.65)$$

dove M la massa totale del corpo. Anche in questo caso il doppio prodotto può essere reso nullo con una opportuna scelta del punto Q . Infatti:

- se Q è fisso sia il primo che il secondo termine nella somma in (7.65) sono nulli e abbiamo

$$T = \frac{1}{2} \sum_{k=1}^N m_k [\boldsymbol{\omega} \wedge (P_k - Q)]^2 ;$$

- oppure se $Q \equiv P_o$, cioè Q coincide con il centro di massa, la somma tra parentesi nel secondo addendo si annulla e si può scrivere

$$T = \frac{1}{2} M v^2(P_o) + \frac{1}{2} \sum_{k=1}^N m_k [\boldsymbol{\omega} \wedge (P_k - P_o)]^2 . \quad (7.66)$$

La formula (7.66) è nota come **Teorema di König**, e dice che l'energia di un corpo rigido è data dalla somma dell'energia che avrebbe un punto materiale di massa uguale alla massa del corpo, indicata con T_{P_o} , e l'energia cinetica che il corpo ha rispetto a un sistema di riferimento che *trasla* con velocità uguale a quella del centro di massa¹⁷, che indicheremo con T_R .

Procedendo nel calcolo in modo analogo a quanto fatto nel calcolo di $L(Q)$, applichiamo la proposizione 7.3.2

$$T_R = \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{\sigma}(P_o) \boldsymbol{\omega} . \quad (7.67)$$

In particolare, una volta selezionato un SdR cartesiano ortogonale centrato nel centro di massa, avremo

$$T_R = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbb{I}_{P_o} \boldsymbol{\omega} , \quad (7.68)$$

essendo \mathbb{I}_{P_o} l'espressione dell'operatore $\boldsymbol{\sigma}(P_o)$ rispetto al SdR selezionato. Anche in tal caso converrà in generale selezionare un SdR solidale che sia principale d'inerzia in modo che sia \mathbb{I}_{P_o} diagonale. Evidentemente dovremo aver cura di esprimere anche $\boldsymbol{\omega}$ nel medesimo SdR.

Riepilogando

$$T = \begin{cases} \frac{1}{2} M v^2(P_o) + \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{\sigma}(P_o) \boldsymbol{\omega} , & \text{se } Q \text{ coincide con } P_o, \\ \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{\sigma}(Q) \boldsymbol{\omega} , & \text{se } Q \text{ è fisso.} \end{cases} \quad (7.69)$$

Esempio 7.6.3 Consideriamo un sistema rigido come in figura 7.15. Il corpo può ruotare attorno alla retta verticale (con cui l'asta forma un angolo α) e su se stesso. Si seleziona un SdR solidale centro in O , in cui l'asse x coincide con l'asta, l'asse z giace sul piano formato dalla retta verticale e dall'asse x , e l'asse y è ortogonale a tale piano. Tale SdR è una terna principale d'inerzia.

Indichiamo con θ l'angolo di rotazione attorno alla retta verticale e con ϕ quello attorno all'asse x . Quindi, l'espressione di $\boldsymbol{\omega}$ nel SdR solidale è

$$\boldsymbol{\omega} = \begin{pmatrix} \dot{\phi} - \dot{\theta} \cos \alpha \\ 0 \\ \dot{\theta} \sin \alpha \end{pmatrix} ,$$

¹⁷Questo teorema è spesso enunciato dicendo che l'energia cinetica di un sistema rigido è uguale alla somma dell'energia cinetica del centro di massa e dell'energia cinetica relativa al centro di massa. E' bene notare che il senso dei termini usati è quello specificato sopra nel testo: presi "alla lettera" *energia cinetica del centro di massa* e *energia cinetica relativa al centro di massa* non hanno significato.

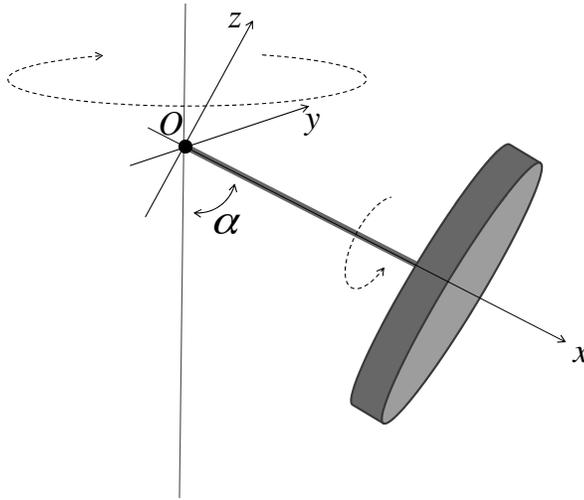


Figura 7.15: Il sistema rigido è composto da un disco di massa M e raggio R , e un'asta, di massa m e lunghezza d , saldata ortogonalmente al disco nel suo centro. L'asta è incernierata nel punto O della retta verticale e mantiene un angolo α con la retta stessa.

dove $\dot{\theta}$ è la velocità di rotazione attorno alla retta verticale mentre $\dot{\phi}$ è la velocità di rotazione del corpo su se stesso. Applicando il teorema 7.3.1 e l'osservazione 7.3.2, la matrice d'inerzia è data da

$$\begin{aligned} \mathbb{I}_O &= \mathbb{I}_O^{asta} + \mathbb{I}_O^{disco} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{md^2}{12} + m\frac{d^2}{4} & 0 \\ 0 & 0 & \frac{md^2}{12} + m\frac{d^2}{4} \end{pmatrix} \\ &\quad + \begin{pmatrix} \frac{MR^2}{2} & 0 & 0 \\ 0 & \frac{MR^2}{4} + Md^2 & 0 \\ 0 & 0 & \frac{MR^2}{4} + Md^2 \end{pmatrix}, \end{aligned}$$

che, per brevità di notazione, indicheremo come

$$\mathbb{I}_O = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_2 \end{pmatrix}.$$

Applicando la (7.69)₂, abbiamo

$$\begin{aligned} T &= \frac{1}{2} \begin{pmatrix} \dot{\phi} - \dot{\theta} \cos \alpha \\ 0 \\ \dot{\theta} \sin \alpha \end{pmatrix} \cdot \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_2 \end{pmatrix} \begin{pmatrix} \dot{\phi} - \dot{\theta} \cos \alpha \\ 0 \\ \dot{\theta} \sin \alpha \end{pmatrix} \\ &= \frac{1}{2} \left[I_1 (\dot{\phi} - \dot{\theta} \cos \alpha)^2 + I_2 \dot{\theta}^2 \sin^2 \alpha \right]. \end{aligned}$$

Esempio 7.6.4 Consideriamo il sistema rigido dell'esempio 7.6.1. Le componenti del vettore $\boldsymbol{\omega}$ nel SdR solidale di figura 7.12, sono

$$\boldsymbol{\omega} = \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix},$$

per cui, applicando sempre la formula (7.69)₂

$$\begin{aligned} T &= \frac{1}{2} \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix} \cdot \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} \omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix} \\ &= \frac{\omega^2}{2} (I_1 \cos^2 \alpha + I_2 \sin^2 \alpha). \end{aligned}$$

Esempio 7.6.5 Calcoliamo l'energia cinetica del disco dell'esempio 4.1.3, applicando la (7.69)₁, dove P_o è il centro Q della figura 4.3. Ricordando la (4.18),

$$\frac{1}{2} M \mathbf{v}^2(Q) = \frac{M}{2} R^2 \dot{\theta}^2.$$

Per quel che riguarda l'energia cinetica dovuta alla rotazione, riferendoci alla figura 6.10, selezioniamo il SdR solidale S' . Come mostrato nell'esempio 6.8.1, la velocità angolare del disco è data da $\boldsymbol{\omega} = \dot{\varphi} \mathbf{e}_{X'} + \dot{\theta} \mathbf{e}_z$, che, come già detto, esprimeremo nel SdR S' . Quindi ricordando la (6.74) abbiamo

$$\boldsymbol{\omega} = \begin{pmatrix} \dot{\varphi} \\ -\dot{\theta} \sin \varphi \\ \dot{\theta} \cos \varphi \end{pmatrix},$$

dove, lo ricordiamo, θ e l'angolo di rotolamento φ sono legati dalla (4.20). Si applica adesso la (7.68), con \mathbb{I}_Q data da

$$\mathbb{I}_Q = \begin{pmatrix} \frac{MR^2}{2} & 0 & 0 \\ 0 & \frac{MR^2}{4} & 0 \\ 0 & 0 & \frac{MR^2}{4} \end{pmatrix}.$$

ottenendo così

$$\frac{1}{2} \boldsymbol{\omega} \cdot \mathbb{I}_Q \boldsymbol{\omega} = \frac{MR^2}{4} \dot{\varphi}^2 + \frac{MR^2}{8} \dot{\theta}^2.$$

La (7.69)₁ diventa quindi

$$T = \frac{5}{8}MR^2\dot{\theta}^2 + \frac{MR^2}{4}\dot{\varphi}^2 \stackrel{(4.20)}{=} \frac{MR^2}{4} \left(\frac{5}{2} + \frac{R^2}{r^2} \right) \dot{\theta}^2.$$

Esempio 7.6.6 Il sistema di figura 7.16 è costituito da un'asta AB, di lunghezza l e massa m , i cui estremi sono vincolati a scorrere su una circonferenza di raggio $l/2$ e massa m . La circonferenza può ruotare attorno all'asse x . I gradi di libertà del sistema rigido sono due: l'angolo φ di rotazione della circonferenza e l'angolo θ di rotazione dell'asta. Rispetto al SdR $\{O, x, y, z\}$ solidale con il disco (ma non con

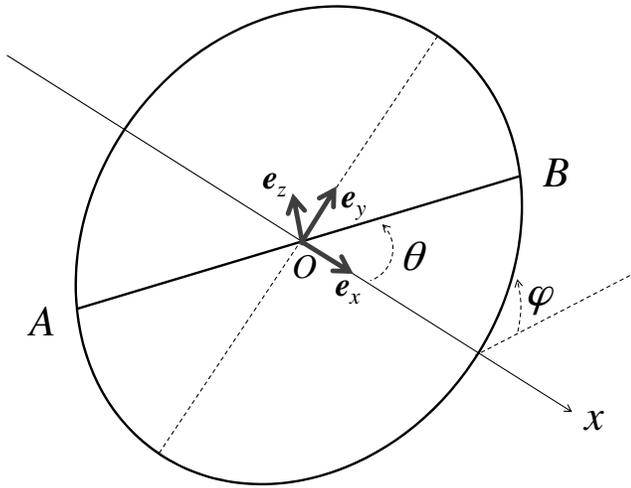


Figura 7.16: Asta vincolata ad una circonferenza ruotante.

l'asta), abbiamo

$$\omega = \dot{\varphi}e_x + \dot{\theta}e_z.$$

In tale SdR dobbiamo esprimere la matrice d'inerzia, che sarà

$$\mathbb{I}_O = \mathbb{I}_O^{\text{anello}} + \mathbb{I}_O^{\text{asta}}.$$

Come già osservato, il SdR $\{O, x, y, z\}$ non è solidale con l'asta. Consideriamo quindi un SdR S , sempre centrato in O , solidale con l'asta, in cui l'asse X coincide con l'asta AB , l'asse Y giace sul piano dell'anello ed è ortogonale ad X , e $Z \equiv z$. In tale SdR abbiamo

$$\mathbb{I}_{O,S}^{\text{asta}} = \frac{ml^2}{12} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Per passare ad \mathbb{I}_O^{asta} , bisogna applicare la formula (7.17), dove \mathbb{B} è la matrice di rotazione di un angolo θ attorno all'asse z , cioè

$$\begin{aligned}\mathbb{I}_O^{asta} &= \frac{ml^2}{12} \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \frac{ml^2}{12} \begin{pmatrix} \sin^2 \theta & -\cos \theta \sin \theta & 0 \\ -\cos \theta \sin \theta & \cos^2 \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.\end{aligned}$$

Come più volte osservato, gli elementi della matrice d'inerzia in un SdR non solidale dipendono, in generale, dal tempo (si ricordi infatti che $\theta = \theta(t)$). Dunque, la matrice d'inerzia della circonferenza e dell'asta è

$$\begin{aligned}\mathbb{I}_O &= \frac{ml^2}{8} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \frac{ml^2}{12} \begin{pmatrix} \sin^2 \theta & -\cos \theta \sin \theta & 0 \\ -\cos \theta \sin \theta & \cos^2 \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \frac{ml^2}{12} \begin{pmatrix} \frac{3}{2} + \sin^2 \theta & -\cos \theta \sin \theta & 0 \\ -\cos \theta \sin \theta & \frac{3}{2} + \cos^2 \theta & 0 \\ 0 & 0 & 4 \end{pmatrix}.\end{aligned}$$

Applicando quindi la (7.69)₂ otteniamo

$$\begin{aligned}T &= \frac{ml^2}{24} \begin{pmatrix} \dot{\varphi} \\ 0 \\ \dot{\theta} \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{2} + \sin^2 \theta & -\cos \theta \sin \theta & 0 \\ -\cos \theta \sin \theta & \frac{3}{2} + \cos^2 \theta & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} \dot{\varphi} \\ 0 \\ \dot{\theta} \end{pmatrix} \\ &= \frac{ml^2}{24} \left[4\dot{\theta}^2 + \dot{\varphi}^2 \left(\sin^2 \theta + \frac{3}{2} \right) \right].\end{aligned}$$

7.7 Le precessioni per inerzia

Un caso importante di moto rigido è quello che va sotto il nome di precessione per inerzia. Ad esso si riducono diversi moti, apparentemente diversi. Il primo caso è quello di un corpo che si muove mantenendo un punto fisso O ed è soggetto a vincoli lisci e a forze esterne che hanno momento risultante $\tau^{ext}(O)$ nullo rispetto al punto O . In questo caso si vede immediatamente che la seconda equazione cardinale è sufficiente per determinare il moto (essa è infatti equivalente all'equazione simbolica della dinamica), mentre la prima equazione cardinale serve solo per determinare la risultante delle forze vincolari.

Il secondo caso è il moto di un corpo rigido "libero" (cioè non soggetto a vincoli oltre a quelli di rigidità) e a cui non siano applicate forze esterne. In questo caso il centro di massa si muove di moto rettilineo uniforme, e il moto "essenziale" è quello del corpo rispetto al riferimento inerziale con origine nel centro di massa.

Infine a questo tipo di moto si può ridurre il moto di un corpo rigido quando sia soggetto solo alla forza di gravità (e quando si possa assumere che questa sia uniforme, ovvero uguale a $-m_k \mathbf{g}$). Infatti in questo caso la prima equazione cardinale determina il moto del centro di massa, mentre il moto del corpo relativo al sistema di riferimento che trasla con la velocità del centro di massa è determinato dalla seconda equazione cardinale, dove ora le forze esterne hanno momento nullo rispetto al centro di massa. E' interessante notare che ora, al contrario del caso precedente, questo il riferimento non è più inerziale.

7.7.1 Le equazioni di Eulero

Le equazioni fondamentali per la descrizione delle precessioni per inerzia sono le **Equazioni di Eulero**. In realtà non si tratta di nuove equazioni; le equazioni di Eulero sono le equazioni scalari ottenute dalla *seconda equazione cardinale*, scritta nella forma (7.54), scomposta nel sistema di riferimento principale di inerzia (quindi un sistema solidale) centrato nel polo della precessione (il punto fisso). Se O è il punto fisso, la matrice d'inerzia ha la forma

$$\mathbb{I}_O = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix},$$

le equazioni di Eulero per precessioni qualsiasi, non solo per quelle per inerzia, hanno la forma

$$\begin{aligned} I_1 \dot{\omega}_1 &= (I_2 - I_3) \omega_2 \omega_3 + \tau_1^{ext}, \\ I_2 \dot{\omega}_2 &= (I_3 - I_1) \omega_3 \omega_1 + \tau_2^{ext}, \\ I_3 \dot{\omega}_3 &= (I_1 - I_2) \omega_1 \omega_2 + \tau_3^{ext}, \end{aligned} \quad (7.70)$$

dove abbiamo indicato con ω_i le componenti di $\boldsymbol{\omega}$ in questo sistema di riferimento. Le componenti τ_i^{ext} del momento delle forze estere $\boldsymbol{\tau}^{ext}(O)$, rispetto ai tre assi solidali del riferimento principale, dipendono in genere dall'orientazione del corpo rispetto al sistema di riferimento che giudica il moto, e quindi sono funzioni degli angoli di Eulero. Quindi nel caso generale le (7.70) formano un sistema di equazioni differenziali del secondo ordine negli angoli di Eulero (per "esplicitare" questo sistema si dovrebbero sostituire le componenti di $\boldsymbol{\omega}$ tramite gli angoli di Eulero e le loro derivate).

Nel caso di una precessione per inerzia, tuttavia, abbiamo $\tau_i^{ext} = 0$, $i = 1, 2, 3$, e le equazioni di Eulero diventano un sistema di equazioni del primo ordine per le componenti di $\boldsymbol{\omega}$. E' bene notare subito che la risoluzione di queste equazioni **non** fornisce direttamente il moto del sistema ma l'evoluzione del vettore velocità angolare nel sistema di riferimento solidale; ovvero come si muove il vettore $\boldsymbol{\omega}$ nel sistema di riferimento di cui stavamo cercando il moto!

In effetti la soluzione delle equazioni di Eulero è solo un passo intermedio nella determinazione del moto. Una volta determinato $\boldsymbol{\omega}$ dobbiamo ricostruire la matrice \mathbb{A} di passaggio tra il sistema di riferimento fisso e quello mobile di cui $\boldsymbol{\omega}$, come espresso dalla formula (6.27) della sezione 6.3, è "la derivata", cioè

$$\frac{d\mathbb{A}}{dt} \mathbb{A}^T = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

7.7.2 Risoluzione dell'equazione di Eulero nel caso di precessioni per inerzia

Le equazioni di Eulero per le precessioni per inerzia sono risolubili “per quadrature”, cioè se ne può trovare la soluzione con un numero finito di integrazioni (e inversioni di funzioni) come avveniva per l'equazione scalare $\ddot{x} = f(x)$, analizzata nella sezione 2.6 capitolo 2. Questo è dovuto all'esistenza di due integrali primi, che permettono di ridurre il sistema a una sola equazione scalare del primo ordine. Riscriviamo il sistema (7.70) ponendo $\tau^{ext}(O) = 0$,

$$\begin{aligned} I_1 \dot{\omega}_1 &= (I_2 - I_3) \omega_2 \omega_3, \\ I_2 \dot{\omega}_2 &= (I_3 - I_2) \omega_3 \omega_1, \\ I_3 \dot{\omega}_3 &= (I_1 - I_2) \omega_1 \omega_2. \end{aligned} \quad (7.71)$$

Il primo integrale è la conservazione dell'energia: poiché non ci sono forze che compiono lavoro, l'energia si riduce alla sola energia cinetica

$$T = \frac{1}{2} (I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2). \quad (7.72)$$

La costanza di T si può ricavare direttamente dalle (7.71) moltiplicando la prima per $\dot{\omega}_1$, la seconda per $\dot{\omega}_2$ e la terza per $\dot{\omega}_3$ e sommando membro a membro: il risultato è la derivata di T eguagliata a zero.

L'altro integrale è la conservazione del momento della quantità di moto che segue dalla seconda equazione cardinale, dal momento che $\tau^{ext}(O) = 0$. In particolare resta costante il modulo quadro del momento

$$L^2 = I_1^2 \omega_1^2 + I_2^2 \omega_2^2 + I_3^2 \omega_3^2. \quad (7.73)$$

Si noti che questi due integrali si riducono, essenzialmente, alla stessa equazione nel caso in cui i tre momenti di inerzia siano uguali: in questo caso ogni asse è principale di inerzia e il moto è una rotazione uniforme attorno all'asse individuato dalla condizione iniziale per ω .

I due integrali primi ci dicono che ad ogni istante il vettore ω si trova sulla curva determinata dall'intersezione dei due ellissoidi determinati da (7.72) e (7.73): questa curva prende il nome di **poloide** (v. figura 7.17).

Dividendo la (7.72) per $2T$

$$\frac{I_1}{2T} \omega_1^2 + \frac{I_2}{2T} \omega_2^2 + \frac{I_3}{2T} \omega_3^2 = 1 \quad (7.74)$$

e la (7.73) per L^2

$$\frac{I_1^2}{L^2} \omega_1^2 + \frac{I_2^2}{L^2} \omega_2^2 + \frac{I_3^2}{L^2} \omega_3^2 = 1, \quad (7.75)$$

e sottraendo membro a membro, otteniamo

$$I_1 \left(I_1 - \frac{L^2}{2T} \right) \omega_1^2 + I_2 \left(I_2 - \frac{L^2}{2T} \right) \omega_2^2 + I_3 \left(I_3 - \frac{L^2}{2T} \right) \omega_3^2 = 0. \quad (7.76)$$

La (7.76) è l'equazione di un cono (detto **cono di Poinsot**¹⁸) passante passante per O (v. figura 7.18), che è la superficie spazzata dall'asse di moto al variare del tempo, ovvero la rigata mobile di questo particolare moto rigido¹⁹.

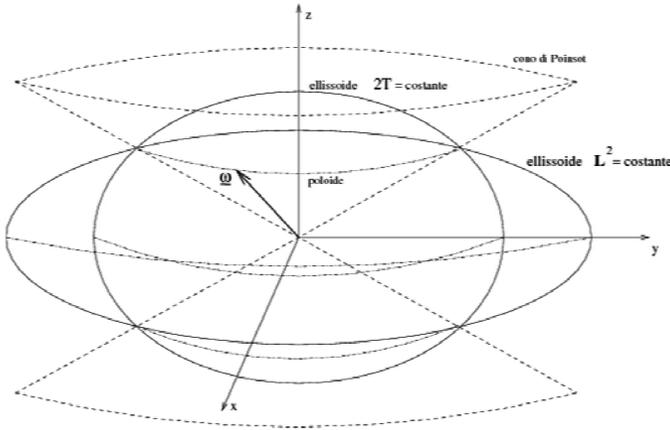


Figura 7.17: Integrali primi del moto per inerzia. Poloide.

Casi degeneri: rotazioni

Vediamo prima cosa succede in alcuni casi particolari. Supponiamo di aver scelto gli assi in modo che sia $I_1 < I_2 < I_3$. Se $\frac{L^2}{2T} = I_1$, il cono rappresentato dalla (7.76) degenera nella retta $\omega_2 = \omega_3 = 0$ (il coefficiente di ω_1 è nullo e gli altri due sono negativi) che è l'asse maggiore dell'ellissoide $2T = I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2$. Questo implica che la velocità angolare ω ha sempre la direzione dell'asse ω_1 . Per valori di $\frac{L^2}{2T}$ leggermente maggiori di I_1 , la poloide è una piccola curva chiusa attorno all'asse maggiore: ne segue che le rotazioni attorno all'asse maggiore sono stabili.

Analogamente se $\frac{L^2}{2T} = I_3$, ω ha sempre la direzione dell'asse ω_3 (l'asse minore dell'ellissoide). Il moto risultante è una rotazione del sistema attorno a questo asse. Anche in questo caso le poloidi corrispondenti a valori leggermente inferiori di $\frac{L^2}{2T}$, sono curve chiuse attorno all'asse minore, e la rotazione è ancora stabile.

¹⁸Louis Poinsot (Clermont, 1777 – Parigi, 1859) è stato un matematico e fisico francese noto per i suoi contributi alla meccanica dei sistemi rigidi.

¹⁹Ovviamente (7.76) è l'equazione di un cono se e solo se $\min\{I_1, I_2, I_3\} < \frac{L^2}{2T} < \max\{I_1, I_2, I_3\}$. Questa condizione è garantita (con i segni non stretti nei casi degeneri) dall'esistenza stessa del moto, ovvero la condizione iniziale per ω fa sì che $\frac{L^2}{2T}$ soddisfi le due disequaglianze.

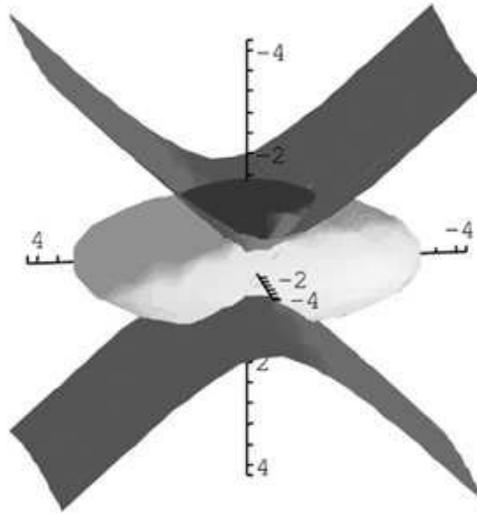


Figura 7.18: Integrali primi: l'ellissoide scuro (si vede solo una calotta) è la superficie dell'energia cinetica; l'ellissoide chiaro è la superficie del modulo quadro del momento; la superficie aperta è il cono di Poinsot solidale.

Nel caso in cui si abbia $\frac{L^2}{2T} = I_2$, il cono degenera nella coppia di piani²⁰

$$I_1 \left(I_1 - \frac{L^2}{2T} \right) \omega_1^2 + I_2 \left(I_2 - \frac{L^2}{2T} \right) \omega_3^2 = 0.$$

Ovviamente vi è anche la soluzione $\omega_1 = \omega_3 = 0$, e $\omega_2 \neq 0$, che corrisponde alla condizione iniziale $\boldsymbol{\omega}$ diretto lungo l'asse intermedio: moto di rotazione attorno all'asse intermedio. Questa volta però il moto di rotazione è instabile.

Il giroscopio

Un altro caso in cui è facile determinare il moto per inerzia è quello che corrisponde ad un corpo a struttura giroscopica, ovvero quando due dei momenti di inerzia risultano uguali (ma non tutti e tre). Volendo caratterizzare un corpo a struttura giroscopica in termini di ellissoide d'inerzia, abbiamo che questo è un *ellissoide di rotazione*. L'asse di simmetria dell'ellissoide è detto **asse giroscopico**, e su di esso giace il centro di massa.

In questo caso i due ellissoidi (7.74) e (7.75) sono "rotondi" (cioè a sezione circolare). Le poloidi sono quindi delle circonferenze e il cono di Poinsot è un cono circolare retto.

²⁰Si noti che i due coefficienti hanno segni opposti.

Senza perdere in generalità, possiamo supporre $I_1 = I_2 = I$, mentre l'asse z ha momento d'inerzia I_3 . Le equazioni di Eulero si riducono a

$$\begin{aligned} I\dot{\omega}_1 &= (I - I_3)\omega_2\omega_3, \\ I\dot{\omega}_2 &= (I_3 - I)\omega_3\omega_1, \\ I_3\dot{\omega}_3 &= 0. \end{aligned}$$

Abbiamo quindi immediatamente $\omega_3(t) = \omega_{30} = \text{costante}$, e

$$\begin{cases} \dot{\omega}_1 = \Omega\omega_2, \\ \dot{\omega}_2 = -\Omega\omega_1, \end{cases} \Rightarrow \begin{cases} \ddot{\omega}_1 = -\Omega^2\omega_1, \\ \ddot{\omega}_2 = -\Omega^2\omega_2, \end{cases} \quad (7.77)$$

dove

$$\Omega = \frac{I - I_3}{I}\omega_{30}. \quad (7.78)$$

Le (7.77) sono immediatamente risolubili: sono le equazioni del moto armonico e ci dicono che ω ruota con velocità di modulo costante attorno all'asse z . In particolare, se supponiamo che all'istante iniziale $\omega_1|_{t=0} = A$, e $\omega_2|_{t=0} = 0$, la soluzione è

$$\begin{cases} \omega_1(t) = A \cos(\Omega t), \\ \omega_2(t) = A \sin(\Omega t), \\ \omega_3(t) = \omega_{30}. \end{cases}$$

Le componenti del momento angolare (ovviamente nel SdR solidale) sono

$$\mathbf{L}(O) = \mathbb{I}_O\boldsymbol{\omega} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} A \cos(\Omega t) \\ A \sin(\Omega t) \\ \omega_{30} \end{pmatrix} = \begin{pmatrix} AI \cos(\Omega t) \\ AI \sin(\Omega t) \\ I_3\omega_{30} \end{pmatrix}.$$

Quindi, rispetto al SdR solidale, le prime due componenti di $\mathbf{L}(O)$, cioè L_1 e L_2 variano nel tempo, mentre la terza componente $L_3 = I_3\omega_{30}$, rimane costante. Ma allora, se osservato dal SdR solidale il vettore $\mathbf{L}(O)$ ruota attorno all'asse z , con velocità angolare Ω . Tale moto viene appunto detto **precessione**. Quindi, se ξ è l'angolo fra l'asse z e il vettore $\mathbf{L}(O)$, esso rimane invariato e, siccome $L_3 = |\mathbf{L}(O)| \cos \xi$, si ha

$$\cos \xi = \frac{I_3\omega_{30}}{|\mathbf{L}(O)|}.$$

Ora poniamoci nel SdR fisso $\{X, Y, Z\}$ in cui l'asse Z coincide con il vettore $\mathbf{L}(O)$, che, ovviamente, è costante. Da questa prospettiva è l'asse z del SdR solidale che forma un angolo (costante) ξ con l'asse Z . In tal caso è l'asse z che precessa attorno all'asse Z (come mostrato dalla figura 7.19) con velocità angolare Ω , data da (7.78), detta anche **velocità di precessione**. In particolare, Ω può anche essere espresso in funzione dell'angolo ξ , infatti

$$\Omega = \frac{I - I_3}{I}\omega_{30} = \frac{I - I_3}{I} \frac{|\mathbf{L}(O)|}{I_3} \cos \xi.$$

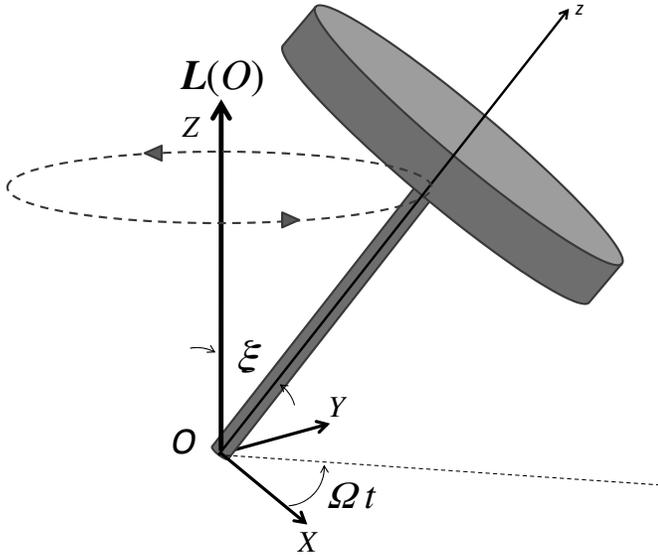


Figura 7.19: SdR fisso $\{X, Y, Z\}$, in cui l'asse Z è allineato con $L(O)$. L'asse z del SdR solido precede attorno a $L(O)$, ovvero attorno all'asse Z .

Caso generico

Torniamo al caso generico in cui sia $I_1 < I_2 < I_3$, e la costante del moto $\frac{L^2}{2T}$ è diversa dai momenti di inerzia. Possiamo utilizzare le due relazioni (7.72) e (7.73) per scrivere un'equazione che coinvolge solo una delle ω_i , per esempio la seconda componente ω_2 . Infatti si ha

$$\omega_1^2 = a_1^2(\nu_1^2 - \omega_2^2), \quad \omega_3^2 = a_3^2(\nu_3^2 - \omega_2^2), \quad (7.79)$$

con

$$a_1^2 = \frac{I_2 I_3 - I_2}{I_1 I_3 - I_1}, \quad a_3^2 = \frac{I_2 I_1 - I_2}{I_3 I_1 - I_3},$$

e

$$\nu_1^2 = \frac{2T}{I_2} \left(\frac{I_3 - \frac{L^2}{2T}}{I_3 - I_2} \right), \quad \nu_3^2 = \frac{2T}{I_2} \left(\frac{I_1 - \frac{L^2}{2T}}{I_1 - I_2} \right).$$

Sostituendo nella seconda equazione di Eulero il prodotto $\omega_1 \omega_3$ tramite la (7.79), otteniamo un'equazione a variabili separabili per la sola ω_2 (l'equazione sarà affetta da una doppia determinazione di segno a causa delle radici quadrate; questo è inevitabile poiché corrisponde al fatto che ad ogni valore di ω_2 corrispondono due punti sulla poloide).

7.7.3 Il moto à la Poincot

Accanto alla descrizione analitica del moto fornita dalle equazioni di Eulero, i due integrali primi T e $L(O)$ ci permettono di dare una elegante descrizione geometrica del moto di precessione per inerzia, anche detto **moto à la Poincot**. Denotiamo con $E = costante$, l'energia cinetica del corpo e consideriamo l'ellissoide dato da

$$\phi(\boldsymbol{\omega}) = \boldsymbol{\omega} \cdot \mathbb{I}_O \boldsymbol{\omega} - 2E = 0. \tag{7.80}$$

Nel SdR solidale $\{e_x, e_y, e_z\}$ l'espressione di $\phi(\boldsymbol{\omega})$ è particolarmente semplice se supponiamo che questo sia una terna principale d'inerzia,

$$\phi(\omega_1, \omega_2, \omega_3) = I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2 - 2E = 0. \tag{7.81}$$

Fissato adesso un qualsiasi $\boldsymbol{\omega}$ che soddisfa la (7.81), calcoliamo il gradiente della superficie $\phi = 0$, in corrispondenza di tale $\boldsymbol{\omega}$

$$\nabla\phi = 2[I_1\omega_1 e_x + I_2\omega_2 e_y + I_3\omega_3 e_z] = 2\mathbb{I}_O \boldsymbol{\omega}.$$

Si osserva facilmente che $\nabla\phi$ è parallelo al momento angolare $\mathbf{L}(O)$

$$\mathbf{L}(O) = \mathbb{I}_O \boldsymbol{\omega} = I_1\omega_1 e_x + I_2\omega_2 e_y + I_3\omega_3 e_z.$$

Adesso consideriamo il piano Π ortogonale a $\nabla\phi$, passante per $\boldsymbol{\omega}$. Esso sarà ovviamente tangente all'ellissoide nel punto $\boldsymbol{\omega}$. Abbiamo quindi provato che, *fissato comunque $\boldsymbol{\omega}$ che soddisfa la (7.81), il piano tangente all'ellissoide in $\boldsymbol{\omega}$ è ortogonale ad $\mathbf{L}(O)$.*

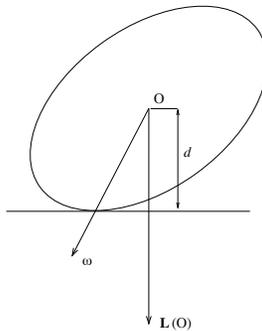


Figura 7.20: Moto à la Poincot.

Adesso facciamo gli stessi ragionamenti nel SdR fisso. Ovviamente l'equazione dell'ellissoide (7.80) avrà una forma molto più complicata e varierà nel tempo dal momento che, in generale, i coefficienti di \mathbb{I}_O dipendono dal tempo. Tuttavia $\mathbf{L}(O)$ è costante e allora anche *la giacitura del piano Π è costante* (essendo perpendicolare ad un vettore costante). Calcoliamo ora la distanza d di questo piano dal punto fisso O . Riferendoci alla figura 7.20 abbiamo

$$d = \boldsymbol{\omega} \cdot \mathbf{n},$$

dove $\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|} = \frac{\mathbf{L}(O)}{|\mathbf{L}(O)|}$, è la normale al piano Π . Ricordando adesso che $\mathbf{L}(O) = \mathbb{I}_O\boldsymbol{\omega}$,

$$d = \boldsymbol{\omega} \cdot \frac{\mathbf{L}(O)}{|\mathbf{L}(O)|} = \frac{\boldsymbol{\omega} \cdot \mathbb{I}_O\boldsymbol{\omega}}{|\mathbf{L}(O)|} = \frac{2E}{|\mathbf{L}(O)|}. \quad (7.82)$$

La (7.82) ci dice dunque che la distanza d del piano Π dall'origine è costante nel tempo. Quindi il piano tangente all'ellissoide nel punto di intersezione con $\boldsymbol{\omega}$, che è l'asse istantaneo di moto, è un piano fisso (rispetto all'osservatore fisso). A sua volta il punto solidale che istantaneamente si trova a essere il punto di tangenza, ha velocità nulla in quanto è un punto dell'asse istantaneo di moto. Ne segue che l'ellissoide rotola senza strisciare sul piano. La curva descritta dal punto sul piano Π dal punto di tangenza prende il nome di **erpoloide**. Essa è l'intersezione del piano con la rigata fissa del moto.

7.8 Lagrangiana del corpo rigido

Le equazioni di Lagrange sono utilizzabili per qualunque sistema meccanico, quali che siano i vincoli, purché olonomi. In particolare, abbiamo visto che il vincolo di rigidità è un vincolo liscio (esempio 5.2.1) per cui le forze vincolari che lo mantengono non compaiono nelle equazioni di Lagrange. In questa sezione illustreremo come si scrivono la funzione di Lagrange e le equazioni di Lagrange per un corpo rigido che:

- sia soggetto alla forza peso;
- sia soggetto a K forze \mathbf{F}_r , $r = 1, \dots, K$, applicate nei punti P_r , $r = 1, \dots, K$, del corpo;

Come nel paragrafo 7.5 si denotano con q_i , le coordinate, rispetto al SdR fisso, del centro del SdR solidale, e con φ_j , gli angoli che descrivono la rotazione. Ovviamente se il corpo è libero $i = 1, 2, 3$ e $j = 1, 2, 3$, ed il numero di gradi di libertà è 6, mentre se il corpo è soggetto a vincoli allora il numero di gradi di libertà sarà $i_{\max} + j_{\max}$, dove $0 \leq i_{\max} \leq 3$, e $0 \leq j_{\max} \leq 3$.

Per quanto riguarda la forza peso, questo è un sistema di vettori applicati negli N punti P_i del corpo, cioè

$$\mathcal{S} = \{(P_i, -m_i g \mathbf{e}_z), i = 1, \dots, N\},$$

dove $-g\mathbf{e}_z$ è l'accelerazione di gravità. Come visto nella sezione 1.5.4 il sistema \mathcal{S} , è equivalente alla forza peso risultante $-Mg\mathbf{e}_z = -\sum_i m_i g \mathbf{e}_z$, applicata in un punto qualunque della retta passante per il centro di massa P_o e parallela a \mathbf{g} . Possiamo quindi supporre che la forza peso risultante sia applicata in P_o , sì che l'energia potenziale ad essa associata è $V_{peso} = Mg z_{P_o}$, dove z_{P_o} è la quota del centro di massa. Esprimendo quest'ultima in termini dei parametri lagrangiani q_i e φ_j , cioè $z_{P_o} = \widehat{z}_{P_o}(q_i, \varphi_j)$, poniamo, come nella sezione 5.3, $\widehat{V}_{peso}(q_i, \varphi_j) = Mg \widehat{z}_{P_o}(q_i, \varphi_j)$, e la funzione di Lagrange è

$$\mathcal{L} = T - \widehat{V}_{peso} = T - Mg \widehat{z}_{P_o},$$

con T energia cinetica del corpo data dalla (7.69). Quindi il sistema di equazioni di Lagrange (5.33) diventa

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} &= \Xi_i, \quad i \leq i_{\max}, \\ \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_j} \right) - \frac{\partial \mathcal{L}}{\partial \varphi_j} &= \Xi_j, \quad j \leq j_{\max}, \end{aligned}$$

dove Ξ_i e Ξ_j sono le forze generalizzate che, ricordando la (5.32), sono date da

$$\Xi_i = \sum_{r=1}^K \mathbf{F}_r \cdot \frac{\partial \mathbf{x}_r}{\partial q_i}, \quad \text{e} \quad \Xi_j = \sum_{h=1}^K \mathbf{F}_r \cdot \frac{\partial \mathbf{x}_r}{\partial \varphi_j},$$

dove, per brevità, $\mathbf{x}_r = (P_r - O)$ denota il vettore posizione del punto P_r , dove è applicata la forza \mathbf{F}_r . Se poi le forze applicate nei K punti del corpo sono conservative (per esempio sono forze generate da molle) allora sono derivabili da un'energia potenziale $V(\mathbf{x}_1, \dots, \mathbf{x}_K)$, per cui

$$\Xi_i = -\frac{\partial \widehat{V}}{\partial q_i}, \quad \Xi_j = -\frac{\partial \widehat{V}}{\partial \varphi_j},$$

con $\widehat{V}(q_i, \varphi_j) = V(\mathbf{x}_1(q_i, \varphi_j), \dots, \mathbf{x}_K(q_i, \varphi_j))$. La funzione di Lagrange diventa $\mathcal{L} = T - \widehat{V}_{peso} - \widehat{V}$.

Esempio 7.8.1 In figura 7.21 è data un'asta AB , di lunghezza l e massa m , vincolata senza attrito nell'estremo A . L'estremo B è soggetto all'azione di una molla di lunghezza a riposo e massa trascurabili fissata nel punto $C \equiv (0, 0, l)$. Il peso è come in figura. Prendiamo come parametri lagrangiani gli angoli $\varphi \in [0, 2\pi)$ e $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$, mostrati in figura. Consideriamo poi il SdR fisso $S : \{A, x, y, z\}$ e quello solidale $S' : \{A, X, Y, Z\}$. La velocità angolare nel SdR S è

$$\begin{aligned} \boldsymbol{\omega} &= \dot{\varphi} \mathbf{e}_z + \dot{\theta} \mathbf{e}_Y = \dot{\varphi} \mathbf{e}_z + \dot{\theta} (-\sin \varphi \mathbf{e}_x + \cos \varphi \mathbf{e}_y) \\ &= -\dot{\theta} \sin \varphi \mathbf{e}_x + \dot{\theta} \cos \varphi \mathbf{e}_y + \dot{\varphi} \mathbf{e}_z, \end{aligned}$$

mentre nel SdR S' diventa

$$\begin{aligned} \boldsymbol{\omega} &= \dot{\varphi} \mathbf{e}_z + \dot{\theta} \mathbf{e}_Y = \dot{\varphi} (\sin \theta \mathbf{e}_X + \cos \theta \mathbf{e}_Z) + \dot{\theta} \mathbf{e}_Y \\ &= \dot{\varphi} \sin \theta \mathbf{e}_X + \dot{\theta} \mathbf{e}_Y + \dot{\varphi} \cos \theta \mathbf{e}_Z. \end{aligned}$$

Poniamoci nel SdR solidale S' per calcolare l'energia cinetica. Applichiamo il teorema 7.3.1 per determinare $\mathbb{I}_A^{S'}$,

$$\mathbb{I}_A^{S'} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{ml^2}{12} + m \left(\frac{l}{2}\right)^2 & 0 \\ 0 & 0 & \frac{ml^2}{12} + m \left(\frac{l}{2}\right)^2 \end{pmatrix} = \frac{ml^2}{3} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

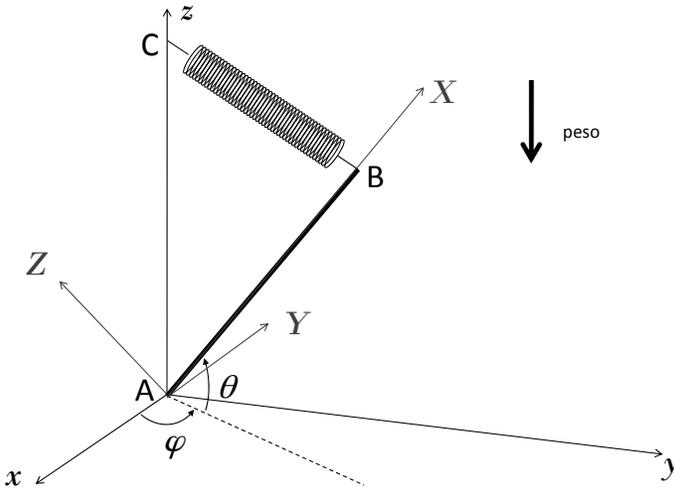


Figura 7.21: Asta vincolata nell'estremo A con l'estremo B soggetto alla forza di una molla.

Sfruttando poi la (7.69)

$$\begin{aligned} T &= \frac{ml^2}{6} \begin{pmatrix} \dot{\varphi} \sin \theta \\ \dot{\theta} \\ \dot{\varphi} \cos \theta \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \dot{\varphi} \sin \theta \\ \dot{\theta} \\ \dot{\varphi} \cos \theta \end{pmatrix} \\ &= \frac{ml^2}{6} (\dot{\theta}^2 + \dot{\varphi}^2 \cos^2 \theta). \end{aligned}$$

Notiamo che se $\theta = 0$, corrispondente all'asta orizzontale, l'energia cinetica è massima (a parità di $\dot{\varphi}$ e $\dot{\theta}$) mentre è minima quando l'asta è verticale, cioè $\theta = \pm\pi/2$.

Nel punto B è applicata una forza elastica, che possiamo derivare dall'energia potenziale $V = \frac{k}{2}(B - C)^2$,

$$\hat{V} = \frac{k}{2} \left((l \cos \theta)^2 + (l - l \sin \theta)^2 \right) = kl^2 (1 - \sin \theta).$$

L'energia potenziale della forza peso è $\hat{V} = mg \frac{l}{2} \sin \theta$, e dunque la Lagrangiana è

$$\mathcal{L} = \frac{ml^2}{6} (\dot{\theta}^2 + \dot{\varphi}^2 \cos^2 \theta) - kl^2 (1 - \sin \theta) - mg \frac{l}{2} \sin \theta.$$

La variabile φ è ciclica e ciò consente di determinare l'energia potenziale efficace (si suggerisce di farlo come esercizio).

Esempio 7.8.2 Riprendiamo l'esempio 4.1.2 e: (i) calcoliamo la funzione di Lagrange; (ii) mostriamo che il vincolo di rotolamento puro è liscio. Cominciamo proprio da questo punto, considerando l'angolo di rotolamento φ come parametro lagrangiano. Se C è il punto che, istante per istante, è a contatto con la guida rettilinea (v. figura 4.2), e \mathbf{F}_{rot} è la forza che impone il vincolo, dobbiamo calcolare

$$\mathbf{F}_{rot} \cdot \frac{\partial(C - \Omega)}{\partial\varphi},$$

dove $(C - \Omega)$ deve essere espresso nel SdR fisso $\{\Omega, x, y, z\}$.

Selezioniamo un generico punto P del disco e prendiamo il centro del disco O come punto solidale. Volendo esprimere il vettore $(P - \Omega)$ nel SdR fisso scriveremo

$$(P - \Omega) = \underbrace{(O - \Omega)}_{-R\varphi\mathbf{e}_x + R\mathbf{e}_y} + \underbrace{(P - O)}_{\mathbb{A}\mathbf{X}},$$

dove \mathbb{A} è la matrice di rotazione

$$\mathbb{A} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

e \mathbf{X} è la terna delle coordinate, rispetto al SdR solidale, del punto P . Dunque

$$\frac{\partial(P - \Omega)}{\partial\varphi} = -R\mathbf{e}_x + \frac{\partial\mathbb{A}}{\partial\varphi}\mathbf{X}. \tag{7.83}$$

Adesso consideriamo come punto P quello che, ad un determinato istante, coincide con C , cioè con il punto di contatto fra la guida ed il disco. Siccome nel SdR fisso il punto di contatto è individuato dal vettore $-R\mathbf{e}_y$, la terna \mathbf{X} è data da

$$\mathbf{X} = \mathbb{A}^T \begin{pmatrix} 0 \\ -R \\ 0 \end{pmatrix}.$$

Se quindi consideriamo il punto di contatto, la (7.83) diventa

$$\frac{\partial(C - \Omega)}{\partial\varphi} = -R\mathbf{e}_x + \frac{\partial\mathbb{A}}{\partial\varphi}\mathbb{A}^T \begin{pmatrix} 0 \\ -R \\ 0 \end{pmatrix}.$$

Siccome

$$\begin{aligned} \frac{\partial\mathbb{A}}{\partial\varphi}\mathbb{A}^T &= \begin{pmatrix} -\sin \varphi & -\cos \varphi & 0 \\ \cos \varphi & -\sin \varphi & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

abbiamo

$$\frac{\partial(C - \Omega)}{\partial\varphi} = \begin{pmatrix} -R \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -R \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Di conseguenza, qualunque sia la forza \mathbf{F}_{rot} , $\mathbf{F}_{rot} \cdot \frac{\partial(C - \Omega)}{\partial\varphi} = 0$: il **vincolo di rotolamento è liscio**.

Per quel che riguarda l'energia cinetica, sfruttando la (7.69)₁, otteniamo

$$T = \frac{m}{2} \left(\frac{d(O - \Omega)}{dt} \right)^2 + \frac{mR^2}{4} \dot{\varphi}^2 = \frac{3}{4} mR^2 \dot{\varphi}^2,$$

da cui $\mathcal{L} = \frac{3}{4} mR^2 \dot{\varphi}^2$.

7.9 Il giroscopio pesante

Un altro caso interessante di moto rigido, che può essere risolto analiticamente, è quello della “trottola”, ovvero di un corpo a struttura giroscopica²¹ soggetto alla *sola forza peso* e vincolato *senza attrito* a precedere attorno a un punto dell'asse giroscopico, come mostrato in figura 7.22. Ovviamente assumeremo che il punto in cui il corpo è vincolato sia diverso dal centro di massa (altrimenti il moto sarebbe una precessione per inerzia).

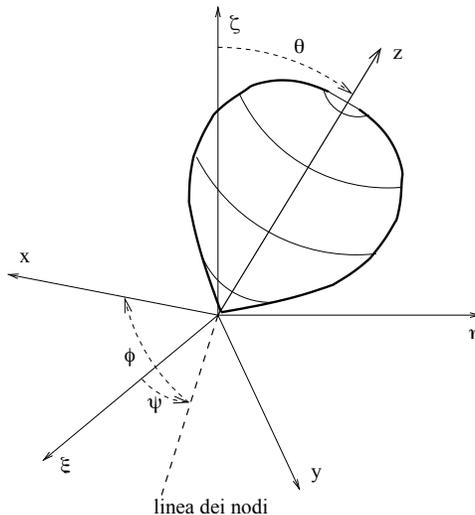


Figura 7.22: Giroscopio pesante

In questo caso l'integrazione delle equazioni di moto è fatta a partire dalle equazioni di Lagrange, utilizzando come parametri lagrangiani gli angoli di Eulero.

²¹La caratteristica di un corpo rigido a struttura giroscopica è stata descritta all'inizio del paragrafo 7.7.2.

Scegliamo il SdR solidale, $S : \{i, j, k\}$, con origine O nel polo della precessione (il punto fisso) con l'asse delle z coincidente con l'asse giroscopico. Ogni coppia di assi x e y (ortogonali) sul piano $z = 0$ definisce un sistema principale di inerzia per la simmetria di rotazione dell'ellissoide rispetto all'asse giroscopico.

Prendiamo poi un riferimento fisso, $\Sigma : \{e_1, e_2, e_3\}$, con asse delle ζ coincidente con la direzione della verticale (ascendente): i parametri lagrangiani sono gli angoli di Eulero definiti da questi due riferimenti. I due SdR sono rappresentati nella figura 7.22.

L'energia potenziale della forza peso è dato da

$$\widehat{V}(\theta, \psi, \varphi) = Mgl \cos \theta$$

dove M è la massa totale del corpo rigido e l la distanza tra P_0 e il polo O .

Per determinare l'energia cinetica esprimiamo la velocità angolare ω in funzione degli angoli di Eulero

$$\omega = \dot{\theta} \mathbf{n} + \dot{\psi} \mathbf{e}_3 - \dot{\varphi} \mathbf{k}.$$

Per determinare $\omega \cdot \sigma(O)\omega$, osserviamo che il calcolo può essere eseguito nel riferimento che ha come assi l'asse giroscopico, la linea dei nodi e l'asse perpendicolare a questi due: questo è una terna principale d'inerzia ad ogni istante (anche se mobile sia rispetto a S che a Σ) per la simmetria dell'ellissoide, e la matrice d'inerzia è del tipo

$$\mathbb{I}_O = \begin{pmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & C \end{pmatrix},$$

con $A > 0$ e $C > 0$. Indichiamo con \mathbf{t} il terzo versore di questo sistema. Basta quindi esprimere ω nelle componenti lungo \mathbf{k} , \mathbf{n} e \mathbf{t} . Siccome \mathbf{e}_3 è ortogonale alla linea dei nodi otteniamo

$$\omega = \dot{\theta} \mathbf{n} + (\dot{\psi} \cos \theta - \dot{\varphi}) \mathbf{k} + \dot{\psi} \sin \theta \mathbf{t},$$

da cui

$$2T = C(\dot{\psi} \cos \theta - \dot{\varphi})^2 + A(\dot{\theta}^2 + \dot{\psi}^2 \sin^2 \theta).$$

Le due variabili ψ e φ sono cicliche, come dovevamo aspettarci visto che il sistema è invariante per rotazioni attorno al suo asse giroscopico (variazioni di φ) e rotazioni del sistema fisso attorno alla verticale (variazioni di ψ). Ne segue che le quantità

$$p_\varphi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = -C(\dot{\psi} \cos \theta - \dot{\varphi}) = Cw_3 = -Aa, \quad (7.84)$$

e

$$p_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = C(\dot{\psi} \cos \theta - \dot{\varphi}) \cos \theta + A\dot{\psi} \sin^2 \theta = Ab, \quad (7.85)$$

sono costanti del moto. I loro valori sono stati indicati con $-Aa$ e Ab . Si noti che il primo integrale (7.84) dice che la componente della velocità angolare lungo l'asse giroscopico è costante.

Avendo supposto che il corpo sia vincolato senza attrito, abbiamo anche la conservazione dell'energia

$$E = \frac{1}{2} \underbrace{C(\dot{\psi} \cos \theta - \dot{\varphi})^2}_{\frac{p_\varphi^2}{C}} + \frac{1}{2} A(\dot{\theta}^2 + \dot{\psi}^2 \sin^2 \theta) + Mgl \cos \theta. \quad (7.86)$$

Possiamo ora utilizzare (7.84) per esprimere $\dot{\varphi}$ in funzione di $\dot{\psi}$ e θ

$$C\dot{\varphi} = -Aa + C\dot{\psi} \cos \theta, \quad (7.87)$$

che sostituita nella (7.85) dà

$$\dot{\psi} = \frac{b - a \cos \theta}{\sin^2 \theta}. \quad (7.88)$$

Sostituendo infine (7.88) nella (7.87) si ottiene

$$\dot{\varphi} = \frac{A}{C}a + \cos \theta \frac{b - a \cos \theta}{\sin^2 \theta}. \quad (7.89)$$

Le due relazioni (7.88) e (7.89) ci forniscono l'espressione delle derivate dei due angoli di Eulero in funzione dell'angolo θ . Quindi una volta ricavato θ in funzione del tempo, possiamo integrarle e ottenere ψ e φ in funzione del tempo.

Possiamo ricavare un'equazione per la sola θ partendo dalla conservazione dell'energia che riscriveremo in forma ridotta, sfruttando la conservazione di p_φ . Infatti, partendo dalla (7.86) e considerando

$$E' = E - \frac{1}{2} \frac{p_\varphi^2}{C} = \frac{1}{2} A(\dot{\theta}^2 + \dot{\psi}^2 \sin^2 \theta) + Mgl \cos \theta, \quad (7.90)$$

abbiamo che E' è ancora un integrale primo del moto. Sostituendo le (7.88) e (7.89) nella (7.90) abbiamo l'equazione cercata

$$\sin^2 \theta \dot{\theta}^2 = \sin^2 \theta (\alpha - \beta \cos \theta) - (b - a \cos \theta)^2, \quad (7.91)$$

dove

$$\alpha = \frac{2E'}{A}, \quad \beta = \frac{2Mgl}{A}.$$

L'equazione (7.91) può essere scritta in termini della variabile $u(t) = \cos \theta(t)$

$$\dot{u}^2 = (1 - u^2)(\alpha - \beta u) - (b - au)^2 = f(u), \quad (7.92)$$

che può essere risolta per separazioni di variabili esprimendo il tempo t in funzione della soluzione u

$$t - t_0 = \pm \int_{u(t_0)}^{u(t)} \frac{du}{\sqrt{(1 - u^2)(\alpha - \beta u) - (b - au)^2}}. \quad (7.93)$$

L'integrale in (7.93) è un integrale ellittico, quindi non esprimibile, con un numero finito di operazioni, in termini di "funzioni elementari". Inoltre la "soluzione" di (7.93) non consente di chiarire la fisica del problema più di quanto non si possa fare con l'analisi qualitativa della soluzione basata direttamente sull'equazione (7.92).

Dalla (7.92) abbiamo che il moto è possibile solo per quei valori di u per cui $f(u) > 0$. Inoltre, essendo $u = \cos \theta$, solo i valori di u compresi nell'intervallo $[-1, 1]$ hanno significato fisico.

Notiamo che $f(u)$ è un polinomio di terzo grado con coefficiente del termine di ordine massimo positivo. Inoltre abbiamo $f(\pm 1) \leq 0$, e il segno di uguaglianza vale solo nei casi $a = b$ e $a = -b$ che corrispondono a dati iniziali con l'asse giroscopico

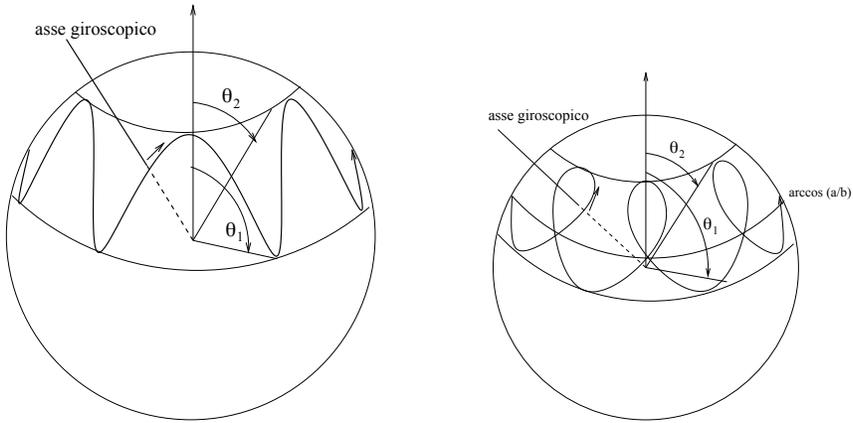


Figura 7.23: $\frac{a}{b} \notin [u_1, u_2]$, e $\frac{a}{b} \in [u_1, u_2]$.

verticale. Nel caso generico abbiamo che, denotando con u_3 il massimo delle soluzioni di $f(u) = 0$, u_3 è sicuramente maggiore di 1 e tale che $f(u)(u - u_3) > 0$ per u in un intorno di u_3 e per tutti gli $u > u_3$.

Poiché i coefficienti della funzione $f(u)$ sono determinati dalle condizioni iniziali, essi saranno tali che $f(u)$ è positiva (o meglio non negativa) in un sottoinsieme (eventualmente ridotto a un punto) contenuto in $[-1, 1]$. Quindi nella situazione generica avremo $f(u) > 0$ per $u \in (u_1, u_2) \subset (-1, 1)$, dove u_1 e u_2 sono le altre due radici del polinomio. Ne segue che il moto dell'asse, per quanto riguarda il suo angolo di nutazione θ , è un moto periodico compreso tra i due valori $\theta_2 = \arccos u_2$ e $\theta_1 = \arccos u_1$ (si noti che $\theta_2 < \theta_1$ dal momento che la funzione coseno decrescente in $(0, \pi)$).

L'angolo di precessione ψ è determinato dall'equazione (7.88). Ciò implica che ψ sarà, o non sarà, una funzione monotona del tempo, a seconda che il valore $\frac{b}{a}$ non appartenga, oppure appartenga, all'intervallo delle radici, si vedano le figure 7.23.

Un caso limite interessante è quello in cui $u_2 = \frac{a}{b}$, schematicamente rappresentato in figura 7.24. Tale caso corrisponde all'asse giroscopico "lasciato con velocità nulla" (ovviamente il giroscopio è inizialmente in rotazione attorno al proprio asse giroscopico). Questo caso può essere trattato in modo sufficientemente esauriente nell'approssimazione di "rotazione veloce" del giroscopio, ovvero nel caso, detto **trottola veloce**, in cui il numero adimensionale

$$\frac{p_\varphi^2}{2C} \gg \frac{2Mgl}{2Mgl} >> 1, \tag{7.94}$$

ovvero quando l'energia cinetica di rotazione attorno all'asse giroscopico è molto maggiore della massima variazione possibile dell'energia potenziale.

Gli effetti dovuti al momento generato dalla forza peso sono ora solo delle piccole perturbazioni del moto della trottola attorno al suo asse giroscopico. Possiamo quindi valutare l'estensione della nutazione e la velocità media della precessione dell'asse giroscopico, e le loro frequenze.

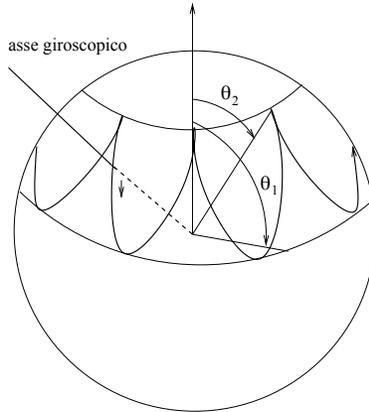


Figura 7.24: Caso: $\frac{a}{b} = u_2$

L'angolo iniziale θ_0 corrisponde alla circonferenza superiore nella figura 7.24, ovvero $u_0 = u_2$. Quindi l'estensione della nutazione dipenderà dalla posizione della seconda radice di $f(u) = 0$. Poiché abbiamo assunto anche $\psi = 0$, le costanti a e b sono legate (vedi la formula (7.88)), da

$$b = au_0.$$

Inoltre $f(u_0) = 0$, quindi avremo anche

$$\alpha = \beta u_0,$$

che significa semplicemente che il valore dell'energia ridotta E' è dato da $Mgl \cos \theta_0$, ovvero l'energia potenziale iniziale. Usando queste relazioni possiamo riscrivere la funzione $f(u)$ fattorizzando la radice u_0

$$f(u) = (u_0 - u) [\beta(1 - u^2) - a^2(u_0 - u)]. \quad (7.95)$$

La seconda radice u_1 è quindi data dalla radice minore di 1 del polinomio di secondo grado tra parentesi quadre, che riscriveremo nella forma

$$x^2 + px - q = 0,$$

con $x = u_0 - u$, e

$$p = \frac{a^2}{\beta} - 2 \cos \theta_0, \quad q = \sin^2 \theta_0.$$

Il termine $2 \cos \theta_0$ in p può essere trascurato, poiché il rapporto $\frac{a^2}{\beta}$ è dato da

$$\frac{a^2}{\beta} = \frac{C}{A} \frac{p_\varphi^2}{2Mgl},$$

che, nell'ipotesi di "trottola veloce", è molto più grande di 2 a meno che il giroscopio non sia di forma molto allungata lungo l'asse giroscopico (un sorta di trottola a forma di cilindretto). Allo stesso modo p^2 sarà molto più grande di $4q$ quindi la radice che ci interessa è data, approssimativamente, da

$$u_0 - u_1 = \frac{\beta \sin^2 \theta_0}{a^2} = \frac{A}{C} \frac{2Mgl}{Cw_3^2} \sin^2 \theta_0. \quad (7.96)$$

La (7.96) ci dice che l'ampiezza della nutazione (ovvero di quanto "scende" l'asse giroscopico) decresce in modo inversamente al quadrato della velocità di rotazione attorno all'asse.

Possiamo anche determinare in modo approssimato la frequenza di nutazione. Infatti, poiché l'ampiezza della nutazione è piccola, possiamo sostituire il termine $(1 - u^2)$ nella (7.95) con il suo valore iniziale $\sin^2 \theta_0$, ottenendo l'equazione differenziale in termini del rapporto $x = u_0/u$

$$\dot{x}^2 = x(\beta \sin^2 \theta_0 - a^2 x),$$

che è l'integrale primo di una equazione di moto armonico e ha soluzione $x = \frac{x_1}{2}(1 - \cos at)$. Ne segue che la frequenza di nutazione è data da $a = \frac{p_\varphi}{A}$, ovvero aumenta con l'aumentare della velocità di rotazione attorno all'asse giroscopico.

Infine, sostituendo questa soluzione nell'equazione per $\dot{\psi}$, (7.88) otteniamo $\dot{\psi} = \frac{\beta}{2a}(1 - \cos at)$, che ci dice che velocità di precessione non è uniforme e non cambia segno (lo sapevamo!). La quantità importante in questo caso è la velocità media della precessione sul suo periodo, che è data da $\overline{\dot{\psi}} = \frac{\beta}{2a} = \frac{Mgl}{p_\varphi}$, che ci dice che la precessione è tanto più lenta quanto la trottola è veloce.

Nota 7.9.1 *Il moto della trottola è un ottimo esempio per capire i cosiddetti **effetti giroscopici**: la **tenacia dell'asse** e la **tendenza al parallelismo**. Il primo effetto consiste nel fatto che spostare l'asse giroscopico dalla sua direzione è tanto più "difficile" (ovvero richiede tanto maggiore momento della forza) quanto più la rotazione attorno all'asse è veloce. Il secondo effetto consiste nell'osservazione di un spostamento dell'asse non nella direzione della forza applicata, ma in quella del suo momento, ovvero in direzione perpendicolare alla forza applicata (questo comportamento paradossale è solo "macroscopicamente vero" nel senso che il moto iniziale dell'asse è nella della direzione della forza ma è piccolo e viene "immediatamente" mascherato dal moto di precessione, che avviene nella direzione del momento).*

Capitolo 8

Principi variazionali

Il est démontré, disait-il, que le choses ne peuvent être autrement: car tout étant fait pour une fin, tout est nécessairement pour la meilleure fin.
Candide, Voltaire

I principi variazionali della meccanica hanno origine da un “principio”, piuttosto oscuro e impregnato di misticismo, enunciato nel 1744 da Pierre Louis Moreau de Maupertuis¹ con il nome di *principe de la moindre quantité d’action*². Esso sosteneva che il moto avviene in modo da rendere minima l’azione che la Natura deve compiere in esso: “*Lorsqu’il arrive quelque changement dans la Nature, la quantité d’action nécessaire pour ce changement est la plus petite qu’il soit possible*”³.

Questo modo “finalistico” di interpretare le leggi della fisica ha però un antecedente nel principio di Fermat⁴ per la propagazione della luce a cui Maupertuis fa esplicito riferimento.

Fermat cerca “...*l’explication des réfractions dans cet unique principe que la Nature agit toujours par les voies les plus courtes*”⁵: enuncia un principio secondo cui la luce segue il percorso lungo il quale in cammino ottico è il più breve, ovvero quello in cui la luce impiega il minor tempo possibile per andare dalla sorgente all’osservatore. Grazie a questo riesce a ricavare la legge della rifrazione di Snell⁶ della costanza del rapporto tra i seni dell’angolo di incidenza e quello di rifrazione (si veda la figura 8.1).

Questo tipo di principi, detti *variazionali*, ha però trovato sempre dei feroci oppositori. Il principio di Fermat fu subito osteggiato dai seguaci di Cartesio, con argomenti abbastanza convincenti: se si sceglie il principio di Fermat come legge fondamentale per la propagazione della luce ci si trova a dover ammettere che il raggio di luce che va

¹Maupertuis (Saint-Malo, 1698 – Basilea, 1759) è stato un matematico, fisico, filosofo, naturalista e astronomo francese.

²Principio della minima quantità d’azione

³Quando in Natura avviene un qualche cambiamento, la quantità di azione necessaria a tale cambiamento è la più piccola possibile.

⁴Pierre de Fermat (Beaumont-de-Lomagne, 1601 – Castres, 1665) fu tra i principali matematici della prima metà del XVII secolo e dette importanti contributi allo sviluppo della matematica moderna.

⁵La spiegazione delle rifrazioni in quell’unico principio per cui la Natura agisce sempre per le vie più brevi.

⁶Willebrord Snell van Royen (Leida, 1580 – Leida, 1626), matematico, astronomo e fisico olandese.

da A a B , quando giunge alla superficie di separazione dei due mezzi deve, per andare proprio in B e non in un altro qualsiasi punto equidistante dal punto O di incidenza, ricordarsi di essere partito da A per andare in B , contrariamente al principio di “localizzazione” che vuole che solo ciò che accade al momento del passaggio da un mezzo all’altro deve influenzare la variazione di direzione del moto.

In effetti il principio di Fermat si può poi ricavare attraverso il principio di Huygens della teoria ondulatoria, principio questo in accordo con la “localizzazione”.

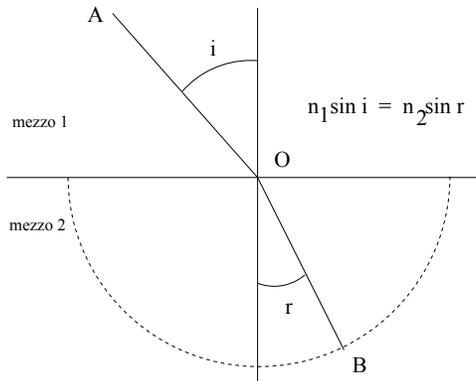


Figura 8.1: La legge di Snell o della rifrazione. n_1 è l’indice di rifrazione del mezzo 1, n_2 è l’indice di rifrazione del mezzo 2.

8.1 La brachistocrona

Esiste un classico problema della meccanica che si presenta direttamente in forma variazionale: la ricerca della **brachistocrona**, ovvero della curva di “minimo tempo”.

Il problema è quello di trovare la curva che unisce due punti A e B su cui un punto materiale, *liberamente cadendo*, ovvero muovendosi senza attrito e sotto l’influenza del suo solo peso, passi dal punto A al punto B nel **minor tempo possibile**. Si assume che il punto parta da A con velocità nulla e quindi B si trovi a una quota inferiore a quella di A .

Il problema aveva già attratto l’attenzione di Galileo, che aveva proposto come soluzione un arco di circonferenza al posto del segmento di retta che unisce A con B . Malgrado questa non sia la soluzione corretta, essa mette in evidenza una caratteristica della possibile soluzione: avere una prima parte molto ripida in modo che il punto acquisti al più presto un’elevata velocità; poi potrà percorrere anche un cammino “più lungo” purché lo faccia con maggiore velocità.

Il problema tornò all’attenzione dei matematici e fisici della fine del XVII secolo con la sfida a trovarne la soluzione lanciata da Johann Bernoulli⁷ nel 1696.

⁷Johann Bernoulli (Basilea 1667 – Basilea 1748) matematico svizzero, è stato uno dei più importanti scienziati della famiglia Bernoulli. Johann è fratello minore di Jacob.

La soluzione di Johann venne pubblicata, assieme a quella del fratello Jacob ed a una nota di Leibniz nel maggio del 1697. Una soluzione anonima apparve nel gennaio 1697 in Inghilterra (opera di Newton).

La soluzione originale di Johann Bernoulli è interessante perché stabilisce una profonda analogia tra il problema della brachistocrona e il problema ottico della rifrazione: il modo di procedere è lo stesso che si utilizza per dare conto dei miraggi e altri fenomeni ottici dovuti alla progressiva variazione di densità dell'atmosfera.

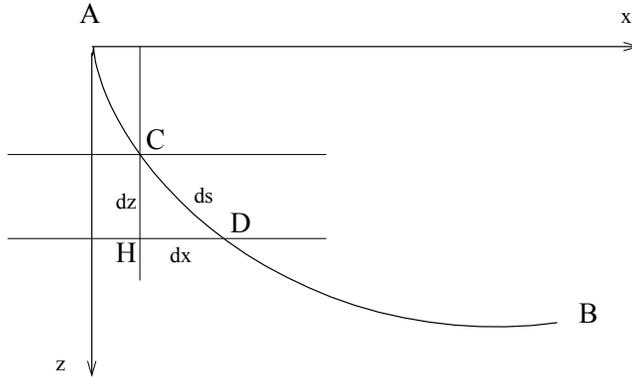


Figura 8.2: La cicloide.

In analogia col principio di Fermat, pensiamo alla curva come al percorso di un raggio luminoso che parte da *A*. Fissiamo un punto *C* sulla curva. L'angolo di deviazione della curva (ovvero della sua tangente) dalla verticale sarà l'angolo di rifrazione alla quota del punto *C*. Il suo seno è dato da

$$\frac{dx}{ds}$$

dove *ds* è la lunghezza della curva tra *C* e il punto "vicino" *D* e *dx* la sua proiezione orizzontale. In accordo con la legge di Snell quindi

$$\frac{dx}{ds} = \frac{v}{a}, \tag{8.1}$$

dove *v* è la velocità del raggio e *a* la costante di proporzionalità (ricordiamo che nella legge di Snell, come vista da Fermat, la velocità è inversamente proporzionale all'indice di rifrazione⁸).

⁸Una delle alternative poste dall'“eterna” dualità tra struttura corpuscolare e ondulatoria della luce era che la velocità della luce fosse maggiore nei mezzi più densi (Cartesio e Newton, corpuscolaristi) oppure in quelli più rarefatti (Fermat, in accordo con l'“intuizione” e soprattutto Huygens per la teoria ondulatoria). Fu proprio la misurazione sperimentale della velocità della luce in laboratorio, effettuata solo alla metà del XIX secolo, a far trionfare, provvisoriamente, la teoria ondulatoria nella fisica pre-quantistica. Al tempo di cui stiamo parlando, Newton era ancora vivo, quindi l'assunzione di Bernoulli, della proporzionalità inversa tra velocità e indice di rifrazione non era una cosa ovvia. Tant'è che nella già citata memoria, Maupertuis introduceva l'*azione*, ovvero una nuova quantità, in modo che la minimizzazione di questa quantità, al posto del tempo di percorrenza, permettesse di ricavare la legge della rifrazione da un principio di minimo à la Fermat anche nell'ipotesi newtoniana che la luce fosse più veloce nei mezzi densi.

Ma $ds^2 = dx^2 + dz^2$ quindi possiamo riscrivere la (8.1) come

$$\frac{dz}{dx} = \frac{\sqrt{a^2 - v^2}}{v}. \quad (8.2)$$

Fino a questo punto la “meccanica” non è ancora intervenuta. Essa interviene nell’esprimere la velocità in funzione della caduta del grave: $v = \sqrt{2gz}$. Inserendo questa velocità nella (8.2), otteniamo l’equazione differenziale cercata

$$\frac{dz}{dx} = \frac{\sqrt{b-z}}{\sqrt{z}}, \quad (8.3)$$

con $b = \frac{a^2}{2g}$, che Bernoulli dimostra essere l’equazione soddisfatta dalla cicloide, rappresentata in figura 8.2.

8.1.1 La trattazione moderna

La soluzione di Johann Bernoulli però non si presta a una generalizzazione a problemi variazionali più generali.

Il metodo che ora seguiamo per dedurre le equazioni differenziali da problemi variazionali è dovuto a Lagrange (e ad Eulero) e permette di trovare le equazioni per tutta una vasta classe di problemi.

Riscriviamo il problema della brachistocrona introducendo il *funzionale* “tempo” in dipendenza della curva scelta. Consideriamo un sistema di coordinate cartesiane in cui l’asse z è la verticale discendente uscente da A (coincidente con l’origine degli assi) e l’asse delle x è l’asse orizzontale. In tale piano le coordinate dei punti A e B , sono rispettivamente $(0, 0)$ e (L, H) (ovviamente escludiamo il caso in cui B non stia sulla verticale condotta da A , ma in questo caso la soluzione è banale!). Assumiamo inoltre che:

- la curva soluzione del problema giaccia nel piano (z, x) ;
- la curva sia esprimibile come il grafico di una funzione $z = z(x)$. In particolare $z(0) = 0$ e $z(L) = H$.

Queste assunzioni, benché ragionevoli, non sono essenziali. Potremmo infatti pensare a una curva espressa parametricamente come $x = x(\tau)$ e $z = z(\tau)$, o addirittura rinunciare all’ipotesi, ovvia, che la curva sia piana e aggiungere anche $y = y(\tau)$, ed ottenere dalla soluzione stessa del problema che in effetti le assunzioni sono, a posteriori, soddisfatte.

Il vettore posizione del punto materiale, espresso in funzione del tempo, è dato da

$$\mathbf{x}(t) = x(t) \mathbf{e}_x + z(x(t)) \mathbf{e}_z,$$

per cui la velocità del punto materiale è $\dot{\mathbf{x}}(t) = \dot{x}(t) [\mathbf{e}_x + z' \mathbf{e}_z]$, dove $z' = dz/dx$. Assumendo che $\dot{x}(t) > 0$, e denotando con v il modulo di $\dot{\mathbf{x}}$, abbiamo

$$v = \dot{x}(t) \sqrt{1 + (z')^2}. \quad (8.4)$$

Il tempo T che il punto impiega per andare da A a B è ovviamente dato da

$$T = \int_0^T dt. \tag{8.5}$$

Ora, se $\dot{x}(t) > 0$, possiamo esprimere t in funzione dell'ascissa x , cioè $t = t(x)$. In particolare, $t(0) = 0$, mentre $t(L) = T$. Effettuando quindi tale cambio di variabile nell'integrale (8.5) si ottiene

$$T = \int_0^L t'(x) dx = \int_0^L \frac{1}{\dot{x}} dx,$$

dove, evidentemente, dobbiamo aver cura di esprimere \dot{x} in termini di x . A tale scopo, ricordando la (8.4) e la conservazione dell'energia⁹

$$\frac{m}{2}v^2 - mgz = 0, \Rightarrow v = \sqrt{2gz},$$

abbiamo

$$\frac{v}{\sqrt{2gz}} = \dot{x} \sqrt{1 + (z')^2}, \Rightarrow \dot{x}(x) = \frac{\sqrt{2gz(x)}}{\sqrt{1 + (z'(x))^2}}.$$

Quindi il tempo impiegato dal punto materiale per andare da A a B sotto l'azione della sola forza peso è dato da

$$T[z] = \int_0^L \frac{\sqrt{(z')^2 + 1}}{\sqrt{2gz}} dx, \tag{8.6}$$

dove abbiamo messo in evidenza la dipendenza dell'integrale dalla scelta della curva.

Il problema della brachistocrona si può ora enunciare facendo riferimento alla (8.6): *Tra tutte le funzioni $z(x)$ tali che $z(0) = 0$ (partenza da A) e $z(L) = H$ (arrivo in B), determinare quella che rende minimo il valore di $T[z]$ in (8.6)*

8.2 L'equazione di Eulero-Lagrange

Generalizziamo ora il problema della brachistocrona a una funzione integranda f "qualsiasi":

Problema: *Data una funzione $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ determinare una funzione $y : [a, b] \rightarrow \mathbb{R}$ tale che:*

1. $y(a) = y_a$ e $y(b) = y_b$;
2. il valore dell'integrale

$$I[y] = \int_a^b f(y'(x), y(x), x) dx, \tag{8.7}$$

sia minimo.

⁹Si ricordi che il punto materiale parte da fermo e che l'energia potenziale iniziale è nulla.

Con y' si intende ancora la derivata della funzione y rispetto alla variabile x . Le ipotesi di regolarità da assumere sulla funzione f e sulla y saranno determinate dal calcolo che faremo, per ora le lasciamo nel vago!

La struttura del problema è: fissato un certo insieme di funzioni (le $y(x)$ che soddisfano alle condizioni agli estremi $y(a) = y_a$ e $y(b) = y_b$) e, data una “funzione” a valori in \mathbb{R} su questo insieme, determinarne il minimo.

Anziché usare il termine “funzione” per $I[y]$ si usa il termine **funzionale**, per sottolineare il fatto che I è “funzione di una funzione”.

Il problema sopra enunciato è quindi simile al problema di determinare i minimi di una funzione in più variabili. Inoltre, anche se ancora non ben specificato, l’insieme dove cercare la soluzione ha una naturale struttura affine: ogni $y(x)$ può essere ottenuta da una funzione $\bar{y}(x)$ di questo insieme più una funzione $h(x)$ (con la stessa regolarità) che soddisfa alle condizioni omogenee al bordo $h(a) = h(b) = 0$. Inoltre, l’insieme delle funzioni h è uno spazio vettoriale. Quindi ci sono molti degli ingredienti che nel calcolo differenziale permettono di “risolvere” il problema della ricerca dei minimi.

La difficoltà è che l’insieme, qualsiasi sia la richiesta di regolarità che si fa sulle y , è “molto grande” (ha dimensione infinita!). Quindi diventa assai difficile definire, per esempio, la derivata di $I[y]$ come limite del rapporto incrementale $(I[y+h] - I[y])/||h||$ (dobbiamo decidere come definire $||h||$). In verità questo programma può essere portato a termine, ma con grande dispendimento di mezzi matematici!

C’è però un modo per dare una **condizione necessaria per il minimo**, simile a quella dell’annullarsi della derivata. Supponiamo che il nostro problema di minimo (assoluto) abbia una o più soluzioni e che $y^*(x)$ sia una di queste soluzioni. Fissiamo ora una qualsiasi delle funzioni $h(x)$ nulle agli estremi e consideriamo solo le funzioni del tipo $y^*(x) + \varepsilon h(x)$ con $\varepsilon \in \mathbb{R}$. Possiamo quindi pensare all’applicazione che manda ε nel valore di $I[y^* + \varepsilon h]$ come ad una *funzione reale di una sola variabile reale* ε . Ora, se $y^*(x)$ rende minimo il valore di $I[y]$, ovvero se $I[y^*] \leq I[y]$ per ogni y , allora la funzione $\varepsilon \rightarrow I[y^* + \varepsilon h]$, avrà un minimo per $\varepsilon = 0$ (un po’ più complesso il caso di “minimi locali” cioè quando $I[y^*] \leq I[y]$ è verificata solo per funzioni y “vicine” a $y^*(x)$, perché non abbiamo definito cosa intendiamo per funzioni vicine; comunque si ha che qualsiasi sia questa definizione, $y^* + \varepsilon h$ sarà “vicina” a $y^*(x)$ purché ε sia sufficientemente piccolo).

Ne risulta che condizione necessaria affinché $y^*(x)$ sia un minimo per il funzionale $I[y]$ è che sia nulla la derivata rispetto a ε della funzione $I[y^* + \varepsilon h]$, calcolata per $\varepsilon = 0$, e questo per ogni funzione h nulla agli estremi¹⁰.

Nota 8.2.1 La derivata $\frac{d}{d\varepsilon} I[y^* + \varepsilon h]|_{\varepsilon=0}$ è detta **variazione prima del funzionale** $I[y]$ in $y = y^*$, e si indica con $\delta I[y^*]$, per cui la condizione necessaria affinché $y^*(x)$ sia un minimo per il funzionale $I[y]$ si esprime dicendo che in y^* si annulla la variazione prima di $I[y]$. La funzione minimizzante $y^*(x)$ va quindi cercata tra le soluzioni dell’equazione variazionale $\delta I[y] = 0$, ovvero

$$\delta I = \delta \int_a^b f(y'(x), y(x), x) dx = 0.$$

¹⁰Si confronti questa condizione con la condizione di annullamento di tutte le derivate direzionali in un punto di minimo di una funzione da \mathbb{R}^n in \mathbb{R}

La funzione arbitraria (purché nulla agli estremi) h è detta **variazione** della y , ed è spesso indicata con δy .

Vogliamo adesso calcolare esplicitamente la variazione prima

$$\frac{d}{d\varepsilon} \int_a^b f((y^*)'(x) + \varepsilon h'(x), y^*(x) + \varepsilon h(x), x) dx \Big|_{\varepsilon=0} .$$

Derivando sotto il segno di integrale, e calcolando il risultato per $\varepsilon = 0$, otteniamo

$$0 = \int_a^b \left(\frac{\partial f}{\partial y'}((y^*)'(x), y^*(x), x) h'(x) + \frac{\partial f}{\partial y}((y^*)'(x), y^*(x), x) h(x) \right) dx ,$$

dove $\frac{\partial f}{\partial y'}$ e $\frac{\partial f}{\partial y}$ denotano rispettivamente la derivata rispetto al primo e al secondo argomento della funzione f . Infine integriamo per parti il primo termine

$$\begin{aligned} 0 &= \int_a^b \left[-\frac{d}{dx} \left(\frac{\partial f}{\partial y'}((y^*)'(x), y^*(x), x) \right) + \frac{\partial f}{\partial y}((y^*)'(x), y^*(x), x) \right] h(x) dx \\ &\quad + \frac{\partial f}{\partial y'}((y^*)'(x), y^*(x), x) h(x) \Big|_{x=a}^{x=b} , \end{aligned}$$

e, tenendo conto che $h(a) = h(b) = 0$, si ha

$$0 = \int_a^b \left[-\frac{d}{dx} \left(\frac{\partial f}{\partial y'}((y^*)'(x), y^*(x), x) \right) + \frac{\partial f}{\partial y}((y^*)'(x), y^*(x), x) \right] h(x) dx . \tag{8.8}$$

La (8.8) deve valere per ogni funzione $h(x)$, quindi deve annullarsi il termine tra parentesi nell'integrale, ovvero la funzione y^* deve soddisfare l'equazione differenziale

$$\frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) - \frac{\partial f}{\partial y} = 0 , \tag{8.9}$$

che è nota come **equazione di Eulero-Lagrange**.

Nota 8.2.2 La (8.9) segue da (8.8) supponendo che

$$\phi(x) = -\frac{d}{dx} \left(\frac{\partial f}{\partial y'}((y^*)'(x), y^*(x), x) \right) + \frac{\partial f}{\partial y}((y^*)'(x), y^*(x), x)$$

sia una funzione continua. Se infatti esistesse un qualche $x_o \in (a, b)$, per cui $\phi(x_o) \neq 0$ (per esempio $\phi(x_o) > 0$), allora per la continuità avremo che $\phi(x) > 0$ in un intervallo $(x_1, x_2) \subset (a, b)$, che contiene al suo interno x_o . A questo punto, se si sceglie la funzione $h(x)$ in modo che: sia continua, valga 0 in $[a, x_1] \cup [x_2, b]$, e sia positiva per $x \in (x_1, x_2)$. Ne segue che $\int_a^b \phi(x) h(x) dx > 0$ contro l'ipotesi che sia nullo per ogni h .

E' chiaro a questo punto quali siano le ipotesi di regolarità: (nel seguito dicendo "derivabile" si sottintende "con derivate continue") dobbiamo avere f derivabile due volte sia a rispetto y' che a y e una volta rispetto a x , inoltre le soluzioni della (8.9) devono

essere funzioni di classe C^2 , il che obbliga anche le variazioni h ad avere la stessa regolarità.

Si confronti la (8.9) con le equazioni di Lagrange per il moto di un sistema meccanico.

Nota 8.2.3 Utilizzando la “notazione del δ ” introdotta nell’osservazione 8.2.1, la procedura pratica per calcolare δI , ovvero la variazione del funzionale, consiste nel calcolo della differenza

$$\delta I = I[y + \delta y] - I[y],$$

dove $I[y + \delta y]$ viene approssimato sviluppando in serie di Taylor l’integrando. Quindi, in pratica,

$$\begin{aligned} \delta I &= \int_a^b f(y'(x) + \delta y'(x), y(x) + \delta y(x), x) dx - \int_a^b f(y'(x), y(x), x) dx \\ &= \int_a^b \left(f(y'(x), y(x), x) + \frac{\partial f}{\partial y'} \delta y' + \frac{\partial f}{\partial y} \delta y \right) dx - \int_a^b f(y'(x), y(x), x) dx \\ &= \int_a^b \left(\frac{\partial f}{\partial y'} \delta y' + \frac{\partial f}{\partial y} \delta y \right) dx, \end{aligned}$$

dove $\delta y' = \frac{d\delta y}{dx}$. Se poi la funzione $\delta y(x)$ è tale che $\delta y(a) = \delta y(b) = 0$, integrando per parti otteniamo

$$\delta I = \int_a^b \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right) \delta y(x) dx.$$

Esempio 8.2.1 Come esempio di calcolo delle variazioni, consideriamo il problema di minimizzare il consumo di un’automobile che deve compiere un percorso di lunghezza $L > 0$ in un tempo fissato $T > 0$. Consideriamo quindi la distanza $s(t)$ coperta al tempo t , per cui si dovrà avere (se il viaggio inizia a $t = 0$)

$$\begin{cases} s(0) = 0. \\ s(T) = L. \end{cases} \quad (8.10)$$

Supponiamo, in prima approssimazione, che il consumo sia proporzionale al lavoro fatto dal motore per opporsi alla forza di attrito viscoso dell’aria. Essendo tale forza proporzionale alla velocità \dot{s} (si veda l’esempio 2.5.1 del capitolo 2) e diretta nella direzione opposta al moto, ne risulta che il lavoro effettuato dal motore è dato da

$$\int_0^L \alpha \dot{s} ds = \int_0^T \alpha \dot{s}^2(t) dt,$$

dove $\alpha > 0$ è un opportuno fattore di proporzionalità. Più in generale, potremo supporre che α dipenda da s (ad esempio, poiché l’efficienza del motore diminuisce nei tratti in salita, si avrà che $\alpha(s)$ è a sua volta proporzionale alla pendenza della strada, ovvero alla derivata della quota raggiunta alla distanza s). Il funzionale consumo

dell'automobile sarà pertanto¹¹

$$C[s] = \int_0^T \alpha(s(t)) \dot{s}^2(t) dt,$$

dove la funzione incognita $s : [0, T] \rightarrow \mathbb{R}$, è soggetta alle condizioni (8.10). Si tratta quindi di un funzionale del tipo (8.7) con

$$f(\dot{s}, s, t) = \alpha(s) \dot{s}^2.$$

Condizione necessaria affinché $s(t)$ sia un minimo del funzionale $C[s]$ è che $s(t)$ soddisfi l'equazione di Eulero-Lagrange (8.9), che in questo caso è

$$2\alpha(s)\ddot{s} + \alpha'(s)\dot{s}^2 = 0, \tag{8.11}$$

dove, ovviamente, $\alpha'(s)$ indica la derivata di $\alpha(s)$ rispetto a s .

Supponiamo ora per semplicità che la strada abbia una pendenza proporzionale a s^q (corrispondente a una quota proporzionale a $s^{(q+1)}$) per una certa potenza $q \in \mathbb{R}$. L'efficienza del motore, $\alpha(s)$, è dunque proporzionale a s^q e l'equazione (8.11) diventa

$$2s^q \ddot{s} + qs^{(q-1)} \dot{s}^2 = 0 \tag{8.12}$$

(indipendentemente dalla costante di proporzionalità tra α e s^q). Per risolvere questa equazione differenziale consideriamo come funzione incognita la derivata di s rispetto al tempo e dunque poniamo

$$\dot{s} = u(s).$$

Poiché $\ddot{s} = u'(s)\dot{s} = u'(s)u$ (al solito, l'apice indica la derivazione rispetto a s e il punto la derivazione rispetto a t), la (8.12) diventa

$$2s^q u(s)u'(s) + qs^{(q-1)}u^2(s) = 0.$$

Osserviamo che possiamo supporre $u(s) \neq 0$, in quanto $u(s) = 0$ corrisponde alla soluzione in cui l'auto è ferma, e le condizioni (8.10) non possono essere soddisfatte. Dividendo per $u(s)$ si ottiene quindi

$$2s^q u'(s) + qs^{(q-1)}u(s) = 0, \tag{8.13}$$

che è un'equazione a variabili separabili la cui soluzione generale è

$$u(s) = cs^{-q/2}, \tag{8.14}$$

dove c è una costante arbitraria. Ricordando che $u(s) = \dot{s}$, la (8.14) è un'equazione differenziale per $s(t)$, che può essere facilmente risolta a sua volta per separazione di variabili. In questo modo si ottiene la seguente soluzione generale dell'equazione (8.12):

$$s(t) = (c_1 t + c_0)^{2/(q+2)}, \tag{8.15}$$

¹¹Il lavoro fatto contro la forza peso sarà semplicemente dato dalla differenza fra la quota finale e quella iniziale ed è quindi una costante ininfluente.

dove c_0 e c_1 sono costanti arbitrarie. Imponendo infine le condizioni (8.10) si ottiene

$$s(t) = L \left(\frac{t}{T} \right)^{2/(q+2)}. \quad (8.16)$$

Il caso $q = 0$ corrisponde a una strada con pendenza costante (compreso naturalmente il caso della strada pianeggiante). In questo caso la (8.16) ci dice che, indipendentemente dal valore della pendenza, per minimizzare il consumo si dovrà viaggiare a velocità costante $\dot{s} = L/T$.

Se, invece, $q > 0$, dalla (8.16) si vede che l'automobile dovrà progressivamente rallentare mano a mano che la pendenza aumenta.

Esempio 8.2.2 Si considerino due punti $A \equiv (a, y_a)$, $B \equiv (b, y_b)$ sul piano. Tra tutte le curve del tipo $y(x)$ che connettono A con B , cioè tali che $y_a = y(a)$, $y_b = y(b)$, si vuole determinare quella di lunghezza minima. Il funzionale che, data la funzione $y(x)$, esprime la lunghezza della curva è

$$I[y] = \int_a^b \sqrt{1 + (y'(x))^2} dx.$$

L'equazione di Eulero-Lagrange è

$$\frac{d}{dx} \left(\frac{\partial \sqrt{1 + (y')^2}}{\partial y'} \right) = 0, \Rightarrow \frac{y' y''}{\sqrt{1 + (y')^2}} = 0,$$

che, escludendo $y' = 0$, ha come soluzione $y'' = 0$. La curva di minima lunghezza è dunque la retta

$$y(x) = y_a + \frac{y_b - y_a}{b - a} (x - a).$$

8.2.1 Un integrale primo e ritorno alla brachistocrona

Possiamo ora ricavare l'equazione differenziale per la brachistocrona direttamente dall'equazione di Eulero-Lagrange e dalla forma della funzione integranda

$$f(z', z, x) = \frac{\sqrt{(z')^2 + 1}}{\sqrt{2gz}}. \quad (8.17)$$

Si ottiene

$$\frac{d}{dx} \left[\frac{\partial}{\partial z'} \left(\frac{\sqrt{(z')^2 + 1}}{\sqrt{2gz}} \right) \right] - \frac{\partial}{\partial x} \left(\frac{\sqrt{(z')^2 + 1}}{\sqrt{2gz}} \right) = 0. \quad (8.18)$$

Così facendo però si giunge ad un'equazione differenziale del secondo ordine per $z(x)$, mentre l'equazione (8.3) è un'equazione del primo ordine con una costante "b" da determinare in funzione delle condizioni iniziali e finali. Questo ci dice che, se abbiamo operato correttamente e le (8.18) e (8.3) sono "equivalenti", la (8.3) deve essere un integrale primo dell'equazione di Eulero-Lagrange della brachistocrona, cioè

della (8.18). Vediamo che è effettivamente così poiché la funzione integranda (8.17) non dipende esplicitamente da x . Considerando un contesto generale, assumiamo che la funzione che definisce il funzionale, cioè la f che compare nella (8.7), *non dipenda esplicitamente da x* , ovvero

$$\frac{\partial f}{\partial x} = 0.$$

Definiamo la funzione¹²

$$u = y' \frac{\partial f}{\partial y'} - f, \tag{8.19}$$

e calcoliamone la derivata rispetto a x . Poiché $\frac{\partial f}{\partial x} = 0$ si ottiene

$$\frac{du}{dx} = y'' \frac{\partial f}{\partial y'} + y' \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) - \frac{\partial f}{\partial y'} y'' - \frac{\partial f}{\partial y} y'.$$

Chiaramente il primo e il terzo termine si cancellano tra loro. Ciò che resta è l'equazione di Eulero-Lagrange, moltiplicata per y' , e quindi è nullo anch'esso. Abbiamo quindi il seguente

Teorema 8.2.1 *Se la funzione integranda in (8.7) non dipende esplicitamente dalla x , la quantità*

$$y' \frac{\partial f}{\partial y'} - f$$

è un integrale primo dell'equazione di Eulero-Lagrange.

□

Torniamo ora alla brachistocrona, e scriviamo l'equazione differenziale corrispondente all'integrale primo (8.19) con $f = f(z', z)$ data dalla (8.17). Abbiamo (tralasciando il fattore $\sqrt{2g}$ che non ha effetto sulla ricerca del minimo)

$$C = \frac{z'}{\sqrt{z}} \frac{z'}{\sqrt{1+(z')^2}} - \frac{\sqrt{1+(z')^2}}{\sqrt{z}},$$

dove C è una costante, da cui si ottiene

$$C \sqrt{1+(z')^2} \sqrt{z} = -1, \tag{8.20}$$

e quindi, ponendo $\sqrt{b} = -\frac{1}{C}$,

$$z' = \frac{\sqrt{b-z}}{\sqrt{z}}, \tag{8.21}$$

che è di nuovo la (8.3). Si noti che per passare da (8.20) a (8.21) si deve estrarre una radice quadrata di $(z')^2$ e quindi abbiamo una ambiguità di segno. Ovviamente questa ambiguità non c'è all'inizio della curva: il punto, per mettersi in moto, deve scendere e quindi $z' > 0$ con la nostra scelta degli assi. Vedremo nella nota 8.2.4, e la cosa

¹²Si confronti la (8.19) con la funzione di Hamilton (3.29), scritta nel caso di un solo parametro lagrangiano q .

è abbastanza sorprendente, che se il rapporto L/H è sufficientemente grande (quindi il punto B è distante da A ma “non troppo sotto” A) la soluzione passerà per punti del piano verticale *più bassi* di B , ovvero il punto dovrà scendere sotto il suo punto di arrivo, e poi risalire, per arrivare nel minor tempo possibile.

Per trovare la soluzione della (8.20) introduciamo un parametro φ e poniamo $x = x(\varphi)$ e $z = z(\varphi)$. Possiamo quindi riscrivere la (8.20) nella forma

$$(x')^2 + (z')^2 = c \frac{(x')^2}{z}, \quad (8.22)$$

dove $c = \frac{1}{C^2}$ e l'apice indica adesso la derivata rispetto al parametro φ .

Si può verificare, a questo punto, che le funzioni

$$x(\varphi) = \frac{c}{2}(\varphi - \sin \varphi), \quad z(\varphi) = \frac{c}{2}(1 - \cos \varphi), \quad (8.23)$$

sono soluzioni di (8.22). Inoltre, le (8.23) soddisfano $x(0) = y(0) = 0$, quindi il valore $\varphi = 0$ del parametro corrisponde al punto di partenza A . Tuttavia, per affermare che (8.23) sono le soluzioni del problema dobbiamo mostrare che esiste un unico valore della costante c ed un unico valore $\varphi_B \in (0, 2\pi)$ tale che

$$\frac{c}{2} x(\varphi_B) = L, \quad \frac{c}{2} z(\varphi_B) = H. \quad (8.24)$$

Ora, se poniamo $h = \frac{H}{L}$ e, teniamo presenti le (8.23), possiamo riscriver la (8.24) in questo modo

$$\frac{c}{2} z(\varphi_B) = \underbrace{\left(\frac{H}{L}\right)}_h \underbrace{L}_{\frac{c}{2} x(\varphi_B)}, \quad \Rightarrow \quad z(\varphi_B) = h x(\varphi_B). \quad (8.25)$$

Quindi, se introduciamo

$$a(\varphi) = (1 - \cos \varphi), \quad \text{e} \quad b(\varphi) = (\varphi - \sin \varphi),$$

la (8.25) ammetterà soluzione se l'equazione

$$a(\varphi) = h b(\varphi), \quad (8.26)$$

ha soluzione. Il valore di c sarà poi determinato da $c = \frac{2L}{x(\varphi_B)}$.

Lemma 8.2.1 *Per ogni $L > 0$ e $H > 0$, ovvero per ogni $h > 0$, l'equazione (8.26) ammette almeno una soluzione.*

Dim. Si osserva che $b(0) = 0$, $b(\varphi) > 0$ in $(0, 2\pi]$, mentre $a(2\pi) = a(0) = 0$, e $a(\varphi) > 0$, per $\varphi \in (0, 2\pi)$. Per provare che (8.26) ha almeno una soluzione bisogna far vedere che, qualunque sia h , esiste un ε_h tale che $b(\varphi) < a(\varphi)$ per $\varphi \in (0, \varepsilon_h)$. Ciò segue facilmente dal confronto degli sviluppi di Taylor in $\varphi = 0$ delle funzioni $a(\varphi)$ e $b(\varphi)$. Infatti a è infinitesima del secondo ordine, mentre b è infinitesima del terzo ordine, qualunque sia h .

□

Lemma 8.2.2 *L'equazione (8.26) ammette una sola soluzione in $(0, 2\pi)$.*

Dim. Poniamo

$$r(\varphi) = \frac{a(\varphi)}{hb(\varphi)} = \frac{1}{h} \frac{1 - \cos \varphi}{\varphi - \sin \varphi}.$$

In ogni soluzione $\tilde{\varphi}$ di (8.26) si ha $r(\tilde{\varphi}) = 1$. Per dimostrare l'unicità basta far vedere che $r(\varphi)$ è una funzione strettamente decrescente in $(0, 2\pi)$, dal momento che $r(\varphi) > 0$, per $\varphi \in (0, 2\pi)$, $r(2\pi) = 0$, e $\lim_{\varphi \rightarrow 0^+} r(\varphi) = +\infty$. Se calcoliamo la derivata si ha

$$r'(\varphi) = \frac{1}{h} \frac{\varphi \sin \varphi - 2(1 - \cos \varphi)}{(\varphi - \sin \varphi)^2}. \tag{8.27}$$

In $(\pi, 2\pi)$ si ha $r'(\varphi) < 0$ in quanto somma di due funzioni negative. Inoltre $r'(\varphi) < 0$ "vicino" a $\varphi = 0$. Infatti

$$r'(\varphi) = \frac{1}{h} \frac{1 - \frac{1}{12}\varphi^4 + \frac{1}{180}\varphi^6 + \mathcal{O}(\varphi^8)}{\frac{1}{36}\varphi^6 + \mathcal{O}(\varphi^8)},$$

per cui $\lim_{\varphi \rightarrow 0^+} r'(\varphi) = -\infty$. Inoltre il numeratore di r' ,

$$n(\varphi) = \varphi \sin \varphi - 2(1 - \cos \varphi),$$

è una funzione decrescente e sempre negativa nell'intervallo $(0, \pi)$, come mostrato dalla figura 8.3. Quindi il rapporto $r(\varphi)$ è strettamente decrescente.

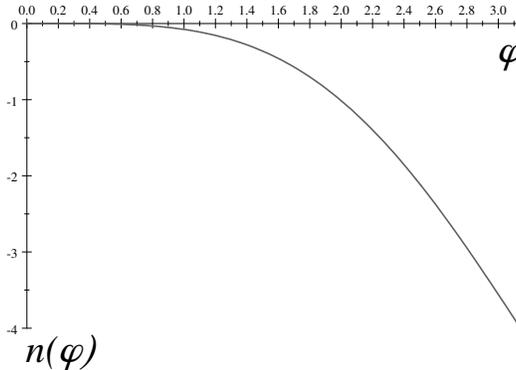


Figura 8.3: Grafico della funzione $n(\varphi) = \varphi \sin \varphi - 2(1 - \cos \varphi)$, per $\varphi \in [0, \pi]$.

□

Ne risulta il seguente

Teorema 8.2.2 *Comunque dati due punti A e B , con B posto a quota inferiore ad A , esiste unica una curva congiungente A con B sulla quale un punto, liberamente cadendo sotto l'azione della forza peso, passa da A a B nel minor tempo possibile. Questa curva è l'unico ramo di cicloide (8.23) sul piano verticale contenente A e B con un vertice in A e passante per B .*

□

Nota 8.2.4 *Se $L > \frac{\pi}{2}H$, ovvero se la distanza tra le verticali passanti per A e B è “abbastanza grande” rispetto alla caduta H , il valore di φ supera π . Ma in π la funzione $z(\varphi)$ assume il suo massimo: questo significa che la brachistocrona discende al di sotto del livello del punto di arrivo.*

Ci resta solo da “giustificare” il nome di cicloide dato alla curva (8.23). Riferendoci alla figura 4.2 dell'esempio 4.1.2, col nome di **cicloide** si indica una famiglia di curve, ognuna delle quali è descritta “cinematicamente” come la *traiettoria sul piano di un punto S solidale con un cerchio che rotola senza strisciare su una guida rettilinea*.

Dovrebbe essere abbastanza evidente da questa descrizione che si hanno due parametri positivi per descrivere la curva. Il primo è il raggio R della circonferenza, il secondo il rapporto r tra la distanza del generico punto dal centro della circonferenza e il raggio della stessa. Il primo è sostanzialmente un fattore di scala, a parità del secondo parametro le curve si ottengono l'una dall'altra per omotetia. Nel caso della brachistocrona si ha $r = 1$, ovvero il punto si trova sulla circonferenza (come mostrato in figura 4.2), mentre il valore $\frac{c}{2}$ in (8.23) è il raggio R della circonferenza. Per $r = 1$ la cicloide ha un punto singolare (cuspidè) di arresto del moto nella direzione parallela alla guida e di *inversione* del moto nella direzione ortogonale. Nella rappresentazione (8.23) questo punto corrisponde a $\varphi = 0$ e $\varphi = 2\pi$ (la curva è ovviamente periodica e, in questa rappresentazione, il periodo è proprio 2π), ovvero al punto A di partenza del grave.

E' opportuno notare che la rappresentazione (8.23) non coincide con quella data dalle (4.15) e (4.16). La differenza risiede nel fatto che le (4.15) e (4.16) sono state ottenute considerando $\varphi > 0$ per rotazioni antiorarie, mentre nelle (8.23) $\varphi > 0$ in caso di rotazioni orarie, come mostrato nella figura 8.4.

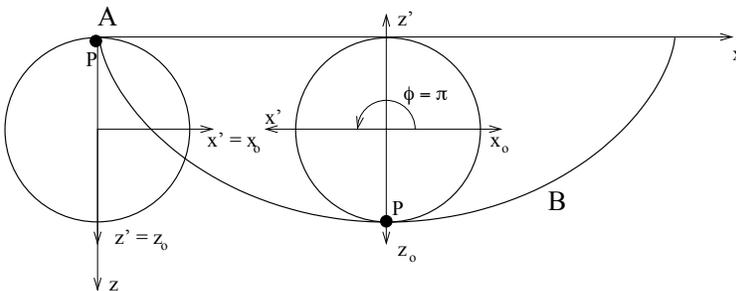


Figura 8.4: La generazione della cicloide

La cicloide, oltre a una notevole quantità di proprietà geometriche¹³, gode di un'altra interessante proprietà meccanica: è l'unica curva per cui le oscillazioni di un punto materiale che si muove su di essa, senza attrito e sotto l'azione della gravità, sono *isocrone*, ovvero il periodo delle oscillazioni è lo stesso per ogni valore (ammissibile) dell'energia¹⁴. A tal riguardo si rimanda all'esempio 5.7.4.

8.3 Funzionali dipendenti da l funzioni

Supponiamo che sia data $f : \mathbb{R}^{2l+1} \rightarrow \mathbb{R}$. Vogliamo determinare l funzioni

$$y_i(x) : [a, b] \rightarrow \mathbb{R}, \quad i = 1, \dots, l,$$

tali da rendere stazionario il funzionale

$$I[y_1, \dots, y_l] = \int_a^b f(y'_1(x), \dots, y'_l(x), y_1(x), \dots, y_l(x), x) dx, \quad (8.28)$$

che, denotando con $\tilde{\mathbf{y}}(x) = (y_1(x), \dots, y_l(x))$ e $\tilde{\mathbf{y}}'(x) = (y'_1(x), \dots, y'_l(x))$, scriveremo nella seguente forma compatta

$$I[\tilde{\mathbf{y}}] = \int_a^b f(\tilde{\mathbf{y}}'(x), \tilde{\mathbf{y}}(x), x) dx.$$

Il problema che stiamo analizzando è più generale di quello trattato nella precedente sezione 8.2, dove il funzionale dipendeva da una sola funzione e dalla sua derivata: adesso il funzionale I dipende da l funzioni e dalle loro derivate.

Definiamo adesso l'insieme dove ricercare le l -uple di funzioni $\tilde{\mathbf{y}}(x)$, ovvero il dominio del funzionale I . Sia

$$\mathcal{V} = \{ \tilde{\mathbf{y}}(x) : [a, b] \rightarrow \mathbb{R} : y_k(x) \in C^2([a, b]) \quad \forall k = 1, 2, \dots, l \}.$$

Se indichiamo con

$$\tilde{\mathbf{y}}(a) = (y_{1a}, \dots, y_{la}) \quad \text{e} \quad \tilde{\mathbf{y}}(b) = (y_{1b}, \dots, y_{lb}) \quad (8.29)$$

le l -uple costituite dai valori assunti dalle funzioni $y_k(x)$, $k = 1, \dots, l$, agli estremi a e b , il dominio del funzionale I è l'insieme

$$\mathcal{D} = \{ \tilde{\mathbf{y}}(x) \in \mathcal{V} : \tilde{\mathbf{y}}(a) = \tilde{\mathbf{y}}_a, \tilde{\mathbf{y}}(b) = \tilde{\mathbf{y}}_b \}. \quad (8.30)$$

Consideriamo adesso l'insieme

$$\mathcal{V}_0 = \{ \tilde{\mathbf{g}}(x) \in \mathcal{V} : \tilde{\mathbf{g}}(a) = \tilde{\mathbf{g}}(b) = 0 \}.$$

E' facile mostrare che \mathcal{V} e \mathcal{V}_0 hanno struttura di spazio vettoriale (in particolare \mathcal{V}_0 è sottospazio di \mathcal{V}) mentre \mathcal{D} , anche se è un sottoinsieme di \mathcal{V} , non è spazio vettoriale

¹³Si consiglia di consultare la pagina dedicata alle curve su: www-groups.dcs.st-and.ac.uk/~history/

¹⁴Ovviamente si assume che la cicloide giaccia su un piano verticale e che i suoi vertici siano alla stessa quota.

(la somma di due elementi di \mathcal{D} non dà un elemento di \mathcal{D}). Si osservi però che \mathcal{D} è uno spazio affine il cui spazio vettoriale è \mathcal{V}_0 (si veda la definizione 1.4.1). In particolare, su \mathcal{V} , e quindi anche su \mathcal{V}_0 , è possibile definire un prodotto scalare $\langle \cdot, \cdot \rangle : \mathcal{V} \rightarrow \mathbb{R}$, ponendo

$$\langle \tilde{\mathbf{f}}, \tilde{\mathbf{z}} \rangle = \int_a^b \sum_{k=1}^l f_k(x) z_k(x) dx, \quad \text{se } \tilde{\mathbf{f}} \text{ e } \tilde{\mathbf{z}} \in \mathcal{V}. \quad (8.31)$$

Seguendo la notazione dell'osservazione 8.2.1, indichiamo con $\delta\tilde{\mathbf{y}}(x) \in \mathcal{V}_0$,

$$\delta\tilde{\mathbf{y}}(x) = (\delta y_1(x), \dots, \delta y_l(x)),$$

una l -upla di funzioni tramite la quale calcolare la variazione del funzionale, cioè $\delta[I]$. Tuttavia prima di calcolare esplicitamente la variazione del funzionale (8.28), dimostriamo il seguente lemma che generalizza l'osservazione 8.2.2.

Lemma 8.3.1 *Sia $\tilde{\mathbf{F}}(x) = (F_1(x), \dots, F_l(x)) \in \mathcal{V}$. Se $\langle \tilde{\mathbf{F}}, \tilde{\mathbf{z}} \rangle = 0, \forall \tilde{\mathbf{z}} \in \mathcal{V}_0$, allora $\tilde{\mathbf{F}} \equiv 0$.*

Dim. Applicando la definizione (8.31) si ha

$$\int_a^b \sum_{k=1}^l F_k(x) z_k(x) dx = 0, \quad \forall (z_1(x), \dots, z_l(x)) \in \mathcal{V}_0. \quad (8.32)$$

Siccome $z_k(x), k = 1, \dots, l$, sono fra loro indipendenti, possiamo scegliere

$$\tilde{\mathbf{z}} = (0, \dots, 0, \underbrace{b(x)}_{\substack{i\text{-esimo} \\ \text{posto}}}, 0, 0, \dots, 0)$$

cioè $z_k(x) = z(x) \delta_{ki}, k = 1, \dots, l$. Con tale scelta la (8.32) si riduce a

$$\int_a^b F_i(x) z(x) dx = 0,$$

dove, lo ricordiamo, $z(x)$ è una funzione arbitraria. Quindi, ricordando la nota 8.2.2, si ha $F_i(x) \equiv 0$. Ma, data l'arbitrarietà con cui è stato scelto l'indice i , abbiamo $F_i(x) \equiv 0$ per ogni $i = 1, 2, \dots, l$, e quindi $\tilde{\mathbf{F}} \equiv 0$.

□

Fissiamo adesso $\delta\tilde{\mathbf{y}} \in \mathcal{V}_0$ e calcoliamo la corrispondente variazione del funzionale. Otteniamo

$$\delta I[\tilde{\mathbf{y}}] = \int_a^b f(\tilde{\mathbf{y}}'(x) + \delta\mathbf{y}'(x), \tilde{\mathbf{y}}(x) + \delta\mathbf{y}(x), x) dx, \quad (8.33)$$

Dato che le variazioni $\delta y_k(x), k = 1, \dots, l$, sono fra loro indipendenti, è legittimo considerare questa l -upla:

$$\delta\tilde{\mathbf{y}}(x) = (\delta y_1(x), \dots, \delta y_l(x)) = (\varepsilon h_1(x), \dots, \varepsilon h_l(x)),$$

dove tutte le funzioni $h_k(x)$, $k = 1, \dots, l$, ancorché arbitrarie, sono fissate. In corrispondenza di tale $\delta\tilde{\mathbf{y}}$ otteniamo la seguente funzione della variabile reale ε ,

$$F(\varepsilon) = \int_a^b f(y'_1 + \varepsilon h'_1, \dots, y'_l + \varepsilon h'_l, y_1 + \varepsilon h_1, \dots, y_l + \varepsilon h_l, x) dx.$$

Se il funzionale è stazionario allora $\left. \frac{dF(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = 0$, e dunque, tenendo conto del fatto che $h_k(a) = h_k(b) = 0$, si ha

$$0 = \int_a^b \sum_{k=1}^l \left[-\frac{d}{dx} \left(\frac{\partial f}{\partial y'_k}(y'_1, \dots, y'_k, \dots, y'_l, y_1, \dots, y_k, \dots, y_l, x) \right) + \frac{\partial f}{\partial y_k}(y'_1, \dots, y'_k, \dots, y'_l, y_1, \dots, y_k, \dots, y_l, x) \right] h_k(x) dx,$$

ovvero

$$\left\langle \frac{d}{dx} \left(\frac{\partial f}{\partial \tilde{\mathbf{y}}'} \right) - \frac{\partial f}{\partial \tilde{\mathbf{y}}}, \delta\tilde{\mathbf{y}} \right\rangle = 0, \tag{8.34}$$

dove con $\frac{d}{dx} \left(\frac{\partial f}{\partial \tilde{\mathbf{y}}'} \right) - \frac{\partial f}{\partial \tilde{\mathbf{y}}}$, intendiamo la l -upla

$$\left(\frac{d}{dx} \left(\frac{\partial f}{\partial y'_1} \right) - \frac{\partial f}{\partial y_1}, \dots, \frac{d}{dx} \left(\frac{\partial f}{\partial y'_l} \right) - \frac{\partial f}{\partial y_l} \right).$$

Ora la (8.34) vale per qualunque l -upla $\delta\tilde{\mathbf{y}}(x) \in \mathcal{V}_0$. Possiamo applicare il lemma 8.3.1 e concludere

$$\frac{d}{dx} \left(\frac{\partial f}{\partial y'_k} \right) - \frac{\partial f}{\partial y_k} = 0, \quad \forall k = 1, 2, \dots, l. \tag{8.35}$$

Il sistema di l equazioni (8.35) è dunque la generalizzazione dell'equazione di Eulero-Lagrange al caso di funzionali dipendenti da l funzioni e dalle loro derivate prime. Concludiamo dunque: *se $\tilde{\mathbf{y}}^*(x)$ rende stazionario il funzionale (8.28) allora è soluzione del sistema (8.35).*

8.3.1 Massimizzazione vincolata

Supponiamo adesso che si richieda di determinare i punti di estremo del funzionale (8.28) assumendo che le funzioni $y_i(x)$, $i = 1, \dots, l$, soddisfino la relazione

$$\Phi(y_1(x), \dots, y_l(x), x) = 0, \quad \forall x \in [a, b], \tag{8.36}$$

che, scritta in forma compatta, diventa $\Phi(\tilde{\mathbf{y}}(x), x) = 0$. Di conseguenza il dominio del funzionale I diventa

$$\widehat{\mathcal{D}} = \{ \tilde{\mathbf{y}}(x) \in \mathcal{D} : \Phi(y_1(x), \dots, y_l(x)) = 0, \forall x \in [a, b] \}.$$

dove \mathcal{D} è definito dalla (8.30).

Nota 8.3.1 Osserviamo che devono essere soddisfatte le seguenti condizioni di compatibilità

$$\underbrace{\Phi(y_{1a}, \dots, y_{la}, a)}_{\Phi(\tilde{\mathbf{y}}(a), a)} = \underbrace{\Phi(y_{1b}, \dots, y_{lb}, b)}_{\Phi(\tilde{\mathbf{y}}(b), b)} = 0.$$

Il vincolo (8.36) potrebbe essere utilizzato per esprimere una delle funzioni $y_i(x)$ in termini delle rimanenti $(l-1)$. Tuttavia, siccome questa procedura può essere estremamente complicata, si preferisce percorrere una via alternativa e, per certi versi, più semplice.

Definiamo innanzitutto qual è l'insieme dove cercare le l -uple $\delta\tilde{\mathbf{y}}(x)$. Fissata una l -upla $\tilde{\mathbf{y}}_o(x) \in \hat{\mathcal{D}}$, una generica l -upla $\tilde{\mathbf{y}}(x) \in \hat{\mathcal{D}}$ può scriversi come

$$\tilde{\mathbf{y}}(x) = \tilde{\mathbf{y}}_o(x) + \delta\tilde{\mathbf{y}}(x),$$

a patto che:

1. $\delta\tilde{\mathbf{y}}(a) = \delta\tilde{\mathbf{y}}(b) = 0$.
2. $\Phi(\tilde{\mathbf{y}}(x)) = 0$, ovvero $\Phi(\tilde{\mathbf{y}}_o(x) + \delta\tilde{\mathbf{y}}(x)) = 0$.

Le l -uple $\delta\tilde{\mathbf{y}}$ ammissibili sono dunque quegli elementi di \mathcal{V}_0 per cui

$$\Phi(\tilde{\mathbf{y}}_o(x) + \delta\tilde{\mathbf{y}}(x)) = 0, \quad \forall \tilde{\mathbf{y}}_o(x) \in \hat{\mathcal{D}}, \quad \text{e } \forall x \in [a, b]. \quad (8.37)$$

La condizione (8.37) è tuttavia "scomoda" dal punto di vista pratico. Una condizione equivalente si ottiene considerandone lo sviluppo di Taylor¹⁵

$$0 = \Phi(\tilde{\mathbf{y}}_o(x) + \delta\tilde{\mathbf{y}}(x)) = \underbrace{\Phi(\tilde{\mathbf{y}}_o(x))}_{=0} + \sum_{k=1}^l \frac{\partial\Phi}{\partial y_k} \delta y_k + \mathcal{O}(\|\delta\tilde{\mathbf{y}}\|^2).$$

Quindi, trascurando i termini del secondo ordine (che però sono influenti ai fini del calcolo delle variazioni), la (8.37) è equivalente a

$$\tilde{\mathbf{\Lambda}} \bullet \delta\tilde{\mathbf{y}} = \Lambda_1(x) \delta y_1(x) + \dots + \Lambda_l(x) \delta y_l(x) = 0, \quad \forall x \in [a, b], \quad (8.38)$$

dove¹⁶

$$\tilde{\mathbf{\Lambda}}(x) = (\Lambda_1(x), \dots, \Lambda_l(x)) = \left(\left. \frac{\partial\Phi}{\partial y_1} \right|_{\tilde{\mathbf{y}}(x)}, \dots, \left. \frac{\partial\Phi}{\partial y_l} \right|_{\tilde{\mathbf{y}}(x)} \right),$$

e dove \bullet denota, come al solito, il prodotto scalare in \mathbb{R}^l . Quindi, l'insieme delle l -uple $\delta\tilde{\mathbf{y}}(x)$ di funzioni con cui calcolare la variazione del funzionale (8.28) non è più \mathcal{V}_0 ma

$$\hat{\mathcal{V}}_0 = \left\{ \delta\tilde{\mathbf{y}}(x) \in \mathcal{V}_0 : \tilde{\mathbf{\Lambda}} \bullet \delta\tilde{\mathbf{y}}(x) = 0, \quad \forall x \in [a, b] \right\}.$$

Quest'ultimo è un sottoinsieme di \mathcal{V}_0 e questo fa sì che anche il lemma 8.3.1 debba essere modificato. Infatti le l funzioni $\delta y_k(x)$, $k = 1, \dots, l$, non sono più fra loro indipendenti (come lo erano nel caso del lemma 8.3.1) ma vincolate dalla (8.38). Vale dunque il seguente

¹⁵Ovviamente si assume che Φ sia sufficientemente regolare.

¹⁶Si assume che $\|\tilde{\mathbf{\Lambda}}(x)\| \neq 0, \forall x \in [a, b]$.

Lemma 8.3.2 Sia $\tilde{\mathbf{F}}(x) \in \widehat{\mathcal{V}}$. Se $\langle \tilde{\mathbf{F}}, \delta\tilde{\mathbf{y}} \rangle = 0, \forall \delta\tilde{\mathbf{y}} \in \widehat{\mathcal{V}}_0$, allora $\tilde{\mathbf{F}} = \lambda(x) \tilde{\mathbf{\Lambda}}$, con $\lambda(x) \in \mathbb{R}$.

Dim. Osserviamo innanzi tutto che la condizione (8.38) implica

$$\int_a^b \tilde{\mathbf{\Lambda}} \bullet \delta\tilde{\mathbf{y}}(x) dx = 0, \iff \langle \tilde{\mathbf{\Lambda}}, \delta\tilde{\mathbf{y}} \rangle = 0, \quad (8.39)$$

cioè $\delta\tilde{\mathbf{y}} \in \widehat{\mathcal{V}}_0$ e $\tilde{\mathbf{\Lambda}}$ sono *ortogonali*, nel senso del prodotto scalare $\langle \cdot, \cdot \rangle$ definito dalla (8.31). Non solo, ma la (8.38) implica anche

$$0 = \langle \gamma(x) \tilde{\mathbf{\Lambda}}, \delta\tilde{\mathbf{y}} \rangle = \int_a^b \sum_{k=1}^l \gamma(x) \Lambda_k(x) \delta y_k(x) dx, \quad (8.40)$$

per ogni funzione $\gamma(x)$ sufficientemente regolare.

Scriviamo adesso la condizione $\langle \tilde{\mathbf{F}}, \delta\tilde{\mathbf{y}} \rangle = 0$, tenendo conto della (8.40) dove $\gamma(x)$ è una generica funzione non identicamente nulla, ovvero

$$\begin{aligned} 0 &= \int_a^b \sum_{k=1}^l F_k(x) \delta y_k(x) dx \\ &= \int_a^b \sum_{k=1}^l F_k(x) \delta y_k(x) dx + \underbrace{\int_a^b \sum_{k=1}^l \gamma(x) \Lambda_k(x) \delta y_k(x) dx}_{=0} \\ &= \int_a^b \sum_{k=1}^l [F_k(x) + \gamma(x) \Lambda_k(x)] \delta y_k(x) dx. \end{aligned} \quad (8.41)$$

Come già osservato, la (8.38) implica che $(l - 1)$ componenti di $\delta\tilde{\mathbf{y}}$ sono indipendenti mentre una può esser espressa in termini delle altre¹⁷. Quindi, senza perdere di generalità, possiamo supporre che le funzioni $\delta y_2(x), \dots, \delta y_l(x)$, siano indipendenti e, tornando alla (8.41), scrivere

$$\begin{aligned} 0 &= \int_a^b [F_1(x) + \gamma(x) \Lambda_1(x)] \delta y_1(x) dx \\ &\quad + \int_a^b \sum_{k=2}^l [F_k(x) + \gamma(x) \Lambda_k(x)] \delta y_k(x) dx. \end{aligned} \quad (8.42)$$

¹⁷Infatti, se per esempio $\Lambda_1 \neq 0$, la (8.38) dà luogo a

$$\delta y_1 = -\frac{1}{\Lambda_1} \sum_{i=2}^l \Lambda_i \delta y_i.$$

Quindi le funzioni $\delta y_2(x), \dots, \delta y_l(x)$, sono indipendenti mentre la prima, $\delta y_1(x)$, viene espressa in termini delle altre.

A questo punto si sceglie la funzione $\gamma(x)$ in modo da far annullare il primo integrale (se, per esempio, $\Lambda_1 \neq 0$ si pone $\gamma(x) = -\frac{F_1(x)}{\Lambda_1(x)}$). La (8.42) si riduce quindi a

$$\int_a^b \sum_{k=2}^l [F_k(x) + \gamma(x) \Lambda_k(x)] \delta y_k(x) dx = 0,$$

dove adesso le funzioni $\delta y_k(x)$, $k = 2, \dots, l$, sono tutte indipendenti. Possiamo quindi applicare il lemma 8.3.1 e concludere che $F_k(x) + \gamma(x) \Lambda_k(x) \equiv 0$, $k = 2, \dots, l$. In definitiva, se $\langle \tilde{\mathbf{F}}, \delta \tilde{\mathbf{y}} \rangle = 0$, per qualunque funzione $\delta \tilde{\mathbf{y}} \in \tilde{\mathcal{V}}_0$, abbiamo provato che $\tilde{\mathbf{F}} = \lambda(x) \tilde{\mathbf{\Lambda}}$, dove $\lambda(x) = -\gamma(x)$ per un'opportuna funzione $\gamma(x)$.

□

Nota 8.3.2 Il lemma appena dimostrato può essere letto anche in chiave geometrica: la condizione $\tilde{\mathbf{F}} = \lambda(x) \tilde{\mathbf{\Lambda}}$ è una condizione di parallelismo fra $\tilde{\mathbf{F}}$ e $\tilde{\mathbf{\Lambda}}$.

Passiamo adesso a calcolare esplicitamente la variazione (8.33) per una generica $\delta \tilde{\mathbf{y}} \in \tilde{\mathcal{V}}_0$. Abbiamo

$$\delta I[\tilde{\mathbf{y}}] = \int_a^b \sum_{k=1}^l \left[\frac{d}{dx} \left(\frac{\partial f}{\partial y'_k} \right) - \frac{\partial f}{\partial y_k} \right] \delta y_k dx,$$

che, ponendo

$$\tilde{\mathbf{F}}(x) = \left(\frac{d}{dx} \left(\frac{\partial f}{\partial y'_1} \right) - \frac{\partial f}{\partial y_1}, \dots, \frac{d}{dx} \left(\frac{\partial f}{\partial y'_l} \right) - \frac{\partial f}{\partial y_l} \right),$$

riscriveremo anche come $\delta I[\tilde{\mathbf{y}}] = \langle \tilde{\mathbf{F}}(x), \delta \tilde{\mathbf{y}} \rangle$. Ora se il funzionale è stazionario in corrispondenza di $\tilde{\mathbf{y}}^*(x) \in \tilde{\mathcal{D}}$, avremo che $\delta I[\tilde{\mathbf{y}}^*] = 0$, per ogni $\delta \mathbf{y}(x) \in \tilde{\mathcal{V}}_0$ e quindi

$$\tilde{\mathbf{F}} = \lambda \tilde{\mathbf{\Lambda}}, \implies \frac{d}{dx} \left(\frac{\partial f}{\partial y'_k} \right) - \frac{\partial f}{\partial y_k} = \lambda(x) \frac{\partial \Phi}{\partial y_k}, \quad \forall k = 1, \dots, l. \quad (8.43)$$

Osserviamo infine che il sistema di equazioni (8.43) può anche riscriversi come

$$\frac{d}{dx} \left(\frac{\partial \mathcal{F}}{\partial y'_k} \right) - \frac{\partial \mathcal{F}}{\partial y_k} = 0, \quad \forall k = 1, \dots, l,$$

dove

$$\mathcal{F}(y'_1, \dots, y'_l, y_1, \dots, y_l, x) = f(y'_1, \dots, y'_l, y_1, \dots, y_l, x) + \lambda(x) \Phi(y_1, \dots, y_l, x).$$

Abbiamo quindi un sistema di l equazioni differenziali nelle l funzioni incognite $y_1(x)$, \dots , $y_l(x)$, con le condizioni agli estremi date dalla (8.29), cioè $y_k(a) = y_{a,k}$ e $y_k(b) = y_{b,k}$, $k = 1, \dots, l$. Per quanto riguarda la funzione $\lambda(x)$, questa sarà determinata imponendo il vincolo (8.36).

Esempio 8.3.1 *Come esempio di massimizzazione/minimizzazione vincolata consideriamo il problema della ricerca della curva di minima lunghezza che connette due punti A e B su una sfera di raggio R . Sia*

$$\mathbf{x}(\tau) = x(\tau) \mathbf{e}_x + y(\tau) \mathbf{e}_y + z(\tau) \mathbf{e}_z, \quad \tau \in [0, 1],$$

la parametrizzazione della generica curva che collega A e B (curva tratteggiata in figura 8.5).

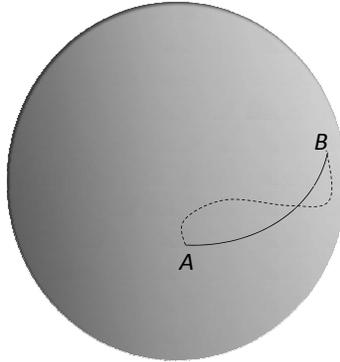


Figura 8.5: Cammino più breve tra i punti A e B con il vincolo di rimanere sulla superficie della sfera.

Il funzionale che ne esprime la lunghezza è

$$I[\mathbf{x}] = \int_0^1 \sqrt{x'^2 + y'^2 + z'^2} \, d\tau,$$

mentre il vincolo cui le funzioni $x(\tau)$, $y(\tau)$ e $z(\tau)$ sono sottoposte è

$$x^2(\tau) + y^2(\tau) + z^2(\tau) - R^2 = 0, \quad \Rightarrow \quad \|\mathbf{x}(\tau)\| = R \quad \forall \tau \in [0, 1].$$

Scrivendo le (8.43) otteniamo

$$\left\{ \begin{array}{l} \frac{d}{d\tau} \frac{x'}{\sqrt{x'^2 + y'^2 + z'^2}} = 2\lambda(\tau) x, \\ \frac{d}{d\tau} \frac{y'}{\sqrt{x'^2 + y'^2 + z'^2}} = 2\lambda(\tau) y, \\ \frac{d}{d\tau} \frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}} = 2\lambda(\tau) z, \end{array} \right. \quad (8.44)$$

ovvero, se indichiamo con

$$\mathbf{t}(\tau) = \frac{1}{\sqrt{x'^2 + y'^2 + z'^2}} (x' \mathbf{e}_x + y' \mathbf{e}_y + z' \mathbf{e}_z) = \frac{\mathbf{x}'(\tau)}{\|\mathbf{x}'(\tau)\|}, \quad (8.45)$$

il versore tangente alla curva, il sistema (8.44) può riscriversi nella seguente forma compatta

$$\frac{d\mathbf{t}}{d\tau} = 2\lambda \mathbf{x}.$$

Se moltiplichiamo vettorialmente per $\mathbf{x}(\tau)$ tale equazione si ottiene

$$\mathbf{x} \wedge \frac{d\mathbf{t}}{d\tau} = 2\lambda \mathbf{x} \wedge \mathbf{x} = 0,$$

e quindi

$$\frac{d}{dt} (\mathbf{x}(\tau) \wedge \mathbf{t}(\tau)) = \underbrace{\mathbf{x} \wedge \frac{d\mathbf{t}}{d\tau}}_{=0} + \frac{d\mathbf{x}}{d\tau} \wedge \mathbf{t} \stackrel{(8.45)}{=} \sqrt{x'^2 + y'^2 + z'^2} \mathbf{t} \wedge \mathbf{t} = 0,$$

ovvero

$$\mathbf{x}(\tau) \wedge \mathbf{t}(\tau) = \mathbf{c}, \quad (8.46)$$

dove \mathbf{c} è un vettore costante. La (8.46) implica che, per ogni τ , $\mathbf{x}(\tau)$ e $\mathbf{x}'(\tau)$ giacciono su un piano passante per l'origine, il cui versore normale è parallelo a \mathbf{c} . Ma allora anche i punti A e B posti agli estremi della curva $\mathbf{x}(\tau)$ giacciono su tale piano, che è appunto un piano equatoriale e la cui intersezione con la sfera è la circonferenza di raggio massimo, ovvero di raggio R . Quindi la curva che connette A e B avente minima lunghezza è l'arco di cerchio massimo che contiene A e B (curva continua in figura 8.5).

8.4 Il principio di Hamilton

Sia dato un sistema meccanico con funzione lagrangiana $\mathcal{L}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l, t) = \mathcal{L}(\tilde{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}, t)$, dove, come nella sezione 4.1, abbiamo posto $\tilde{\mathbf{q}}(t) = (q_1(t), \dots, q_l(t))$. Ipotizziamo inoltre che il moto sia governato dal sistema (5.36), ovvero che il **sistema sia soggetto a vincoli lisci e che le uniche forze attive siano conservative**.

Definizione 8.4.1 Siano dati due tempi t_1 e t_2 , con $t_1 < t_2$, e due configurazioni di coordinate lagrangiane

$$\tilde{\mathbf{q}}_a = (q_{1a}, \dots, q_{la}) \text{ e } \tilde{\mathbf{q}}_b = (q_{1b}, \dots, q_{lb}). \quad (8.47)$$

Diremo **classe dei moti variati sincroni** l'insieme \mathcal{Q} costituito dalle l -uple di funzioni $\tilde{\mathbf{q}}(t) = (q_1(t), \dots, q_l(t))$ tali che: (1) $q_k(t) : [t_1, t_2] \rightarrow \mathbb{R}$, e $q_k(t) \in C^2([t_1, t_2]) \forall k = 1, \dots, l$; (2) $\tilde{\mathbf{q}}(t_1) = \tilde{\mathbf{q}}_a$, $\tilde{\mathbf{q}}(t_2) = \tilde{\mathbf{q}}_b$.

Definizione 8.4.2 Diremo **moti naturali** per il sistema le funzioni $\tilde{\mathbf{q}}(t)$ **soluzioni delle equazioni di Lagrange**¹⁸ (5.36).

¹⁸Ovvero distinguiamo tra un "moto" compatibile con la struttura cinematica del sistema meccanico, ivi compresi i vincoli imposti, che è dato da una qualsiasi funzione $\tilde{\mathbf{q}}(t)$, e un moto che sia compatibile con la dinamica risultante dalle forze applicate.

Definizione 8.4.3 Si dice *azione hamiltoniana* il funzionale $\mathcal{A} : \mathcal{Q} \rightarrow \mathbb{R}$, che opera così

$$\mathcal{A}[\tilde{q}] = \int_{t_1}^{t_2} \mathcal{L}(\tilde{q}(t), \dot{\tilde{q}}(t), t) dt. \tag{8.48}$$

Vale il seguente teorema, anche noto come **Principio di Hamilton**,

Teorema 8.4.1 Se $\tilde{q}^*(t) \in \mathcal{Q}$ è moto naturale del sistema allora $\tilde{q}^*(t)$ rende stazionario il funzionale¹⁹ (8.48), ovvero la **variazione del funzionale** (8.48), cioè $\delta\mathcal{A}[\tilde{q}]$, è nulla in corrispondenza di \tilde{q}^* .

Vale anche il viceversa: se $\tilde{q}^*(t) \in \mathcal{Q}$ rende stazionaria l'azione (8.48), cioè $\delta\mathcal{A}[\tilde{q}^*] = 0$, allora $\tilde{q}^*(t)$ è moto naturale secondo la definizione 8.4.2.

Dim. La dimostrazione del teorema è una ripetizione del calcolo che abbiamo effettuato nella sezione 8.3 per ricavare il sistema (8.35) (equazioni di Eulero-Lagrange) nel caso in cui il funzionale dipenda da l funzioni e dalle loro derivate prime. Infatti, utilizzando la notazione dell'osservazione 8.2.3, abbiamo

$$\delta\mathcal{A} = \int_{t_1}^{t_2} \sum_{k=1}^l \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} \right] \delta q_k dt,$$

dove $\delta\tilde{q}(t) = (\delta q_1(t), \dots, \delta q_l(t)) \in \mathcal{Q}_0$, con

$$\mathcal{Q}_0 = \{ \delta\tilde{q}(x) : [t_1, t_2] \rightarrow \mathbb{R} : \delta q_k(x) \in C^2([t_1, t_2]) \forall k = 1, \dots, l \text{ e } \delta\tilde{q}(t_1) = \delta\tilde{q}(t_2) = 0 \}.$$

E' quindi evidente che se $\tilde{q}^*(t)$ è moto naturale, cioè se le $q_1^*(t), \dots, q_l^*(t)$, soddisfano

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} = 0, \quad k = 1, \dots, l,$$

allora $\delta\mathcal{A}[\tilde{q}^*] = 0$. Viceversa, se l'azione è stazionaria in corrispondenza di $\tilde{q}^*(t)$, cioè $\delta\mathcal{A}[\tilde{q}^*] = 0$, allora

$$\int_{t_1}^{t_2} \sum_{k=1}^l \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} \right] \delta q_k dt = 0, \quad \Leftrightarrow \quad \left\langle \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\tilde{q}}} \right) - \frac{\partial \mathcal{L}}{\partial \tilde{q}}, \delta\tilde{q} \right\rangle = 0,$$

per ogni l -upla $\delta\tilde{q}(t) \in \mathcal{Q}_0$. Ora le funzioni $\delta q_k(t)$, $k = 1, \dots, l$, sono indipendenti e quindi, applicando il lemma 8.3.1, concludiamo che le funzioni $q_1^*(t), \dots, q_l^*(t)$, costituenti la l -upla $\tilde{q}^*(t)$, risolvono le equazioni di Lagrange (5.36). Dunque, in virtù della definizione 8.4.2, $\tilde{q}^*(t)$ è moto naturale.

□

¹⁹Questo enunciato è una forma più generale del classico enunciato *Il moto naturale di un sistema meccanico rende minima l'azione stazionaria nella classe dei moti che rispettano le condizioni iniziali e finali*; così enunciato il principio è molto più "significativo" in quanto fa appello alla proprietà di minimo, come nel principio di Fermat. Questo però non è generalmente vero, l'azione è minima solo se le condizioni iniziale e finale sono "sufficientemente vicine". Nella enunciazione del teorema (8.4.1) si ottiene l'equivalenza tra il principio e le equazioni di Lagrange. Il principio è così matematicamente corretto, ma meno significativo!

Esempio 8.4.1 Consideriamo l'oscillatore armonico la cui Lagrangiana è

$$\mathcal{L} = \frac{m}{2} \sum_{i=1}^3 (\dot{q}_i^2 - \omega^2 q_i^2),$$

e, sull'intervallo $[0, T]$, consideriamo il seguente moto naturale

$$q_1^*(t) = a \sin \omega t, \quad q_2^*(t) = q_3^*(t) = 0.$$

Mostriamo che tale moto naturale rende minima l'azione se T è sufficientemente piccolo. Calcoliamo esplicitamente

$$\delta \mathcal{A} = \mathcal{A}[\tilde{\mathbf{q}}^* + \delta \tilde{\mathbf{q}}] - \mathcal{A}[\tilde{\mathbf{q}}^*],$$

e mostriamo che $\delta \mathcal{A} > 0$, per ogni $\delta \tilde{\mathbf{q}} \in \mathcal{Q}_0$. Abbiamo

$$\begin{aligned} \delta \mathcal{A} &= \frac{m}{2} \int_0^T \sum_{i=1}^3 \left[\left(\frac{d}{dt} (q_i^* + \delta q_i) \right)^2 - \omega^2 (q_i^* + \delta q_i)^2 \right] dt \\ &\quad - \frac{m}{2} \int_0^T \sum_{i=1}^3 \left[\left(\frac{dq_i^*}{dt} \right)^2 - \omega^2 q_i^{*2} \right] dt \\ &= \frac{m}{2} \int_0^T \sum_{i=1}^3 (\delta \dot{q}_i^2 - \omega^2 \delta q_i^2) dt \\ &\quad + m \int_0^T \left[\underbrace{\dot{q}_1^* \delta q_1}_{\substack{\text{si integra per parti} \\ \text{con } \delta q_1(0) = \delta q_1(T) = 0}} - \omega^2 q_1^* \delta q_1 \right] dt \\ &= \frac{m}{2} \sum_{i=1}^3 \int_0^T (\delta \dot{q}_i^2 - \omega^2 \delta q_i^2) dt - m \int_0^T \underbrace{(\dot{q}_1^* + \omega^2 q_1^*)}_{=0} dt, \end{aligned}$$

ovvero

$$\delta \mathcal{A} = \frac{m}{2} \sum_{i=1}^3 \int_0^T (\delta \dot{q}_i^2 - \omega^2 \delta q_i^2) dt.$$

Si sfrutta adesso $\delta q_i(t) = \int_0^t \delta \dot{q}_i d\tau$, e la stima²⁰ $\left| \int_0^t \delta \dot{q}_i d\tau \right| \leq \sqrt{t} \|\delta \dot{q}_i\|_{L^2(0,T)}$, per cui

$$\delta q_i(t) \leq \sqrt{t} \|\delta \dot{q}_i\|_{L^2(0,T)}, \quad \Rightarrow \quad \delta q_i^2(t) \leq t \|\delta \dot{q}_i\|_{L^2(0,T)}^2. \quad (8.49)$$

²⁰ $\|\delta \dot{q}_i\|_{L^2(0,T)} = \sqrt{\int_0^T (\delta \dot{q}_i(t))^2 dt}$.

Tornando a $\delta\mathcal{A}$, otteniamo

$$\begin{aligned} \delta\mathcal{A} &= \frac{m}{2} \left[\sum_{i=1}^3 \int_0^T \delta\dot{q}_i^2 dt - \omega^2 \sum_{i=1}^3 \int_0^T \delta q_i^2 dt \right] \\ &\stackrel{(8.49)}{\geq} \frac{m}{2} \sum_{i=1}^3 \|\delta\dot{q}_i\|_{L^2(0,T)}^2 - \omega^2 \sum_{i=1}^3 \|\delta q_i\|_{L^2(0,T)}^2 \underbrace{\int_0^T t dt}_{\frac{T^2}{2}} \\ &= \frac{m}{2} \sum_{i=1}^3 \|\delta\dot{q}_i\|_{L^2(0,T)}^2 \left(1 - \frac{\omega^2 T^2}{2} \right), \end{aligned}$$

per cui, se T è sufficientemente piccolo si che $\left(1 - \frac{\omega^2 T^2}{2} \right) > 0$, si ha $\delta\mathcal{A} > 0$.

8.4.1 Sistemi vincolati ed equazioni di Lagrange di prima specie

Riprendiamo quanto visto nell'osservazione 4.1.3 della sezione 4.1 supponendo che il sistema meccanico, con l gradi di libertà sia soggetto ad un ulteriore vincolo olonomo e liscio del tipo (4.11), che per brevità indicheremo con $\Phi(\vec{q}) = 0$. Evidentemente si può sfruttare l'equazione vincolare e ridurre così i gradi di libertà passando da l ad $(l - 1)$. Tuttavia si può anche procedere come nella sezione 8.3.1, considerando il funzionale (8.48) con in più il vincolo $\Phi(\vec{q}) = 0$. Quindi, ricordando la (8.43) e supponendo che tutti gli altri vincoli siano lisci e che le uniche forze attive siano conservative, le equazioni di Eulero-Lagrange per il sistema diventano

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} = \lambda \frac{\partial \Phi}{\partial q_k}, \quad k = 1, \dots, l, \quad (8.50)$$

che coincidono proprio con le equazioni di Lagrange di prima specie (5.56) nel caso in cui le forze non vincolari siano conservative. Come già specificato nella sezione 5.3.2, il termine $\lambda \frac{\partial \Phi}{\partial q_k}$, $k = 1, \dots, l$, si interpreta come la componente Lagrangiana della forza dovuta all'ulteriore vincolo (4.11).

Esempio 8.4.2 Riprendiamo l'esempio 4.1.2, solo che adesso consideriamo il disco che rotola senza strisciare lungo un piano inclinato come mostrato nella figura 8.6.

Il sistema ha due gradi di libertà: l'ascissa s del centro e l'angolo di rotolamento φ . La relazione tra s e φ è data dalla (4.14), ovvero da

$$\Phi(s, \varphi) = R\varphi + s = 0. \quad (8.51)$$

Adesso però, anziché sfruttare la (8.51) per esprimere φ in funzione di s , o viceversa, manteniamo i due gradi di libertà φ ed s e consideriamo il vincolo (8.51). La funzione di Lagrange del sistema è la differenza fra l'energia cinetica del disco rigido

$$T = \frac{M}{2} \dot{s}^2 + \frac{1}{2} \left(\frac{MR^2}{2} \right) \dot{\varphi}^2,$$

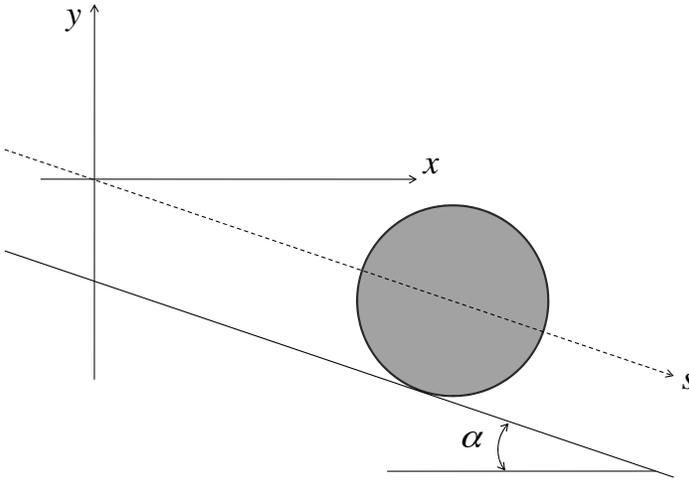


Figura 8.6: L'inclinazione del piano rispetto all'orizzontale è α . Il disco è omogeneo ed M è la sua massa. Il raggio del disco è R . Il peso è diretto nel verso opposto dell'asse y .

e l'energia potenziale dovuta alla forza peso, $V = -Mgs \sin \alpha$, cioè

$$\mathcal{L} = \frac{M}{2} \left(\dot{s}^2 + \frac{R^2}{2} \dot{\varphi}^2 \right) + Mgs \sin \alpha .$$

Il sistema (8.50), (8.51) diventa

$$\begin{cases} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{s}} \right) - \frac{\partial \mathcal{L}}{\partial s} = \lambda \frac{\partial \Phi}{\partial s}, \\ \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \right) - \frac{\partial \mathcal{L}}{\partial \varphi} = \lambda \frac{\partial \Phi}{\partial \varphi}, \\ \Phi(s, \varphi) = 0, \end{cases} \Rightarrow \begin{cases} \ddot{s} - g \sin \alpha = \frac{\lambda}{M}, \\ \frac{R}{2} \ddot{\varphi} = \frac{\lambda}{M}, \\ s = -R\varphi, \end{cases}$$

da cui otteniamo

$$\begin{cases} \ddot{s} = \frac{2}{3}g \sin \alpha, \\ \ddot{\varphi} = -\frac{2}{3R}g \sin \alpha, \\ \lambda = -\frac{Mg}{3} \sin \alpha. \end{cases}$$

8.5 Il principio di Jacobi

8.5.1 Coordinate cicliche nell'ambito del principio di Hamilton

Ricordiamo che se una delle coordinate lagrangiane, diciamo q_k , non compare esplicitamente nella funzione di Lagrange \mathcal{L} (ma la \dot{q}_k ovviamente compare nella \mathcal{L}) la coordinata si dice *ciclica* (si veda la definizione definizione 3.3.1). Come conseguenza delle equazioni di Lagrange (si ricordi la proposizione 3.3.2) si ha che il momento coniugato $p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k}$ è costante.

Abbiamo già visto nella sezioni 3.3.1 e 5.6 l'importanza delle coordinate cicliche nel processo di soluzione delle equazioni di moto, in quanto esse permettono di ridurre la "dimensione" del problema, ovvero il numero di equazioni che restano da risolvere. Si ricordi, per esempio, come nella sezione 3.4 è stata sfruttata la ciclicità della variabile angolare nel problema di Keplero al fine di ridursi a risolvere un problema unidimensionale per la distanza dal centro della forza.

In particolare, nelle sezioni 3.3.1 e 5.6 abbiamo introdotto la *funzione di Routh*, come quella funzione da utilizzare "al posto" della funzione di Lagrange quando si vogliono sfruttare le semplificazioni indotte dalle variabili cicliche. In questa sezione vogliamo mostrare come la funzione di Routh emerga naturalmente nell'ambito del principio variazionale di Hamilton.

Supponiamo di avere un sistema caratterizzato da l variabili lagrangiane e assumiamo, per semplicità, che una sola variabile, la q_l , sia ciclica²¹. Abbiamo quindi

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_l} = p_l, \quad \text{con } p_l = \text{costante}, \quad (8.52)$$

da cui, invocando la proposizione 5.6.1, possiamo ricavare \dot{q}_l in funzione della costante p_l e delle altre variabili. Pertanto, utilizzando la notazione della sezione 5.6, scriveremo

$$\dot{q}_l = \widehat{q}_l(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, p_l, t). \quad (8.53)$$

Dobbiamo poi tener conto del fatto che la relazione (8.53) rende la q_l funzione delle altre variabili: una volta assegnate le (q_1, \dots, q_{l-1}) , la q_l si ottiene integrando la (8.53) rispetto al tempo. In altri termini, le funzioni $q_1(t), \dots, q_l(t)$, sono *vincolate*. Il vincolo che le lega però non è del tipo analizzato nella sezione 8.3.1 (per intendersi un vincolo del tipo (8.36)) ma è un vincolo che coinvolge un integrale. La teoria sviluppata nella sezione 8.3.1 non è quindi applicabile al presente caso. Si procede pertanto seguendo una strada diversa da quella illustrata nel paragrafo 8.3.1. L'idea è quella di utilizzare l'equazione (8.53) per eliminare fin dall'inizio la dipendenza dalla variabile \dot{q}_l nell'integrale (8.48).

Nota 8.5.1 *Si osservi che, comunque si varino le funzioni $q_1(t), \dots, q_{l-1}(t)$, se la \dot{q}_l è definita dalla (8.53), la relazione (8.52) viene automaticamente soddisfatta: cioè, per ogni t , vale*

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_l}(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, \widehat{q}_l(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, p_l, t), t) = p_l,$$

²¹Il caso di $m < l$ variabili cicliche è stato analizzato nella sezione 5.6.

dove p_l assume il valore costante scelto. In altri termini, le l funzioni $q_1(t), \dots, q_l(t)$, con $q_l(t)$ ottenuta integrando rispetto al tempo (8.53), giacciono sulla "ipersuperficie" definita dalla (8.52).

Partendo dalla (8.48), si sostituisce nell'integrando la funzione \widehat{q}_l al posto di \dot{q}_l , ottenendo così un nuovo funzionale che dipende soltanto dalle funzioni $q_1(t), \dots, q_{l-1}(t)$, e dalle loro derivate. In altri termini, rifacendosi sempre alla notazione della sezione 5.6, introduciamo il funzionale

$$\begin{aligned} \widehat{\mathcal{A}}[q_1, \dots, q_{l-1}] &= \int_{t_1}^{t_2} \widehat{\mathcal{L}}(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, t) dt \\ &= \int_{t_1}^{t_2} \mathcal{L}\left(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, \widehat{q}_l(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, p_l, t), t\right) dt. \end{aligned} \quad (8.54)$$

Osserviamo che il funzionale definito dalla (8.54) è diverso da quello di partenza, cioè da quello definito dalla (8.48). Quest'ultimo infatti dipende da²² (q_1, \dots, q_{l-1}) , e da $(\dot{q}_1, \dots, \dot{q}_l)$ mentre $\widehat{\mathcal{A}}$ dipende da (q_1, \dots, q_{l-1}) , e da $(\dot{q}_1, \dots, \dot{q}_{l-1})$.

La differenza fra \mathcal{A} ed $\widehat{\mathcal{A}}$ è ancora più profonda: mentre \mathcal{A} soddisfa il principio di Hamilton, $\widehat{\mathcal{A}}$ non lo soddisfa. Vale infatti il seguente

Teorema 8.5.1 *Siano dati t_1 e t_2 , con $t_1 < t_2$, e le configurazioni*

$$\begin{cases} (q_1(t_1), \dots, q_{l-1}(t_1)) = (q_{1a}, \dots, q_{l-1a}), \\ (q_1(t_2), \dots, q_{l-1}(t_2)) = (q_{1b}, \dots, q_{l-1b}), \end{cases} \quad (8.55)$$

e sia $(q_1(t_1), \dots, q_{l-1}(t_1))$ il moto naturale del sistema che soddisfa le condizioni (8.55). La variazione del funzionale $\widehat{\mathcal{A}}$, cioè $\delta\widehat{\mathcal{A}}$, calcolata in corrispondenza di tale moto **non è nulla** in generale, ma vale

$$\delta\widehat{\mathcal{A}} = \int_{t_1}^{t_2} p_l \sum_{i=1}^{l-1} \left[\frac{\partial \widehat{q}_l}{\partial q_i} - \frac{d}{dt} \frac{\partial \widehat{q}_l}{\partial \dot{q}_i} \right] \delta q_i dt, \quad (8.56)$$

dove p_l è la costante data dalla (8.52) e la funzione \widehat{q}_l è data dalla (8.53).

Tale teorema ricalca lo stesso risultato illustrato nella sezione 3.3.1: la funzione di Lagrange $\widehat{\mathcal{L}}(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, t)$ non soddisfa le equazioni di Lagrange. Il teorema 8.5.1 afferma un fatto analogo: il funzionale $\widehat{\mathcal{A}}$, pur dipendendo da $l-1$ funzioni anziché da l come l'originaria azione \mathcal{A} , non soddisfa il teorema 8.4.1.

Dim. del teorema 8.5.1. Consideriamo una $(l-1)$ -upla di funzioni

$$(\delta q_1(t), \dots, \delta q_{l-1}(t)),$$

tali che $\delta q_i(t_1) = \delta q_i(t_2) = 0$, per ogni $i = 1, \dots, l-1$. Calcoliamo esplicitamente la variazione di $\widehat{\mathcal{A}}$ facendo uso della procedura illustrata nella nota 8.2.3, nell'ipotesi

²²Si ricordi che la q_n è ciclica e quindi non compare nella funzione di Lagrange \mathcal{L} .

che $q_1(t), \dots, q_{l-1}(t)$, soddisfino le equazioni di Lagrange (5.36). Otteniamo

$$\delta \widehat{\mathcal{A}} = \int_{t_1}^{t_2} \widehat{\mathcal{L}}((q_1 + \delta q_1), \dots, (q_{l-1} + \delta q_{l-1}), (\dot{q}_1 + \delta \dot{q}_1), \dots, (\dot{q}_{l-1} + \delta \dot{q}_{l-1}), t) dt - \int_{t_1}^{t_2} \widehat{\mathcal{L}}(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, t) dt,$$

dove $\widehat{\mathcal{L}}$ deve essere espressa in termini di \mathcal{L} come mostrato esplicitamente nella (8.54). Abbiamo dunque

$$\delta \widehat{\mathcal{A}} = \int_{t_1}^{t_2} \left\{ \sum_{i=1}^{l-1} \left[\frac{\partial \mathcal{L}}{\partial q_i} \delta q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta \dot{q}_i \right] + \sum_{i=1}^{n-1} \underbrace{\frac{\partial \mathcal{L}}{\partial q_i}}_{p_i} \left[\frac{\partial \widehat{q}_l}{\partial q_i} \delta q_i + \frac{\partial \widehat{q}_l}{\partial \dot{q}_i} \delta \dot{q}_i \right] \right\} dt,$$

dove, lo ricordiamo, $\delta \dot{q}_i = \frac{d}{dt}(\delta q_i)$. Integrando per parti e sfruttando la condizione $\delta q_i(t_1) = \delta q_i(t_2) = 0, i = 1, \dots, l-1$, ed il fatto che $(q_1(t_1), \dots, q_{l-1}(t_1))$ sono il moto naturale, cioè soddisfano le equazioni di Lagrange (5.36), otteniamo

$$\delta \widehat{\mathcal{A}} = \int_{t_1}^{t_2} \sum_{i=1}^{l-1} \underbrace{\left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right]}_{=0} \delta q_i dt + \int_{t_1}^{t_2} p_l \sum_{i=1}^{l-1} \left[\frac{\partial \widehat{q}_l}{\partial q_i} - \frac{d}{dt} \frac{\partial \widehat{q}_l}{\partial \dot{q}_i} \right] \delta q_i dt,$$

ovvero la (8.56).

□

Il principio di Hamilton 8.4.1 non è quindi soddisfatto a meno di non modificare l'azione Hamiltoniana considerando

$$\mathcal{A}_{\mathcal{R}} [q_1, \dots, q_{l-1}] = \int_{t_1}^{t_2} \left[\widehat{\mathcal{L}}(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, t) - p_l \widehat{q}_l(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, p_l, t) \right] dt, \quad (8.57)$$

ovvero

$$\mathcal{A}_{\mathcal{R}} [q_1, \dots, q_{l-1}] = \int_{t_1}^{t_2} \mathcal{R}(q_1, \dots, q_{l-1}, \dot{q}_1, \dots, \dot{q}_{l-1}, t) dt, \quad (8.58)$$

essendo \mathcal{R} la funzione di Routh definita dalla (5.72). Possiamo enunciare il principio variazionale di Hamilton come segue

Proposizione 8.5.1 *Se la variabile q_l è ciclica, allora la $\delta\mathcal{A}_{\mathcal{R}} = 0$ se e sole se $q_1(t), \dots, q_{l-1}(t)$, sono moto naturale cioè risolvono le (5.75), ovvero*

$$\frac{d}{dt} \left(\frac{\partial \mathcal{R}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{R}}{\partial q_i} = 0, \quad i = 1, \dots, l-1.$$

Dim. Calcolando la variazione di $\mathcal{A}_{\mathcal{R}}$ si ottiene

$$\delta\mathcal{A}_{\mathcal{R}} = \int_{t_1}^{t_2} \sum_{k=1}^{l-1} \left[\frac{d}{dt} \left(\frac{\partial \mathcal{R}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{R}}{\partial q_k} \right] \delta q_k dt.$$

Quindi se $\delta\mathcal{A}_{\mathcal{R}} = 0$, allora siccome le funzioni $\delta q_k, k = 1, \dots, l-1$, sono indipendenti possiamo applicare il lemma 8.3.1 e così ottenere il sistema (5.75). Viceversa se $q_1(t), \dots, q_{l-1}(t)$, soddisfano le (5.75) allora $\delta\mathcal{A}_{\mathcal{R}} = 0$.

□

Abbiamo quindi provato che, se vogliamo tener conto del vincolo $p_l = \frac{\partial \mathcal{L}}{\partial \dot{q}_l} = \text{costante}$, la minimizzazione del funzionale \mathcal{A} si riduce alla minimizzazione del funzionale (8.58). La soluzione di tale problema variazionale permette di determinare l'evoluzione temporale delle prime $l-1$ variabili indipendentemente dalla l -esima. Quest'ultima potrà essere ricavata *a posteriori* dalla (8.53), il cui secondo membro è diventato una funzione nota del tempo e della costante p_l .

8.5.2 Il tempo come variabile

Consideriamo ora per semplicità il caso di una Lagrangiana \mathcal{L} che non dipende esplicitamente dal tempo.

Invece di studiare il problema del moto negli usuali termini in cui le coordinate lagrangiane q_k sono funzioni del tempo e il tempo è la variabile indipendente, introduciamo un cambiamento di variabili, sotto forma parametrizzazione delle q_k e del tempo stesso tramite una nuova variabile indipendente²³ τ . Avremo quindi²⁴

$$t = \tilde{t}(\tau), \Leftrightarrow \tau = \tau(t), \quad \text{cioè } \tilde{t}(\tau) \text{ è invertibile,} \quad (8.59)$$

$$\tilde{q}_k(\tau) = q_k(\tilde{t}(\tau)), \Leftrightarrow q_k(t) = \tilde{q}_k(\tau(t)), \quad (8.60)$$

$$\tilde{q}'_k(\tau) = \frac{dq_k(\tilde{t}(\tau))}{d\tau} = \dot{q}_k(\tilde{t}(\tau)) \tilde{t}'(\tau), \Rightarrow \dot{q}_k(t(\tau)) = \frac{\tilde{q}'_k(\tau)}{\tilde{t}'(\tau)}, \quad (8.61)$$

dove l'apice ' indica la derivata rispetto a τ .

Con questo cambiamento di variabili l'azione Hamiltoniana (8.48) diventa

$$\mathcal{A} = \int_{\tau_1}^{\tau_2} \left[\tilde{\mathcal{L}}(\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau), \tilde{q}'_1(\tau), \dots, \tilde{q}'_l(\tau)) \tilde{t}'(\tau) \right] d\tau, \quad (8.62)$$

²³Per esempio si può pensare a τ come al parametro d'arco s lungo la traiettoria. Ovviamente possiamo ragionare in questo modo anche se la Lagrangiana dipende esplicitamente dal tempo.

²⁴Assumendo $\tilde{t}(\tau)$ invertibile $\tilde{t}'(\tau)$ ha segno definito. Noi considereremo sempre $\tilde{t}'(\tau) > 0$.

dove $\tilde{t}'(\tau) d\tau$ sostituisce dt e dove

$$\begin{aligned} & \tilde{\mathcal{L}}(\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau), \tilde{q}'_1(\tau), \dots, \tilde{q}'_l(\tau)) \\ &= \mathcal{L}(q_1(\tilde{t}(\tau)), \dots, q_l(\tilde{t}(\tau)), \dot{q}_1(\tilde{t}(\tau)), \dots, \dot{q}_l(\tilde{t}(\tau))) \\ &\stackrel{(8.61)}{=} \mathcal{L}\left(\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau), \frac{\tilde{q}'_1(\tau)}{\tilde{t}'(\tau)}, \dots, \frac{\tilde{q}'_l(\tau)}{\tilde{t}'(\tau)}\right). \end{aligned} \tag{8.63}$$

Possiamo calcolare la variazione prima di questo funzionale, facendo variare oltre alle \tilde{q}_k anche la funzione \tilde{t} con le condizioni $\delta\tilde{q}_k = 0$ e $\delta\tilde{t} = 0$ per²⁵ $\tau = \tau_1$ e $\tau = \tau_2$. Se ora scriviamo le equazioni di Eulero-Lagrange per l'annullarsi della variazione di (8.62) abbiamo un sistema di $l + 1$ equazioni. Le prime l sono equivalenti alle ordinarie equazioni di Lagrange (a parte il fattore moltiplicativo \tilde{t}'). La $(l + 1)$ -esima equazione è invece

$$\frac{d}{d\tau} \left[\frac{\partial(\tilde{\mathcal{L}} \tilde{t}')}{\partial \tilde{t}'} \right] = 0, \tag{8.64}$$

dal momento che $\tilde{t}(\tau)$ è *ciclica*, cioè non compare esplicitamente nella $\tilde{\mathcal{L}}$. Quindi, se poniamo

$$p_t = \frac{\partial(\tilde{\mathcal{L}} \tilde{t}')}{\partial \tilde{t}'}, \stackrel{(8.64)}{\implies} p_t = \text{costante}. \tag{8.65}$$

Calcoliamo adesso esplicitamente p_t , detto anche momento coniugato a \tilde{t} ,

$$\begin{aligned} p_t &= \tilde{t}' \left(\frac{\partial \tilde{\mathcal{L}}}{\partial \tilde{t}'} \right) + \tilde{\mathcal{L}} \\ &= \tilde{t}' \frac{\partial}{\partial \tilde{t}'} \left[\mathcal{L}\left(\tilde{q}_1, \dots, \tilde{q}_l, \frac{\tilde{q}'_1}{\tilde{t}'}, \dots, \frac{\tilde{q}'_l}{\tilde{t}'}\right) \right] + \mathcal{L}\left(\tilde{q}_1, \dots, \tilde{q}_l, \frac{\tilde{q}'_1}{\tilde{t}'}, \dots, \frac{\tilde{q}'_l}{\tilde{t}'}\right) \\ &= -\frac{1}{\tilde{t}'} \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \tilde{q}_i} \tilde{q}'_i + \underbrace{\mathcal{L}\left(\tilde{q}_1, \dots, \tilde{q}_l, \frac{\tilde{q}'_1}{\tilde{t}'}, \dots, \frac{\tilde{q}'_l}{\tilde{t}'}\right)}_{\tilde{\mathcal{L}}(\tilde{q}_1, \dots, \tilde{q}_l, \tilde{q}'_1, \dots, \tilde{q}'_l)}. \end{aligned} \tag{8.66}$$

Se introduciamo la funzione Hamiltoniana²⁶

$$\begin{aligned} & \tilde{\mathcal{H}}(\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau), \tilde{q}'_1(\tau), \dots, \tilde{q}'_l(\tau)) \\ &= \mathcal{H}(q_1(\tilde{t}(\tau)), \dots, q_l(\tilde{t}(\tau)), \dot{q}_1(\tilde{t}(\tau)), \dots, \dot{q}_l(\tilde{t}(\tau))), \end{aligned} \tag{8.67}$$

²⁵Si assume $\tilde{t}(\tau_1) = t_1$ e $\tilde{t}(\tau_2) = t_2$.

²⁶Osserviamo che la (8.67) può anche essere derivata se, in accordo con la (5.78), definiamo

$$\tilde{\mathcal{H}} = \sum_{i=1}^l \frac{\partial \tilde{\mathcal{L}}}{\partial q'_i} \tilde{q}'_i - \tilde{\mathcal{L}},$$

e poi sfruttiamo la (8.61) e la (8.63).

per cui, dalla definizione (5.78),

$$\begin{aligned}\tilde{\mathcal{H}} &= \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i(\tilde{t}(\tau)) - \mathcal{L}(q_1(\tilde{t}(\tau)), \dots, q_l(\tilde{t}(\tau)), \dot{q}_1(\tilde{t}(\tau)), \dots, \dot{q}_l(\tilde{t}(\tau))) \\ &= \frac{1}{\tilde{t}'} \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i' - \mathcal{L}\left(\tilde{q}_1, \dots, \tilde{q}_l, \frac{\tilde{q}_1'}{\tilde{t}'}, \dots, \frac{\tilde{q}_l'}{\tilde{t}'}\right),\end{aligned}\quad (8.68)$$

abbiamo

$$\tilde{\mathcal{H}}(\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau), \tilde{q}_1'(\tau), \dots, \tilde{q}_l'(\tau)) = -p_t = \text{costante}.\quad (8.69)$$

In particolare, tornando alle variabili originarie, cioè considerando t come variabile indipendente e ricordando le (8.60), (8.61), abbiamo

$$p_t = \mathcal{L}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l) - \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i,\quad (8.70)$$

ovvero

$$\mathcal{H}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l) = -p_t = \text{costante}.\quad (8.71)$$

Abbiamo quindi ottenuto, ma con un formalismo completamente diverso, lo stesso risultato dimostrato nella sezione 5.7: se la funzione Lagrangiana non dipende esplicitamente dal tempo, se le uniche forze attive sono conservative e se i vincoli sono lisci allora vale la (5.81).

La (8.71) ci dà anche un'informazione in più: si può guardare all'Hamiltoniana (cambiata di segno) come alla variabile coniugata della variabile "tempo" t . La (8.65) ricalca infatti la definizione 9.0.1 del capitolo 9 se consideriamo \tilde{t}' al posto di \dot{q}_k e $\tilde{\mathcal{L}}$ al posto di \mathcal{L} .

8.5.3 Formulazione del principio di Jacobi

La condizione (8.71) rappresenta un *vincolo* per le $q_i(t)$, $\dot{q}_i(t)$, $i = 1, \dots, l$. Tale vincolo tuttavia è diverso da quello analizzato nel paragrafo 8.5.1 dal momento che la (8.71), coinvolgendo tutte le q_i e tutte le \dot{q}_i , non consente, in generale, di esprimere una variabile lagrangiana in funzione delle rimanenti ($l - 1$). Se quindi vogliamo tener conto di tale vincolo sorge il problema di come incorporarlo nel principio di Hamilton. In pratica, come mostrato nella sezione 8.5.1, dobbiamo modificare l'azione Hamiltoniana. In definitiva, tener conto del vincolo (8.71) significa individuare, come già accennato nella nota 3.3.1, una particolare funzione di Routh.

La strada da seguire è stata essenzialmente tracciata nella sezioni 8.5.1 e 8.5.2: introduciamo una nuova variabile indipendente τ e consideriamo il tempo come variabile dipendente. Osservando poi che la funzione riportata nella parentesi quadra di (8.62), ovvero $\tilde{\mathcal{L}} \tilde{t}'$, *non dipende esplicitamente dalla variabile \tilde{t}* ma solo dalla sua derivata \tilde{t}' , si può attuare la strategia di riduzione delle variabili del paragrafo 8.5.1, tornando così a "sole" l variabili $\tilde{q}_1(\tau_1), \dots, \tilde{q}_l(\tau_1)$, non più vincolate.

Seguendo i passaggi illustrati nel caso generale, dobbiamo considerare solo l variabili $\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau)$, e sfruttare la (8.65), ovvero la (8.66), per ricavare

$$\tilde{t}' = \tilde{t}'(\underbrace{\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau)}_{\tilde{q}(\tau)}, \underbrace{\tilde{q}_1'(\tau), \dots, \tilde{q}_l'(\tau)}_{\tilde{q}'(\tau)}),\quad (8.72)$$

e da essa t in funzione di τ , cioè $\tilde{t}(\tau)$. In tal modo, come osservato nella nota 8.5.1, il vincolo (8.69) sarà automaticamente soddisfatto.

Definiamo adesso l'insieme \tilde{Q} , detto **classe dei moti variati asincroni isoenergetici**, che costituisce il dominio della nuova azione. L'insieme \tilde{Q} è formato da tutte le l -uple di funzioni $\tilde{q}(\tau)$, tali che:

- (1). $\tilde{q}_i(\tau) \in C^2([\tau_1, \tau_2]), i = 1, \dots, l$.
- (2). $\tilde{q}(\tau_1) = \tilde{q}_a$, e $\tilde{q}(\tau_2) = \tilde{q}_b$, con \tilde{q}_a, \tilde{q}_b definite dalla (8.47).

Nota 8.5.2 *E' importante sottolineare la differenza fra gli insiemi \tilde{Q} e Q (quest'ultimo introdotto con la definizione 8.4.1). Nelle l -uple di \tilde{Q} non compare il tempo e quindi \tilde{Q} è l'insieme di tutte le curve di \mathbb{R}^l che connettono \tilde{q}_a e \tilde{q}_b , non sottoposte ad alcun vincolo. Infatti data una qualunque $\tilde{q}(\tau) \in \tilde{Q}$, si determinano, per mezzo della (8.66), la funzione $\tilde{t}'(\tilde{q}(\tau), \tilde{q}'(\tau))$ e, successivamente t in funzione di τ , sì che il vincolo (8.69) risulta automaticamente soddisfatto.*

La classe Q è invece l'insieme di tutti i moti che partono al tempo t_1 da \tilde{q}_a e terminano al tempo t_2 in \tilde{q}_b . Gli elementi di Q sono dunque le l -uple $\tilde{q}(t)$ che rappresentano i possibili moti (quindi elementi "fisici") non soggetti ad alcun vincolo. Gli elementi di \tilde{Q} sono invece enti geometrici che non danno informazioni sullo sviluppo temporale del moto, ovvero sulla legge oraria con cui la traiettoria è percorsa.

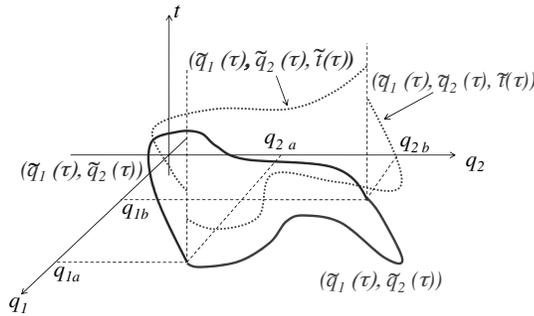


Figura 8.7: Le due curve continue corrispondono a due generiche traiettorie $(\tilde{q}_1(\tau_1), \tilde{q}_2(\tau_1))$. I relativi moti $(\tilde{q}_1(\tau_1), \tilde{q}_2(\tau_1), \tilde{t}(\tau))$ sono rappresentati dalle curve tratteggiate. In particolare, una volta fissate le traiettorie, la legge oraria con cui si esprime il tempo t in funzione di τ , viene determinata dalla (8.65), ovvero dalla (8.66).

Nella figura 8.7 sono riportate sul piano (q_1, q_2) due possibili traiettorie $(\tilde{q}_1(\tau_1), \tilde{q}_2(\tau_1))$, cui corrispondono, nello spazio (q_1, q_2, t) , i relativi moti $(\tilde{q}_1(\tau_1), \tilde{q}_2(\tau_1), \tilde{t}(\tau))$ (curve tratteggiate). Quest'ultimi danno informazioni sulla traiettoria e anche sulla legge oraria con cui la stessa è percorsa. Si nota che gli istanti di partenza e di arrivo non sono gli stessi per entrambe i moti. Infatti, una volta selezionati: punto di

partenza, punto di arrivo, e la particolare traiettoria $(\tilde{q}_1(\tau_1), \tilde{q}_2(\tau_1))$ che li connette, il corrispondente moto $(\tilde{q}_1(\tau), \tilde{q}_2(\tau), \tilde{t}(\tau))$ si determina calcolando la legge oraria $\tilde{t}(\tau)$. Questo si fa tramite la (8.66) che permette di determinare $\tilde{t}'(\tau)$ e poi, integrando la (8.72), di esprimere t come funzione di τ . Il moto così determinato è tale da soddisfare automaticamente il vincolo (8.69).

Seguendo le (8.57) e (8.58), la nuova Lagrangiana, o, se vogliamo, la funzione di Routh, è data da

$$\mathcal{L}_R = \tilde{\mathcal{L}}\tilde{t}' - p_t \tilde{t}' = \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \tilde{q}'_i,$$

e pertanto la nuova azione è così definita: $\mathcal{A}_R : \tilde{\mathcal{Q}} \rightarrow \mathbb{R}$, ed opera così

$$\mathcal{A}_R[\tilde{\mathbf{q}}(\tau)] = \int_{\tau_1}^{\tau_2} \mathcal{L}_R d\tau = \int_{\tau_1}^{\tau_2} \left(\sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \tilde{q}'_i \right) d\tau,$$

che, ricordando la (8.61), si può anche scrivere come

$$\mathcal{A}_R = \int_{\tau_1}^{\tau_2} \sum_{i=1}^l \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i \tilde{t}' d\tau.$$

Se adesso ipotizziamo che l'energia cinetica del sistema sia²⁷ $T = T_2$, sfruttando la (5.83) abbiamo

$$\sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i = 2T,$$

e l'azione diventa quindi

$$\mathcal{A}_R = \int_{\tau_1}^{\tau_2} 2T \tilde{t}' d\tau. \quad (8.73)$$

In questa forma però compare ancora \tilde{t}' data dalla (8.72) che è stata determinata sfruttando la conservazione del momento p_t .

Ora supponiamo che in \mathbb{R}^l sia dato un sistema di coordinate curvilinee (q_1, \dots, q_l) la cui corrispondente matrice metrica (si veda la sezione 1.4) è la matrice²⁸ \mathbb{A} dell'energia cinetica definita nella sezione 4.1.2. Se fissiamo una generica $\tilde{\mathbf{q}}(\tau) \in \tilde{\mathcal{Q}}$, questa definisce una curva, o meglio una traiettoria, $\gamma : [\tau_1, \tau_2] \rightarrow \mathbb{R}^l$,

$$\tau \xrightarrow{\gamma} \mathbf{x}(\tau) = \sum_{i=1}^l x_i(\tilde{q}_1(\tau), \dots, \tilde{q}_l(\tau)) \mathbf{e}_i. \quad (8.74)$$

Il vettore tangente, $\mathbf{x}'(\tau)$, può essere espresso utilizzando la base locale $\{\mathbf{u}_i\}$ legata alle coordinate curvilinee

$$\mathbf{x}'(\tau) = \sum_{i=1}^l \tilde{q}'_i(\tau) \mathbf{u}_i.$$

²⁷Si ricordi la (4.33).

²⁸Si ricordi che la matrice \mathbb{A} , di componenti a_{ij} , è una matrice simmetrica definita positiva.

Di conseguenza

$$\begin{aligned} \|\mathbf{x}'(\tau)\|^2 &= \sum_{i,j=1}^l a_{ij}(\tilde{\mathbf{q}}(\tau)) \dot{q}_i(\tau) \dot{q}_j(\tau) \\ &\stackrel{(8.61)}{=} \underbrace{\left[\sum_{i,j=1}^l a_{ij}(\tilde{\mathbf{q}}(\tau)) \dot{q}_i(\tau) \dot{q}_j(\tau) \right]}_{\substack{2T \\ (4.38)}} \tilde{t}'^2, \end{aligned}$$

per cui la lunghezza della curva γ è

$$\int_{\gamma} ds = \int_{\tau_1}^{\tau_2} \|\mathbf{x}'(\tau)\| d\tau = \int_{\tau_1}^{\tau_2} \sqrt{2T} \tilde{t}' d\tau,$$

dove con ds si denota la lunghezza dell'elemento d'arco, cioè

$$ds = \|\mathbf{x}'(\tau)\| d\tau = \sqrt{2T} \tilde{t}' d\tau. \tag{8.75}$$

Possiamo allora riscrivere il funzionale (8.73) da minimizzare come²⁹

$$\begin{aligned} \mathcal{A}_R &= \int_{\tau_1}^{\tau_2} 2T \tilde{t}' d\tau = \int_{\tau_1}^{\tau_2} \sqrt{2T} \underbrace{\sqrt{2T} \tilde{t}'}_{ds} d\tau \\ &= \int_{\gamma} \frac{\sqrt{2T}}{\sqrt{2(E-V)}} ds = \int_{\gamma} \sqrt{2(E-V)} ds. \end{aligned}$$

dove γ è la traiettoria data da (8.74). Nell'ultimo integrale è “scomparso” definitivamente il tempo, quindi la soluzione delle equazioni che si ottengono annullandone la variazione prima conterrà solo delle *informazioni geometriche*: saranno le equazioni della traiettoria del punto rappresentativo nello spazio delle configurazioni. Una volta determinata la traiettoria del moto, la legge oraria potrà essere ricavata dalla (8.75), cioè

$$\tilde{t}' = \frac{\|\mathbf{x}'(\tau)\|}{\sqrt{2T}} = \frac{\|\mathbf{x}'(\tau)\|}{\sqrt{2(E-V)}}, \tag{8.76}$$

poiché a questo punto il secondo membro della (8.76) sarà una funzione nota del parametro τ e la funzione $\tilde{t}(\tau)$, appunto la legge oraria, potrà essere ricavata per integrazione rispetto a τ .

Abbiamo quindi determinato un principio variazionale la cui soluzione determina completamente la traiettoria del moto nello spazio delle configurazioni. Questo principio prende il nome di **Principio di Jacobi**³⁰ ed è la riscrittura “formalmente corretta” di un più antico principio variazionale, noto anche come **Principio di Minima**

²⁹Si ricordi che $T + V = E$, con $E =$ costante.

³⁰Carl Gustav Jacob Jacobi (Potsdam, 1804 – Berlino, 1851) è stato un matematico tedesco.

Azione, che prevedeva che il moto annullasse la variazione dell'integrale di azione³¹ $\int_{t_1}^{t_2} (\sum_{i=1}^n p_i \dot{q}_i) dt = \int_{t_1}^{t_2} 2T dt$, nella classe dei moti variati isoenergetici.

Si può dare anche un'altra formulazione dell'azione se consideriamo come matrice metrica

$$\mathbb{G} = 2T \mathbb{A} = 2(E - V) \mathbb{A}, \quad (8.77)$$

dove \mathbb{A} è definita dalla (4.38). In questo caso

$$ds = \|\mathbf{x}'(\tau)\| d\tau = \sqrt{2(E - V) \sum_{i,j=1}^l a_{ij}(\tilde{\mathbf{q}}(\tau)) \dot{q}_i(\tau) \dot{q}_j(\tau) \tilde{t}' d\tau},$$

e quindi

$$\mathcal{A}_R = \int_{\gamma} ds,$$

rappresenta proprio la lunghezza della traiettoria γ . La matrice metrica (8.77) viene detta **metrica di Jacobi**. Possiamo quindi riformulare come segue il principio variazionale: *il moto di un sistema meccanico soggetto a forze conservative e vincoli lisci avviene, nello spazio delle configurazioni munito della metrica di Jacobi, lungo le geodetiche (linee di più breve cammino).*

E' interessante analizzare il caso in cui un singolo punto materiale si muova per inerzia (ovvero senza forze direttamente applicate, e quindi $V \equiv 0$) su una superficie liscia (come descritto nella sezione 4.1.3). In questo $2T = 2E$, e quindi ds definito dall'energia cinetica coincide proprio con la lunghezza dell'elemento di arco superficie (moltiplicato per la massa e $\sqrt{2E}$, che sono fattori costanti).

Ne segue l'importante risultato: *il moto per inerzia di un punto materiale su una superficie avviene lungo le linee (geodetiche).*

³¹Il principio di minima azione, nella forma presente, è dovuto a Lagrange, ma è in genere noto sotto il nome di Principio di Maupertuis che lo anticipò nella sua memoria del 1744. L'“errore” messo in evidenza da Jacobi consiste nel fatto che non è corretto assumere il tempo come variabile indipendente. Questa “sottigliezza” era però chiara a Lagrange

Capitolo 9

Il Sistema Canonico

Abbiamo osservato nel paragrafo 5.4 che la risolubilità delle equazioni di Lagrange è legata alla possibilità di mettere il sistema delle equazioni in forma normale, cioè di “risolverle” (algebricamente) rispetto alle variabili di ordine massimo, le \dot{q}_k . Poi si trasforma il sistema originario di l equazioni differenziali del secondo ordine in un sistema di $2l$ equazioni del primo ordine, definendo delle nuove variabili $\eta_k = \dot{q}_k$.

Esiste un'altra riduzione al primo ordine che porta a un sistema noto come **sistema canonico** o **sistema di Hamilton**. Il punto di partenza è sempre lo stesso: *la non-singolarità della matrice delle derivate seconde rispetto alle \dot{q}_k* . Ricordando la definizione della matrice \mathbb{A} tramite (4.34) ed il teorema 4.1.3, abbiamo $\det \mathbb{A} \neq 0$. Ora, partendo dalle (5.34) e (4.33), cioè

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^l a_{ij}(\tilde{\mathbf{q}}, t) \dot{q}_i \dot{q}_j + \sum_{i=1}^l b_i(\tilde{\mathbf{q}}, t) \dot{q}_i + \frac{1}{2} c(\tilde{\mathbf{q}}, t) - \widehat{V}(\tilde{\mathbf{q}}, t),$$

otteniamo

$$\det \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{q}_i \partial \dot{q}_j} \right) = \det \mathbb{A} \neq 0. \quad (9.1)$$

Estendiamo adesso la definizione 3.3.2 al caso generale di l variabili lagrangiane.

Definizione 9.0.1 Si dicono **momenti coniugati delle variabili q_k** le funzioni

$$p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k}(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l, t), \quad k = 1, \dots, l. \quad (9.2)$$

Le p_k sono l funzioni delle variabili $(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l)$ ed eventualmente del tempo se la Lagrangiana ne dipende esplicitamente. Infatti, esplicitando la (9.2) abbiamo

$$p_k = \sum_{j=1}^l a_{kj}(\tilde{\mathbf{q}}, t) \dot{q}_j + b_k(\tilde{\mathbf{q}}, t), \quad k = 1, \dots, l,$$

ovvero, con ovvio significato di simboli

$$\tilde{\mathbf{p}} = \mathbb{A}(\tilde{\mathbf{q}}, t) \tilde{\dot{\mathbf{q}}} + \tilde{\mathbf{b}}(\tilde{\mathbf{q}}, t),$$

La condizione (9.1) ci dice che, per ogni fissato (q_1, \dots, q_l) , la mappa tra le coordinate $(\dot{q}_1, \dots, \dot{q}_l)$ e le (p_1, \dots, p_l) definite da 9.2 è **invertibile**, cioè

$$\tilde{\mathbf{q}} = \mathbb{A}^{-1} (\tilde{\mathbf{p}} - \tilde{\mathbf{b}}), \quad (9.3)$$

che ci consente di esprimere le $\tilde{\mathbf{q}}$ in funzione delle $\tilde{\mathbf{p}}$ e delle $\tilde{\mathbf{q}}$ (e ovviamente del tempo t se, al solito, la Lagrangiana ne presenta una esplicita dipendenza). Possiamo quindi definire un “nuovo sistema di coordinate” per lo stato cinematico del sistema (e non solo per le configurazioni) usando al posto delle $(q_1, \dots, q_l, \dot{q}_1, \dots, \dot{q}_l)$ le variabili $(q_1, \dots, q_l, p_1, \dots, p_l)$.

Per determinare a quali equazioni debbano soddisfare le nuove variabili, quando le originarie soddisfano alle equazioni di Lagrange, richiamiamo la definizione della funzione di Hamilton 5.78, che scriviamo ora nella forma

$$\mathcal{H}(q_1, \dots, q_l, p_1, \dots, p_l, t) = \sum_{i=1}^l p_i \dot{q}_i - \mathcal{L}, \quad (9.4)$$

dove le \dot{q}_i nella sommatoria e nella Lagrangiana \mathcal{L} sono ora pensate come funzioni delle variabili¹ $(q_1, \dots, q_l, p_1, \dots, p_l)$.

Quindi, pensando la \mathcal{H} come funzione delle q_i e delle p_i , calcoliamo la derivata di \mathcal{H} rispetto alla q_i

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial q_i} &= \frac{\partial}{\partial q_i} \left(\sum_{k=1}^l p_k \dot{q}_k - \mathcal{L} \right) \\ &= \sum_{k=1}^l p_k \frac{\partial \dot{q}_k}{\partial q_i} - \underbrace{\frac{\partial \mathcal{L}}{\partial q_i}}_{\substack{\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \\ p_i}} - \sum_{h=1}^l \underbrace{\frac{\partial \mathcal{L}}{\partial \dot{q}_h}}_{p_h} \frac{\partial \dot{q}_h}{\partial q_i} \\ &= -\frac{dp_i}{dt}. \end{aligned}$$

E' doveroso rimarcare che nel calcolo di $\frac{\partial \mathcal{H}}{\partial q_i}$ abbiamo sfruttato le equazioni di Lagrange nella forma (5.36). Ciò significa che stiamo implicitamente assumendo che il sistema sia soggetto soltanto a vincoli olonomi e lisci e che le forze attive siano tutte conservative.

¹Le \dot{q}_i , $i = 1, \dots, l$, sono espresse tramite la (9.3).

Calcoliamo adesso da derivata di \mathcal{H} rispetto alla variabile p_i

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial p_i} &= \frac{\partial}{\partial p_i} \left(\sum_{k=1}^l p_k \dot{q}_k - \mathcal{L} \right) \\ &= \dot{q}_i + \sum_{k=1}^l p_k \frac{\partial \dot{q}_k}{\partial p_i} - \sum_{h=1}^l \underbrace{\frac{\partial \mathcal{L}}{\partial \dot{q}_h}}_{p_h} \frac{\partial \dot{q}_h}{\partial p_i} \\ &= \frac{dq_i}{dt} . \end{aligned}$$

Otteniamo quindi

$$\begin{cases} \dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \\ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \end{cases}, \quad i = 1, \dots, l \quad (9.5)$$

assieme con la relazione

$$\frac{d\mathcal{H}}{dt} = \sum_{k=1}^l \underbrace{\frac{\partial \mathcal{H}}{\partial p_k}}_{\dot{q}_k} \dot{p}_k + \underbrace{\frac{\partial \mathcal{H}}{\partial q_k}}_{-\dot{p}_k} \dot{q}_k + \frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \mathcal{H}}{\partial t} .$$

La funzione \mathcal{H} è dunque una costante del moto se la Lagrangiana, e la Hamiltoniana di conseguenza, non dipendono esplicitamente dal tempo.

Il dominio delle variabili $(q_1, \dots, q_l, p_1, \dots, p_l)$ è detto **spazio delle fasi**, mentre il sistema di equazioni differenziali (9.5) è detto **sistema canonico**.

9.1 Il Teorema di Liouville

La struttura del sistema di Hamilton, oltre alla sua “gradevolezza estetica” dovuta alla simmetria assunta dal sistema rispetto alla “coppia di variabili” p e q , è ricca di conseguenze, che richiedono però, per essere messe in evidenza, una trattazione molto tecnica che supera gli scopi di questo corso. Una conseguenza è però semplice da ottenere come corollario di un teorema generale sui sistemi di equazioni differenziali, che ci limitiamo a enunciare senza dimostrazione.

Sia dato un sistema di equazioni differenziali in \mathbb{R}^n

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}), \quad \Leftrightarrow \quad \begin{cases} \dot{x}_1 = F_1(x_1, \dots, x_n), \\ \dot{x}_2 = F_2(x_1, \dots, x_n), \\ \vdots \\ \dot{x}_n = F_n(x_1, \dots, x_n). \end{cases} \quad (9.6)$$

Sia $\Omega_0 \subset \mathbb{R}^n$, aperto. Indichiamo con $\Phi_t(\xi)$ l’applicazione

$$\xi \xrightarrow{\Phi_t(\xi)} \mathbf{x} = \Phi_t(\xi),$$

che opera così: ad ogni $\xi \in \Omega_0$, associa $x = \Phi_t(\xi)$, soluzione del problem di Cauchy

$$\begin{cases} \dot{x} = F(x), \\ x(0) = \xi. \end{cases}$$

In particolare, $\Phi_0(\xi) = \xi$, mentre, se indichiamo con Ω_t l'immagine al tempo t dell'aperto Ω_0 , abbiamo $\Omega_t = \Phi_t(\Omega_0)$ (si veda la figura 9.1).

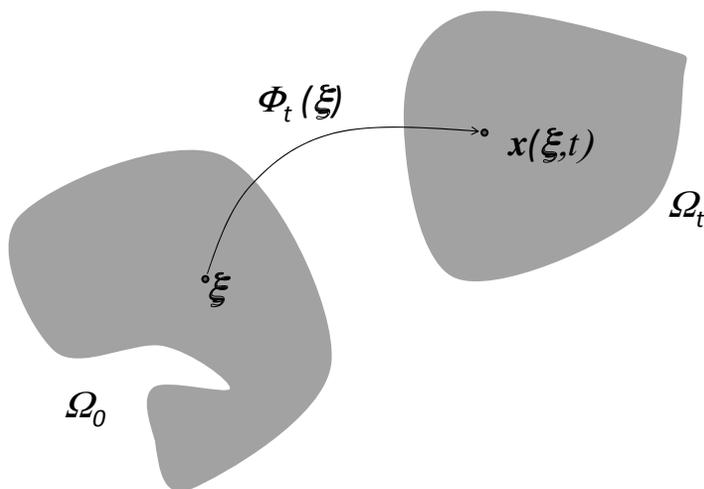


Figura 9.1: Rappresentazione schematica dell'applicazione $\Phi_t(\xi)$.

Se $F(x)$ è tale che possiamo applicare le ipotesi del teorema 2.3.1, allora l'applicazione Φ_t è definita (almeno in un intervallo di tempo finito) ed invertibile (ad ogni dato iniziale ξ corrisponde uno ed un solo x). Inoltre, si può provare (ma non lo faremo) che Φ_t è continua (cioè Φ_t porta punti "vicini" in "punti vicini"), sì che Ω_t è un dominio aperto se lo è Ω_0 . Vale poi questo importante:

Teorema 9.1.1 *Le seguenti affermazioni sono equivalenti:*

1. $\operatorname{div} F(x) = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i} = 0$, cioè il campo vettoriale F ha divergenza nulla;
2. l'applicazione $\xi \rightarrow \Phi_t(x)$ preserva il volume per ogni t , ovvero il volume dell'insieme Ω_t è uguale al volume di Ω_0 , per ogni dominio $\Omega_0 \subset \mathbb{R}^n$ e per ogni tempo t .

Dim. Per semplicità proviamo il teorema in \mathbb{R}^2 . Indichiamo con $|\Omega_0|$ la misura, cioè la superficie, del dominio Ω_0 , e con $|\Omega_t|$ la misura di Ω_t . Abbiamo

$$|\Omega_0| = \int_{\Omega_0} d\xi_1 d\xi_2, \quad |\Omega_t| = \int_{\Omega_t} dx_1 dx_2.$$

Calcoliamo la derivata rispetto al tempo di $|\Omega_t|$, cioè $\frac{d}{dt} \int_{\Omega_t} dx_1 dx_2$. Ovviamente tale calcolo non è assolutamente agevole dal momento che il dominio di integrazione varia nel tempo. Tuttavia, ricordando l'applicazione Φ_t fra Ω_0 e Ω_t , possiamo effettuare il cambio di variabile

$$\begin{cases} x_1 = x_1(\xi_1, \xi_2, t), \\ x_2 = x_2(\xi_1, \xi_2, t), \end{cases}$$

e scrivere

$$|\Omega_t| = \int_{\Omega_0} \det \mathbb{J} d\xi_1 d\xi_2,$$

dove \mathbb{J} è la matrice jacobiana

$$\mathbb{J} = \begin{pmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_2} \\ \frac{\partial x_2}{\partial \xi_1} & \frac{\partial x_2}{\partial \xi_2} \end{pmatrix}, \Rightarrow \det \mathbb{J} = \frac{\partial x_1}{\partial \xi_1} \frac{\partial x_2}{\partial \xi_2} - \frac{\partial x_1}{\partial \xi_2} \frac{\partial x_2}{\partial \xi_1}. \quad (9.7)$$

Tornando adesso alla derivata rispetto al tempo di $|\Omega_t|$, possiamo scrivere

$$\frac{d|\Omega_t|}{dt} = \int_{\Omega_0} \frac{d}{dt} (\det \mathbb{J}) d\xi_1 d\xi_2, \quad (9.8)$$

e quindi dobbiamo calcolare la derivata temporale del $\det \mathbb{J}$. Questo la facciamo sfruttando la formula esplicita (9.7)

$$\begin{aligned} \frac{d}{dt} (\det \mathbb{J}) &= \frac{d}{dt} \left(\frac{\partial x_1}{\partial \xi_1} \right) \frac{\partial x_2}{\partial \xi_2} - \frac{d}{dt} \left(\frac{\partial x_1}{\partial \xi_2} \right) \frac{\partial x_2}{\partial \xi_1} \\ &\quad + \frac{\partial x_1}{\partial \xi_1} \frac{d}{dt} \left(\frac{\partial x_2}{\partial \xi_2} \right) - \frac{\partial x_1}{\partial \xi_2} \frac{d}{dt} \left(\frac{\partial x_2}{\partial \xi_1} \right). \end{aligned} \quad (9.9)$$

Considerando le derivate $\frac{d}{dt} \left(\frac{\partial x_i}{\partial \xi_k} \right)$, $i = 1, 2, k = 1, 2$, abbiamo

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial x_i}{\partial \xi_k} \right) &= \frac{\partial \dot{x}_i}{\partial \xi_k} \stackrel{(9.6)}{=} \frac{\partial F_i(x_1(\xi_1, \xi_2, t), x_2(\xi_1, \xi_2, t))}{\partial \xi_k} \\ &= \frac{\partial F_i}{\partial x_1} \frac{\partial x_1}{\partial \xi_k} + \frac{\partial F_i}{\partial x_2} \frac{\partial x_2}{\partial \xi_k}. \end{aligned}$$

Sostituendo nella (9.9) otteniamo

$$\frac{d}{dt}(\det \mathbb{J}) = \frac{\partial F_1}{\partial x_1}(\det \mathbb{J}) + \frac{\partial F_2}{\partial x_2}(\det \mathbb{J}) = (\det \mathbb{J}) \operatorname{div} \mathbf{F}.$$

Quindi, se $\operatorname{div} \mathbf{F} = 0$, abbiamo $\frac{d|\Omega_t|}{dt} = 0$: il volume viene preservato. Viceversa se, per ogni dominio e per ogni tempo, $\frac{d|\Omega_t|}{dt} = 0$, cioè se il volume rimane costante, allora nella (9.8) deve annullarsi l'integrando e di conseguenza $\operatorname{div} \mathbf{F} = 0$.

□

E' immediato verificare che il secondo membro del sistema di Hamilton ha divergenza nulla, quindi

Teorema 9.1.2 (Teorema di Liouville) *Il moto preserva il volume nello spazio delle fasi.*

□

Questo teorema ha una straordinaria importanza in molti campi, in special modo nella meccanica statistica. Esso ha però anche una conseguenza immediata sulla struttura delle posizioni di equilibrio di un sistema meccanico conservativo.

Corollario 9.1.1 *Un sistema meccanico conservativo non ammette posizioni di equilibrio asintoticamente stabili.*

Dim. Se $(\mathbf{q}^e, \mathbf{p}^e) = (q_1^e, \dots, q_l^e, p_1^e, \dots, p_l^e)$ fosse un punto di equilibrio asintoticamente stabile, dopo un tempo finito T tutti i punti contenuti in una "sfera" di centro $(\mathbf{q}^e, \mathbf{p}^e)$ e raggio R qualsiasi, si troverebbero all'interno di una sfera di raggio $R/2$, contraddicendo la conservazione del volume.

□

9.2 Le parentesi di Poisson

Sia $f(\mathbf{q}, \mathbf{p})$ una funzione definita sullo spazio delle fasi, e sia $\mathcal{H}(\mathbf{q}, \mathbf{p})$ la funzione di Hamilton (che per semplicità assumiamo indipendente dal tempo).

Possiamo calcolare la derivata della funzione f lungo le soluzioni del sistema hamiltoniano (9.5). Abbiamo

$$\frac{df}{dt} = \sum_{k=1}^l \frac{\partial f}{\partial q_k} \dot{q}_k + \sum_{k=1}^l \frac{\partial f}{\partial p_k} \dot{p}_k = \sum_{k=1}^l \frac{\partial f}{\partial q_k} \frac{\partial \mathcal{H}}{\partial p_k} - \sum_{k=1}^l \frac{\partial f}{\partial p_k} \frac{\partial \mathcal{H}}{\partial q_k}.$$

Questa ultima formula suggerisce di introdurre per ogni coppia di funzioni definite sullo spazio delle fasi un prodotto tramite

$$\{f, g\} = \sum_{k=1}^l \frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \sum_{k=1}^l \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k}. \quad (9.10)$$

Il “prodotto” così definito $\{f, g\}$ prende il nome di **parentesi di Poisson** tra le funzioni f e g .

La derivata totale di una funzione f lungo le soluzioni del sistema con Halmiltoniana \mathcal{H} può essere espressa come

$$\frac{df}{dt} = \{f, \mathcal{H}\}.$$

Quindi possiamo affermare che

Corollario 9.2.1 *Una funzione $f(\mathbf{q}, \mathbf{p})$ è un integrale primo del moto, cioè $\frac{df}{dt} = 0$, se e solo se la sua parentesi di Poisson con la funzione Hamiltoniana è nulla.*

E' immediato verificare alcune proprietà delle parentesi di Poisson:

- antisimmetria: $\{f, g\} = -\{g, f\}$
- linearità: $\{\alpha f_1 + \beta f_2, g\} = \alpha\{f_1, g\} + \beta\{f_2, g\}$ per ogni tripla di funzioni f_1, f_2 e g e ogni coppia di numeri reali α e β . Analoga formula vale per le combinazioni lineari fatte sul secondo elemento della parentesi.

Un po' meno ovvia ma sempre di verifica elementare (grazie al teorema di Schwarz) è la cosiddetta **identità di Jacobi**

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \tag{9.11}$$

Nota 9.2.1 *Le stesse tre proprietà algebriche elencate per la parentesi di Poisson sono soddisfatte dal prodotto vettoriale di vettori di \mathbb{R}^3 introdotto nella sezione 1.5.1. In particolare, l'identità di Jacobi (1.17) segue dalla bilinearità e dal fatto che $\mathbf{i} \wedge (\mathbf{j} \wedge \mathbf{k}) = 0$.*

Possiamo poi utilizzare le Parentesi di Poisson per riscrivere il sistema canonico (9.5) nella forma

$$\begin{cases} \dot{q}_i = \{q_i, H\}, \\ \dot{p}_i = \{p_i, H\}, \end{cases} \quad i = 1, \dots, l. \tag{9.12}$$

Dal corollario 9.2.1 e dalle proprietà della parentesi di Poisson seguono alcune semplici conseguenze:

- ogni combinazione lineare di integrali primi è ancora un integrale primo;
- la parentesi di Poisson di due integrali primi è ancora un integrale primo (basta applicare l'identità di Jacobi).

Quest'ultima proprietà potrebbe sembrare una strada per “generare” nuovi integrali primi: in realtà i nuovi integrali primi che si ottengono in questo modo sono quasi sempre integrali che “dipendono funzionalmente” dagli integrali già noti.

L'antisimmetria permette di ottenere immediatamente una versione parziale in ambito hamiltoniano del Teorema di Noether. Infatti possiamo usare una qualsiasi funzione f per scrivere un “sistema di equazioni di Hamilton” di cui f sia la Hamiltoniana.

Lungo le soluzioni di questo sistema la derivata totale di una funzione g sarà data da $\{g, f\}$. Ma se f è un integrale primo del sistema di Hamiltoniana H sia ha $\{f, H\} = -\{H, f\} = 0$. Quindi la funzione di Hamilton H risulta costante lungo le traiettorie del flusso generato dal sistema di Hamiltoniana f , che quindi è un gruppo (locale) di simmetrie della Hamiltoniana H .

9.3 Derivazione variazionale delle equazioni di Hamilton

Come le equazioni di Lagrange, anche le equazioni di Hamilton possono essere derivate da un principio variazionale che non differisce di molto da quello che ci ha condotto alle equazioni di Lagrange.

Infatti il punto di partenza resta l'integrale di azione, ovvero l'integrale della Lagrangiana tra due tempi fissati. In questo caso però la Lagrangiana sarà espressa tramite la funzione di Hamilton e le variabili (\mathbf{q}, \mathbf{p}) , e l'azione prenderà la forma

$$\int_{t_1}^{t_2} \left[\sum_{k=1}^l p_k \dot{q}_k - \mathcal{H}(\mathbf{q}, \mathbf{p}) \right] dt. \quad (9.13)$$

Si noti che ora la funzione integranda contiene le derivate prime delle \mathbf{q} ma non le derivate delle \mathbf{p} . Di conseguenza **non** si chiederà alle variazioni di soddisfare delle condizioni agli estremi per le \mathbf{p} ma solo per le \mathbf{q} . Quindi il principio variazionale richiederà l'annullarsi della variazione dell'azione espressa dalla (9.13) con le condizioni $\mathbf{q}(t_1) = \mathbf{q}_1$ e $\mathbf{q}(t_2) = \mathbf{q}_2$, cioè

$$\begin{aligned} 0 &= \delta \int_{t_1}^{t_2} \left(\sum_{k=1}^l p_k \dot{q}_k - \mathcal{H}(\mathbf{q}, \mathbf{p}) \right) dt \\ &= \int_{t_1}^{t_2} \sum_{k=1}^l \left(p_k \delta \dot{q}_k + \delta p_k \dot{q}_k - \frac{\partial \mathcal{H}}{\partial q_k} \delta q_k - \frac{\partial \mathcal{H}}{\partial p_k} \delta p_k \right) dt \end{aligned}$$

Ora “scarichiamo” le derivate nei termini $\delta \dot{q}_k$ integrando per parti e, utilizzando il fatto che la variazione della \mathbf{q} è nulla agli estremi, otteniamo

$$\int_{t_1}^{t_2} \sum_{k=1}^l \left(-\dot{p}_k \delta q_k + \delta p_k \dot{q}_k - \frac{\partial \mathcal{H}}{\partial q_k} \delta q_k - \frac{\partial \mathcal{H}}{\partial p_k} \delta p_k \right) dt, \quad (9.14)$$

da cui, per l'arbitrarietà delle δq_k e δp_k , seguono le equazioni di Hamilton (9.5).

MANUALI

BIOMEDICA

- Branchi R., *Le impronte nel paziente totalmente edentulo*
- Branchi R., *Riabilitazione protesica del paziente oncologico testa-collo*
- Martinelli E., *Rieducazione del piede. Problematiche biomeccaniche e posturali statiche e dinamiche*
- Martinelli E., *Sviluppo del dolore rachideo in gravidanza. Mutamenti della biomeccanica rachidea, problematiche posturali, prevenzione e attività fisica adattata pre e post parto*
- Martinelli E., *Rieducazione posturale. Fondamenti per la progettazione della postura*
- Martinelli E., *Prevenzione del mal di schiena di origine meccanica con attività motoria e comportamentale. Approfondimenti di patomeccanica e biomeccanica rachidea*
- Rossetti R., *Manuale di batteriologia clinica. Dalla teoria alla pratica in laboratorio*
- Rucci L., *Testo Atlante di embriologia clinica della Laringe. La chirurgia conservativa compartimentale della regione glottica*
- Rucci L., *Clinical Embryology of Human Larynx for Conservative Compartmental Surgery. A Text and Atlas*

SCIENZE

- Bart J.C.J., *Polymer Additive Analytics. Industrial Practice and Case Studies*
- Caramelli D., *Antropologia molecolare. Manuale di base*
- Lo Nostro P., Peruzzi N., *Spontaneamente. Esercizi Risolti di Chimica Fisica*
- Ricci R., *Lezioni di Sistemi Dinamici*, a cura di Lorenzo Barletti, Angiolo Farina, Lorenzo Fusi, Federico Talamucci
- Scialpi A., Mengoni A. (a cura di), *La PCR e le sue varianti. Quaderno di laboratorio*
- Simonetta M.A., *Short history of Biology from the Origins to the 20th Century*
- Spinicci R., *Elementi di chimica* (nuova edizione)

SCIENZE SOCIALI

- Ciampi F., *Fondamenti di economia e gestione delle imprese*
- Giovannini P. (a cura di), *Teorie sociologiche alla prova*
- Maggino F., *L'analisi dei dati nell'indagine statistica. Volume 1. La realizzazione dell'indagine e l'analisi preliminare dei dati*
- Maggino F., *L'analisi dei dati nell'indagine statistica. Volume 2. L'esplorazione dei dati e la validazione dei risultati*

- Magliulo A., *Elementi di economia del turismo*
- Visentini L., Bertoldi, M., *Conoscere le organizzazioni. Una guida alle prospettive analitiche e alle pratiche gestionali*

SCIENZE TECNOLOGICHE

- Borri C., Pastò S., *Lezioni di ingegneria del vento*
- Borri C., Betti M., Marino E., *Lectures on Solid Mechanics*
- Cidronali A., Paolo Colantonio P. e Lucci L., *Antenne Integrate Attive: Quaderni del Dottorato di Ricerca in Ingegneria dell'Informazione dell'Università di Firenze*
- Gulli R., *Struttura e costruzione / Structure and Construction*
- Policicchio F., *Lineamenti di infrastrutture ferroviarie*

UMANISTICA

- Bertini F., *Risorse, conflitti, continenti e nazioni. Dalla rivoluzione industriale alle guerre irachene, dal Risorgimento alla conferma della Costituzione repubblicana*
- Bombardieri L., Graziadio G., Jasink A.M., *Preistoria e Protostoria egea e cipriota*
- Bombi A.S., Pinto G., Cannoni E., *Pictorial Assessment of Interpersonal Relationships (PAIR). An analytic system for understanding children's drawings*
- Baldini S., Marini D., Vorrei. *Corso di lingua italiana di livello elementare 1. Libro di testo e libro degli esercizi*
- Borello E., Mannori S., *Teoria e tecnica delle comunicazioni di massa*
- Brandi L., Salvadori B., *Dal suono alla parola. Percezione e produzione del linguaggio tra neurolinguistica e psicolinguistica*
- Coniglione F., Lenoci M., Mari G., Polizzi G. (a cura di), *Manuale di base di Storia della filosofia*
- Marcialis N., *Introduzione alla lingua paleoslava*
- Michelazzo F., *Nuovi itinerari alla scoperta del greco antico. Le strutture fondamentali della lingua greca: fonetica, morfologia, sintassi, semantica, pragmatica*
- Peruzzi A., *Il significato inesistente. Lezioni sulla semantica*
- Peruzzi A., *Modelli della spiegazione scientifica*
- Sandrini M.G., *Filosofia dei metodi induttivi e logica della ricerca*
- Trisciuzzi L., Zappaterra T., Bichi L., *Tenersi per mano. Disabilità e formazione del sé nell'autobiografia*

Quest'opera si basa sugli *Appunti per il corso di sistemi dinamici* scritti dal professor Riccardo Ricci nel 2005 e successivamente rielaborati fino all'anno della sua scomparsa, avvenuta nel 2013. Partendo dall'ultima versione, i curatori hanno intrapreso un'opera di rivisitazione e ampliamento che ha portato al presente manuale. Il testo si rivolge agli studenti triennali dei corsi di Matematica, Fisica e Ingegneria e intende presentare gli argomenti fondamentali della meccanica Lagrangiana, della dinamica dei corpi rigidi e dei principi variazionali, con qualche cenno alla meccanica Hamiltoniana.

LUIGI BARLETTI è professore associato di Fisica Matematica presso l'Università di Firenze, dove svolge attività di ricerca nell'ambito delle teorie cinetiche classiche e quantistiche con applicazioni ai dispositivi ottici ed elettronici.

ANGIOLO FARINA è ricercatore in Fisica Matematica presso l'Università di Firenze. Ha al suo attivo numerose pubblicazioni nei settori dei problemi a frontiera libera e della fluidodinamica.

LORENZO FUSI è ricercatore di Fisica Matematica presso l'Università di Firenze, dove svolge attività di ricerca nell'ambito fluidodinamica non-Newtoniana.

FEDERICO TALAMUCCI è ricercatore presso l'Università di Firenze. Dal 2000 svolge attività didattica per lezioni ed esercitazioni dei corsi di Meccanica nella Scuola di Scienze Matematiche, Fisiche e Naturali.