LUCIANO BATTAIA

MATEMATICA DI BASE

1 - *Richiami teorici*

www.batmath.it

Matematica di Base
1 - Richiami teorici

Luciano Battaia

http://www.batmath.it

Versione 1.1 del 21 giugno 2021

La matematica è una meravigliosa apparecchiatura spirituale
fatta per pensare in anticipo tutti i casi possibili.
Robert Musil

Indice

Elenco delle figure

Elenco delle tabelle

Premessa

Questo lavoro è rivolto agli studenti che si accingono ad affrontare un corso universitario nel quale la Matematica sia un requisito essenziale del processo formativo: Ingegneria, Informatica, Fisica, Matematica, ecc.

È ormai consuetudine che i candidati all'iscrizione ad uno di questi corsi di tipo tecnico-scientifico si cimentino in un test di matematica di base, contenente quesiti relativi agli argomenti matematici principali che sono stati oggetto del loro studio alle scuole secondarie superiori (calcolo differenziale ed integrale esclusi). Il primo scopo di questo test è quello di rendere anzitutto i ragazzi stessi consapevoli di eventuali lacune, o carenze culturali, di matematica elementare che, per essere tale, è nel contempo assolutamente necessaria e irrinunciabile.

In questo primo volume si sono voluti richiamare i concetti essenziali di questa *matematica elementare*, corredando i richiami teorici con una serie di esempi ed esercizi esplicativi. In ogni caso gli esercizi fanno parte integrante del testo e, spesso, molti concetti sono richiamati proprio nel corso della risoluzione degli stessi.

Il volume è diviso in due parti: nella prima sono proposti gli argomenti sviluppati nella quasi totalità dei corsi di scuola media superiore, nella seconda alcuni argomenti più specialistici.

Trattandosi di un testo di riepilogo di concetti in gran parte noti dagli studi precedenti, gli argomenti non sono proposti in ordine strettamente sequenziale e, spesso, si usano in alcuni capitoli concetti che saranno trattati in capitoli successivi. Questo in particolare succede negli esercizi, dove si fa quasi sempre uso non solo dei concetti trattati nel capitolo, ma anche di altre nozioni note dagli studi secondari. Inoltre non è escluso che alcuni argomenti siano ripresi più volte in diversi capitoli, magari sotto differenti angolature, oppure semplicemente per ragioni di completezza nel corso della trattazione di certi argomenti. In sostanza questo non è un testo adatto per una "prima lettura", quanto piuttosto un utile strumento per una accurata sistemazione dei concetti della matematica di base. Si segnala altresì che non sono proposte, salvo qualche eccezione, le dimostrazioni dei risultati via via enunciati, dimostrazioni che si possono trovare in qualunque buon testo di scuola secondaria superiore.

Questo volume ha, almeno nelle intenzioni, anche un altro scopo: troppo spesso gli studenti sono abituati, nel corso degli studi a livello di scuola media superiore, a utilizzare il libro di testo esclusivamente come fonte da cui prelevare gli esercizi via via assegnati dal docente. Così non si acquisisce quella che, invece, dovrebbe essere una abilità fondamentale e cioè la capacità di leggere un testo scientifico: la lezione di un docente non può mai essere sostitutiva dello studio di un libro di testo che, proprio per le sue caratteristiche, è sempre frutto di un lungo e paziente lavoro di scrittura e riscrittura, di raffinamento, di perfezionamento, di sistemazione. In effetti una delle difficoltà più evidenti che gli studenti incontrano ai primi approcci con il metodo di lavoro universitario è proprio quella di non sapersi orientare nella lettura e comprensione del libro di testo adottato o dei testi consigliati come approfondimento, e quasi tutti finiscono per studiare solo sugli appunti presi a lezione, magari dal compagno più bravo, senza tenere conto che quegli appunti possono contenere degli inevitabili strafalcioni: chi prende appunti non è quasi mai un esperto della materia! Proprio per abituare lo studente alla lettura dei testi che incontrerà nei successivi corsi universitari, il linguaggio e la forma utilizzata sono spesso più vicini a quelli di un testo specialistico che non a quelli di un testo propedeutico.

Questo manuale è frutto del lavoro di oltre quaranta anni di insegnamento nella scuola media superiore e nei corsi universitari di Analisi matematica, Geometria, Meccanica razionale, Fisica matematica, Matematica e statistica, Matematica generale, Istituzioni di analisi superiore, Matematica di base, delle università di Padova, Udine e Trieste, e tiene anche conto dell'esperienza maturata nella gestione del sito web `www.batmath.it`, e nei relativi rapporti con le migliaia di utenti dello stesso. Come ogni testo di matematica, anche questo non può essere esente da errori, imperfezioni, lacune: chiunque abbia qualcosa da segnalare è pregato di usare l'indirizzo di mail collegato al già citato sito web dell'autore.

Parte I.

Argomenti fondamentali

La prima parte del testo contiene gli argomenti di matematica di base sviluppati nella quasi totalità dei corsi di scuola secondaria superiore e sostanzialmente presenti nel Syllabus dell'UMI, anche se alcuni argomenti sono trattati in maniera più approfondita.

1. Elementi di logica. Insiemi

In questo capitolo saranno brevemente richiamati alcuni concetti di base di logica e teoria degli insiemi, anche con lo scopo di fissare con la dovuta precisione le notazioni utilizzate. Il concetto di insieme è alla base di qualunque trattazione matematica ed è quindi opportuno accertarsi di padroneggiare con sicurezza i contenuti qui proposti. In particolare gli elementi di logica delle proposizioni e dei predicati, seppur ridotti all'essenziale, sono di grande utilità nel rispondere ai quesiti sempre più frequentemente proposti nei test di ammissione alle facoltà universitarie.

Utilizzeremo fin da subito, specie negli esempi, gli insiemi dei numeri naturali ($\mathbb{N}$), interi ($\mathbb{Z}$), razionali ($\mathbb{Q}$) e reali ($\mathbb{R}$), le cui proprietà essenziali saranno richiamate in seguito.

1.1. Logica proposizionale

La frase "*La neve è bianca*" esprime un fatto ritenuto da tutti *vero*, anzi universalmente vero. La frase "*La terra è una stella*" esprime invece un fatto ritenuto da tutti *falso*, anzi universalmente falso. La frase "*Roma è una bella città*" esprime un fatto che può essere ritenuto vero da certi individui e falso da altri. Alle frasi "*Non disturbare mentre faccio lezione*", "*Vai a comperare il pane*", "*Se lancio un dado esce il sei*", "*Domani pioverà*", non può essere attribuito un valore di verità o falsità.

Questi esempi mostrano che alcune frasi, o proposizioni, della lingua italiana (ma lo stesso succede in tutte le lingue) assumono uno ed uno solo tra i valori *vero* e *falso*, in altri casi o non c'è accordo sull'attribuzione di un valore di verità o falsità, oppure non ha proprio senso tale attribuzione.

Esistono anche esempi più complessi, come l'affermazione "*Tutti i numeri naturali pari maggiori di* 2 *sono somma di due numeri primi*". Ebbene, a tutt'oggi (2012), non è possibile sapere se tale affermazione sia *vera* o *falsa*[(1)], benchè non si sia trovato alcun caso in cui essa non è verificata. Chi è interessato a questo problema è invitato a leggere il bellissimo thriller matematico *Zio Petros e la Congettura di Goldbach*, di Apostolos Doxiadis, Bompiani, 2007.

1.1.1. Definizioni

Tenendo conto di queste osservazioni, daremo ora una definizione di *enunciato*, o *proposizione*, segnalando comunque che il concetto di *verità* è estremamente delicato e un'analisi del problema esula dagli scopi di questa trattazione.

[1]Si tratta della famosa *Congettura di Goldbach*, proposta sostanzialmente da Christian Goldbach nel 1742. Per esempio si ha

1. $4 = 2 + 2$
2. $6 = 3 + 3$
3. $8 = 3 + 5$
4. $10 = 3 + 7 = 5 + 5$
5. $\ldots$

Definizione 1.1 (Proposizione). *Si chiama* proposizione *o* enunciato *ogni affermazione che assume uno e un solo dei due valori:* vero *o* falso.

Si noti che è implicito nella definizione data il fatto che ammettiamo che la logica che utilizziamo sia *bivalente*, cioè preveda che le espressioni di cui ci occupiamo possano avere uno solo dei due valori di verità "vero" o "falso".

Gli enunciati possono essere costituiti da una sola affermazione, come negli esempi che abbiamo proposto sopra, e li chiameremo *enunciati atomici*, oppure possono essere costituiti da più affermazioni, collegate tra di loro. Un esempio è costituito dall'enunciato "*Il sole è una stella* e *la terra è un pianeta*", che si può considerare composto da due enunciati atomici (entrambi veri) connessi dalla parola "e". Un altro modo per costruire nuovi enunciati è quello di usare la negazione "non". Per esempio "*La terra* non *è una stella*" è ottenuto dalla *negazione* dell'enunciato (falso) "*La terra è una stella*".

Si chiamano *connettivi* le parole (come la "e" dell'esempio) che collegano tra di loro due enunciati, oppure che operano su un enunciato (come il "non" dell'esempio) per ottenere un nuovo enunciato. A volte il "non" è chiamato un *operatore* invece che un connettivo, in quanto in realtà non connette due enunciati, ma agisce, "opera", su un singolo enunciato.

Si deve notare che i connettivi collegano tra di loro due enunciati senza alcun riguardo al significato che questi possono assumere; per esempio è perfettamente legittimo l'enunciato "*Parigi è la capitale del Brasile* o $2+2$ *vale* 4", che è la connessione, tramite la parola "o", di due enunciati (uno falso e uno vero). L'unica cosa che conta è il valore di verità complessivo dell'enunciato risultante.

Poiché nel linguaggio comune le parole non hanno sempre un senso univoco, in logica al posto delle parole si utilizzano dei simboli speciali per formalizzare in maniera rigorosa i connettivi e si costruiscono delle *tavole di verità* che stabiliscono le regole che permettono di dedurre la verità o meno di un enunciato composto, una volta che sia noto il valore di verità degli enunciati componenti: queste tavole di verità possono essere pensate come delle vere e proprie *definizioni* dei connettivi stessi.

1.1.2. Connettivi logici

Nel seguito indicheremo le proposizioni con simboli come $\mathscr{P}$, $\mathscr{Q}$, ... I connettivi che ci interesseranno sono i seguenti:

- non, oppure $\neg$, *negazione*: $non\mathscr{P}$ (oppure $\neg\mathscr{P}$) è vera, se $\mathscr{P}$ è falsa mentre è falsa, se $\mathscr{P}$ è vera;
- $\wedge$, "et", oppure "e", *congiunzione*: $\mathscr{P} \wedge \mathscr{Q}$ è vera se tutte due le proposizioni sono vere, altrimenti è falsa;
- $\vee$, "vel", oppure "o", *disgiunzione*: $\mathscr{P} \vee \mathscr{Q}$ è vera se almeno una delle due proposizioni è vera, altrimenti è falsa;
- $\Rightarrow$, "implica", *implicazione*: $\mathscr{P} \Rightarrow \mathscr{Q}$ è falsa solo quando $\mathscr{P}$ è vera e $\mathscr{Q}$ è falsa, in particolare da una proposizione falsa si può dedurre qualsiasi cosa;
- $\Leftrightarrow$, "se e solo se", "condizione necessaria e sufficiente", *equivalenza*: $\mathscr{P} \Leftrightarrow \mathscr{Q}$ è vera se $\mathscr{P}$ e $\mathscr{Q}$ sono entrambe vere o entrambe false.

La tabella 1.1 (dove "V" indica *vero* e "F" indica *falso*) riassume in maniera formale le definizioni dei connettivi. Si noti che la tabella 1.1 è costruita tenendo conto che ciascuno dei due enunciati atomici ha due possibili valori di verità, e che quindi per esaminare il valore di verità di un enunciato che li coinvolga entrambi devo esaminare tutte le situazioni che si possono presentare (si tratta di un caso molto elementare di disposizioni con ripetizione, in particolare delle disposizioni con ripetizione di due oggetti a due a due, ovvero $2^2 = 4$ casi). Per il solo connettivo “non” basterebbero evidentemente due sole righe nella tabella, in quanto in questo caso è coinvolto un solo enunciato atomico. Tabelle di questo tipo si chiamano *Tavole di verità*.

Tabella 1.1.: *Connettivi logici e relativa tavola di verità*

$\mathscr{P}$	$\mathscr{Q}$	$\neg\mathscr{P}$	$\mathscr{P}\wedge\mathscr{Q}$	$\mathscr{P}\vee\mathscr{Q}$	$\mathscr{P}\Rightarrow\mathscr{Q}$	$\mathscr{P}\Leftrightarrow\mathscr{Q}$
V	V	F	V	V	V	V
V	F	F	F	V	F	F
F	V	V	F	V	V	F
F	F	V	F	F	V	V

Il connettivo $\Rightarrow$ ha molta importanza in matematica. Dimostrare un teorema significa infatti dimostrare la verità di $\mathscr{P} \Rightarrow \mathscr{Q}$, sapendo che $\mathscr{P}$ è vera: $\mathscr{P}$ è detta *ipotesi* e $\mathscr{Q}$ è detta *tesi*.

Nel costruire enunciati complessi, che utilizzino ripetutamente i connettivi indicati, bisogna prestare attenzione all'ordine in cui le “operazioni” sugli enunciati vengono eseguite, ed eventualmente usare le parentesi, come nelle normali espressioni matematiche.

Vediamo un esempio in cui siano coinvolti solo due enunciati, per evitare di dover costruire una tabella troppo ampia: dati due enunciati $\mathscr{P}$ e $\mathscr{Q}$, costruire la tavola di verità dell'enunciato $\neg\big[(\mathscr{P}\wedge\mathscr{Q})\vee(\mathscr{P}\wedge\neg\mathscr{Q})\big]$. Si veda la tabella 1.2.

Tabella 1.2.: *Tavola di verità della proposizione* $\neg\big[(\mathscr{P}\wedge\mathscr{Q})\vee(\mathscr{P}\wedge\neg\mathscr{Q})\big]$

$\mathscr{P}$	$\mathscr{Q}$	$\mathscr{P}\wedge\mathscr{Q}$	$\neg\mathscr{Q}$	$\mathscr{P}\wedge\neg\mathscr{Q}$	$(\mathscr{P}\wedge\mathscr{Q})\vee(\mathscr{P}\wedge\neg\mathscr{Q})$	$\neg\big[(\mathscr{P}\wedge\mathscr{Q})\vee(\mathscr{P}\wedge\neg\mathscr{Q})\big]$
V	V	V	F	F	V	F
V	F	F	V	V	V	F
F	V	F	F	F	F	V
F	F	F	V	F	F	V

È abbastanza frequente l'uso delle seguenti due formule che stabiliscono i legami tra negazione, unione e intersezione e sono chiamate *leggi di de Morgan*:

$$\neg(\mathscr{P}\vee\mathscr{Q}) = (\neg\mathscr{P})\wedge(\neg\mathscr{Q}) \quad \text{e} \quad \neg(\mathscr{P}\wedge\mathscr{Q}) = (\neg\mathscr{P})\vee(\neg\mathscr{Q}), \tag{1.1}$$

formule che possono essere verificate facilmente costruendo le relative tavole di verità.

1.2. Logica dei predicati

Come abbiamo detto, il senso di una proposizione sta nel poter stabilire se è vera o se è falsa. Un'affermazione del tipo $x < -2$ non è una proposizione, perchè il suo valore di verità dipende da x. Facendo variare x in un opportuno insieme (che deve essere precisato) si possono ottenere proposizioni vere o proposizioni false. Possiamo dire che si tratta di un proposizione dipendente da x, e indicarla con $\mathscr{P}(x)$:

x sia chiama una *variabile* e $\mathscr{P}(x)$ un *predicato*. Naturalmente si possono avere predicati che dipendono da più variabili, per esempio $x+y>0$, e in questo caso i predicati sono anche chiamati *relazioni*.

Come abbiamo già osservato è indispensabile precisare in quale ambiente si deve scegliere la variabile (o le variabili) di un predicato. Per esempio l'affermazione "*x è pari*" ha senso se x è un numero naturale, non ha alcun senso se x è una frazione.

Fissato uno dei possibili valori di x, diciamolo x_0, il predicato diventa una proposizione (che sarà vera o falsa a seconda del valore di x_0), proposizione che si indica con $\mathscr{P}(x_0)$.

1.2.1. Quantificatori

Nella costruzione dei predicati si usano comunemente espressioni del tipo

— *esiste (almeno) un x tale che valga* $\mathscr{P}(x)$;
— *per ogni x è verificato* $\mathscr{P}(x)$.

Per formalizzare queste frasi si usano due simboli logici, detti *quantificatori*

— $\exists$, "esiste (almeno) un", *quantificatore esistenziale*;
— $\forall$, "per ogni", *quantificatore universale*.

Si usa anche spesso il simbolo $\exists!$, oppure $\exists^1$ col significato: *esiste uno e uno solo*.

Nel caso di uso contemporaneo di più quantificatori si deve prestare particolare attenzione all'ordine con cui sono scritti. Un esempio chiarirà il senso di questa affermazione.

Consideriamo il predicato $\mathscr{P}(x,y)=$ "x è uno studente in grado di risolvere il problema y". Allora

$$\forall y\ \exists x \text{ tale che } \mathscr{P}(x,y)$$

significa: "qualunque sia il problema y c'è uno studente in grado di risolverlo". Invece

$$\exists x\ \forall y \text{ tale che } \mathscr{P}(x,y)$$

significa: "c'è uno studente in grado di risolvere qualsiasi problema". Evidentemente si tratta di due situazioni radicalmente diverse.

Osservazione 1.2. È opportuno rendersi conto, su un esempio classico, di come la simbologia comunemente usata in matematica possa facilmente dar luogo a equivoci, senza un'effettiva conoscenza delle relazioni tra i connettivi logici.

Consideriamo l'equazione

$$x^2=1\,,$$

le cui soluzioni si trovano scritte usualmente nella forma

$$x=\pm 1\,,$$

ove si intende che sia il numero 1 che il numero -1 soddisfano l'equazione (in termini logici: rendono vero, nell'insieme dei numeri reali, il predicato "$x^2=1$"). Questo risultato andrebbe, più correttamente, espresso nella forma

$$x=1 \lor x=-1\,.$$

Consideriamo ora la scrittura

$$x^2 \neq 1$$

la cui “soluzione” è usualmente scritta nella forma

$$x \neq \pm 1.$$

Ebbene, questa scrittura *non* deve essere tradotta in $x \neq 1 \vee x \neq -1$, che porterebbe alla conclusione che $x^2 \neq 1$ è verificata da ogni numero reale; la traduzione logica corretta è, invece,

$$x \neq 1 \wedge x \neq -1$$

in quanto quello che si intende scrivendo $x \neq \pm 1$ è proprio il *contemporaneo* verificarsi delle due condizioni su x.

Si può notare che $x^2 \neq 1$ equivale a $\neg(x^2 = 1)$ che porta a $\neg(x = 1 \vee x = -1)$ ovvero a $\neg(x = 1) \wedge \neg(x = -1)$, che viene abitualmente scritta $x \neq 1 \wedge x \neq -1$: per provare questo fatto in maniera formale si può dimostrare, usando le tavole di verità, che $\neg(\mathscr{P} \vee \mathscr{Q})$ e $(\neg\mathscr{P}) \wedge (\neg\mathscr{Q})$ hanno gli stessi valori di verità (si tratta di una delle già citate *Leggi di De Morgan*).

Può essere utile osservare che i due quantificatori introdotti sono tra di loro strettamente collegati: ciascuno dei due può essere considerato un'abbreviazione di una espressione più complessa coinvolgente l'altro e l'operatore di negazione. È in effetti abbastanza evidente che dire

esiste almeno un x tale che valga $\mathscr{P}(x)$

equivale a dire

non per ogni x $\mathscr{P}(x)$ *non vale*,

ovvero, in formule,

$$\exists x \textit{ tale che } \mathscr{P}(x) \quad \text{equivale a} \quad \neg\big(\forall x \textit{ vale } \neg\mathscr{P}(x)\big).$$

Analogamente, dire

per ogni x vale $\mathscr{P}(x)$

equivale a dire

non esiste alcun x per cui $\mathscr{P}(x)$ *non vale*,

ovvero, in formule,

$$\forall x \textit{ vale} \mathscr{P}(x) \quad \text{equivale a} \quad \neg\big(\exists x \textit{ tale che } \neg\mathscr{P}(x)\big).$$

Ad esempio, invece degli enunciati

“*Esiste almeno un triangolo rettangolo*”,
“*Tutti gli italiani sono europei*”,

possiamo usare i seguenti

“*Non tutti i triangoli sono non rettangoli*”, "*Non esiste alcun italiano che non sia europeo*".

Questi stessi esempi mostrano, tuttavia, che la prima forma (quella senza la negazione) è più facilmente comprensibile della seconda.

È altresì opportuno notare come si costruisce la negazione di una frase con quantificatore:

— la negazione di: $\exists x$ *tale che valga* $\mathscr{P}(x)$ è: $\forall x$ *vale* $\neg\mathscr{P}(x)$;
— la negazione di: $\forall x$ *vale* $\mathscr{P}(x)$ è: $\exists x$ *tale che valga* $\neg\mathscr{P}(x)$.

1.3. Altri simboli di base

Tra i molti simboli che si usano nella pratica matematica ne richiamiamo qui due, per la loro importanza.

Se dobbiamo scrivere la somma dei numeri 1, 2, 3, possiamo tranquillamente scrivere $1+2+3$, ma se dobbiamo scrivere la somma dei numeri da 1 a 100[2], la scrittura esplicita diventa oltremodo pesante. Si potrebbe pensare di ovviare con l'uso dei puntini di sospensione:

$$1+2+\cdots+99+100\,.$$

La cosa però non è esente da critiche e, soprattutto, non è sempre praticabile. Per questo si introduce il cosiddetto *simbolo di sommatoria*, col quale la somma precedente si scrive

$$\sum_{i=1}^{100} i\,,$$

che traduce in forma compatta esattamente quello che si deve fare: sommare i numeri naturali, rappresentati genericamente dalla "variabile" i, partendo dal numero 1 e arrivando fino al numero 100. Il fatto che i numeri da sommare siano solo naturali viene sottinteso: non avrebbe alcun senso infatti sommare numeri razionali (o, peggio ancora, reali) partendo da 1 e arrivando fino a 100.

In generale gli addendi di una somma possono essere più complessi, per esempio:

— i reciproci dei numeri naturali: $1/i$,
— i quadrati dei numeri naturali: i^2,
— un'espressione qualunque coinvolgente i numeri naturali, come il rapporto tra un naturale e il suo successivo: $i/(i+1)$.

Se indichiamo con $a(i)$, o a_i, l'espressione coinvolgente il numero naturale i, la scrittura

$$\sum_{i=m}^{n} a_i \tag{1.2}$$

[2]Un aneddoto, abbastanza verosimile, relativo al grande matematico tedesco Carl Friedrich Gauss (1777-1855) racconta che all'età di otto-nove anni il maestro, per metterlo a tacere per un bel po', gli ordinò di sommare i numeri da 1 a 100: in brevissimo tempo Gauss fornì la risposta $50\times 101 = 5050$, sorprendendo anche il maestro che aveva sottovalutato l'intelligenza del suo allievo…

indicherà la somma di tante "copie" di quell'espressione, dove al posto di i si devono mettere, successivamente, tutti i numeri naturali dal valore iniziale m al valore finale n. Proponiamo alcuni esempi per chiarire ancora meglio il senso di quanto detto[(3)].

- $\displaystyle\sum_{i=5}^{10}\frac{1}{i^2}=\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\frac{1}{8^2}+\frac{1}{9^2}+\frac{1}{10^2}$;
- $\displaystyle\sum_{i=2}^{100}\frac{i}{i-1}=\frac{2}{2-1}+\frac{3}{3-1}+\cdots+\frac{99}{99-1}+\frac{100}{100-1}$;
- $\displaystyle\sum_{i=0}^{5}(-1)^i=(-1)^0+(-1)^1+(-1)^2+(-1)^3+(-1)^4+(-1)^5=1-1+1-1+1-1\,(=0)$

È opportuno osservare che al posto di i (che si chiama *indice della sommatoria*) si può usare una qualunque altra lettera: le scritture

$$\sum_{i=m}^{n}a_i \quad , \quad \sum_{j=m}^{n}a_j \quad \text{e} \quad \sum_{k=m}^{n}a_k$$

sono del tutto equivalenti (naturalmente purché i valori iniziale e finale restino gli stessi e le espressioni che coinvolgono numeri naturali siano identiche): per questo motivo l'indice i è spesso detto una *variabile muta*.

Giova anche ricordare che, trattandosi di somme, si possono applicare le usuali proprietà, in particolare ci interessa segnalare quella associativa. Si vedano gli esempi che seguono.

- $\displaystyle\sum_{i=2}^{100}\frac{2i+4}{i-1}=2\sum_{i=2}^{100}\frac{i+2}{i-1}$;
- $\displaystyle\sum_{i=0}^{20}\frac{(-1)^i}{i}=(-1)\sum_{i=0}^{20}\frac{(-1)^{i-1}}{i}$.

In maniera perfettamente analoga si possono considerare prodotti al posto di somme: si parla di simbolo di *produttoria*, anziché di sommatoria e si usa la scrittura:

$$\prod_{i=m}^{n}a_i\,, \tag{1.3}$$

ove il significato dei simboli è ormai evidente.

Sommatorie e produttorie possono combinarsi tra di loro e ripetersi nella stessa espressione. In casi del genere è sufficiente applicare *ordinatamente* le definizioni che abbiamo dato.

[3] Alcuni scrivono lo stesso simbolo disponendo gli estremi in maniera leggermente diversa:

$$\sum_{n}^{m}{}_{i}\,a_i\,;$$

si tratta praticamente solo di una questione di gusto, e nulla cambia ovviamente per quanto riguarda il significato.

Esempio 1.1.

$$
\begin{aligned}
\sum_{i=1}^{5}\Big(\sum_{j=3}^{6}(i-2j)\Big) &= \sum_{i=1}^{5}\Big((i-2\times 3)+(i-2\times 4)+(i-2\times 5)+(i-2\times 6)\Big)=\\
&=\sum_{i=1}^{5}\Big((i-6)+(i-8)+(i-10)+(i-12)\Big)=\\
&=\sum_{i=1}^{5}(4i-36)=4\sum_{i=1}^{5}(i-9)=\\
&=4\Big((1-9)+(2-9)+(3-9)+(4-9)+(5-9)\Big)=\\
&=4(-8-7-6-5-4)=-120
\end{aligned}
$$

1.4. Assiomi, teoremi, dimostrazioni, ...

Prima di affrontare con dimestichezza lo studio di una qualunque parte della matematica è opportuno chiarire il significato di termini di uso quotidiano, come assiomi, teoremi, ecc. Ci accontenteremo, naturalmente, di una introduzione semplice e schematica di questi concetti: la sola comprensione accurata di che cosa si intenda con teoria assiomatica richiederebbe interi volumi...

Concetti primitivi. Termini non definiti

Proviamo ad eseguire alcuni esperimenti sulle parole utilizzando un dizionario di italiano. L'idea è questa: partiamo da una parola, ne leggiamo la definizione e cerchiamo sul vocabolario una delle parole usate nella definizione; procediamo così fin che non ritroviamo una delle parole già usate: a questo punto siamo in un circolo vizioso. Abbiamo fatto alcuni esperimenti con vari dizionari e ottenuto più o meno gli stessi risultati. Ne riportiamo alcuni, dall'edizione del 1990 del Grande Dizionario De Agostini della Lingua Italiana. Abbiamo cercato parole attinenti la geometria, ma le conclusioni che trarremo sarebbero sostanzialmente valide con tutte le parole.

— *Punto*: in geometria, ente elementare intuitivo, privo di dimensioni.
— *Dimensione*: ciascuna delle misure di cui ci si vale per determinare l'estensione di una superficie o di un corpo nelle varie direzioni.
— *Estensione*: spazio, superficie.
— *Spazio*: l'estensione...
— Stop: circolo vizioso!

— *Retta*: ente geometrico definibile intuitivamente come la più corta delle linee che congiungono due punti e che si prolungano all'infinito nei due sensi.
— *Linea*: in geometria, concetto intuitivo che l'immaginazione raffigura come la traiettoria tracciata da un punto mobile.
— *Traiettoria*: la linea immaginaria percorsa da un oggetto che si muove nello spazio.
— *Spazio*: l'estensione vuota e illimitata nella quale si immaginano immersi gli enti geometrici solidi e nella quale sono collocati gli oggetti reali.
— *Estensione*: spazio, ...

— Stop: circolo vizioso!

Per dirla con Samuel Coxeter (vedi [3]): "Un circolo vizioso illustra l'importante principio che qualunque definizione di una parola richiede necessariamente altre parole, le quali richiedono ulteriori definizioni. Il solo modo di evitare un circolo vizioso è di pensare certi concetti primitivi come cose così semplici e ovvie che concordiamo nel non definirle."

Dunque in ogni teoria dobbiamo assumere alcuni concetti come *primitivi* o come *termini non definiti*: si tratta di parole sul cui significato non si dà alcuna spiegazione. Nella geometria euclidea, per esempio, tali termini sono: *punto*, *retta*, *piano*, *giace su*, *tra*, *congruente*. Nella teoria (ingenua) degli insiemi tali termini sono: *insieme*, *appartiene*.

Assiomi

Sono proposizioni che vengono introdotte senza dimostrazione e che si riferiscono a proprietà dei termini non definiti. Sono proprio gli assiomi che ci permettono di farci un'idea (anche fisica) di che che cosa siano gli enti primitivi: noi non ci chiediamo che cosa siano gli enti primitivi, ci basta solo sapere che tra di essi intercorrono certe relazioni e caratteristiche che noi stessi soggettivamente attribuiamo loro. In sostanza gli assiomi delimitano e caratterizzano gli enti primitivi, togliendo loro buona parte di quella arbitrarietà che sembrava assoluta quando li abbiamo introdotti solo come "parole". L'unica cosa di cui dovremo preoccuparci è di controllare che gli assiomi stessi non si contraddicano l'uno con l'altro. Purtroppo questo controllo non è per niente semplice e può richiedere lunghi e complessi dibattiti. L'esempio più clamoroso è l'assioma delle parallele di Euclide, di cui tutti abbiamo sentito parlare.

Teoremi

I teoremi sono proposizioni dedotte, mediante ragionamento, dagli assiomi o da altri teoremi già dimostrati, utilizzando gli enti primitivi ed eventualmente altri oggetti via via introdotti.

Per capire la situazione consideriamo qualche esempio:

— Se scaldo la cera di una candela essa fonde.
— Se un triangolo ha due lati uguali ha anche due angoli uguali.

In queste frasi si può individuare il seguente schema: se si verifica una prima condizione, allora consequenzialmente se ne deve verificare una seconda. Questo tipo di proposizioni si chiamano *implicazioni logiche*. La prima condizione prende il nome di *ipotesi*, la seconda di *tesi*. La tesi è dunque una conseguenza dell'ipotesi e questa dipendenza può essere verificata sperimentalmente (come nel caso della cera scaldata) o mediante un preciso ragionamento, detto *dimostrazione*. Questo tipo di implicazioni logiche da provare mediante un opportuno ragionamento vengono dette teoremi. La proposizione che si intende dimostrare viene detta *enunciato*.

In un teorema si distinguono dunque:

— L'enunciato, cioè l'affermazione da provare, contenente:
  - l'ipotesi, ovvero la condizione di partenza;
  - la tesi, ovvero la condizione che deve essere provata.
— La dimostrazione, cioè l'insieme dei ragionamenti, basati sugli assiomi o su teoremi precedenti, che ci permettono di dedurre logicamente la tesi dall'ipotesi.

È molto importante segnalare che l'idea che sta alla base del concetto di dimostrazione è che un qualsiasi studioso deve poter essere in grado di seguire e rifare tutti i ragionamenti utilizzati, se conosce i "precedenti" (cioè quello che è già stato assunto o dimostrato). Ovvero: la dimostrazione è un ragionamento mediante il quale un matematico può convincere un altro matematico, che la *legga*, della verità di una affermazione. Naturalmente questo avviene solo in linea di principio: man mano che la teoria diventa più complessa, tali diventano anche le dimostrazioni. Citiamo qui, a titolo di esempio, il famoso *teorema enorme* (un teorema di classificazione dei gruppi semplici): la dimostrazione originale occupa circa quindicimila pagine, sparse in oltre cinquecento articoli di riviste di matematica e ha richiesto il contributo di un centinaio di matematici e circa quarant'anni di lavoro. È altamente improbabile che una dimostrazione del genere possa essere rifatta da una sola persona, seguendo rigorosamente il metodo del ragionamento ipotetico-deduttivo; nonostante ciò questa ha diritto di chiamarsi dimostrazione.

Condizione necessaria e/o sufficiente

I teoremi possono esprimere *condizioni necessarie*, *condizioni sufficienti* o *condizioni necessarie e sufficienti*: vediamo di capire le differenze.

Condizione necessaria. A è *condizione necessaria* per B se

$$B \land (\neg A) \text{ è falsa, cioè è una contraddizione.}$$

La definizione formale appena data può anche essere espressa a parole, in vari modi:

- A è *condizione necessaria* per B se B implica A, cioè se in presenza di B è presente sempre A.
- A è *condizione necessaria* per B se mancando A, B non può essere presente (attenzione però: la presenza di A non permette, in generale, di assicurare la presenza di B).

In sostanza si può dire che quando A è condizione necessaria per B, in mancanza di A è inutile andare a controllare la presenza di B, in quanto B è assente sicuramente.

Condizione sufficiente. A è *condizione sufficiente* per B se

$$A \land (\neg B) \text{ è falsa, cioè è una contraddizione.}$$

La definizione formale appena data può anche essere espressa a parole, in vari modi:

- A è *condizione sufficiente* per B se A implica B, cioè se in presenza di A è sempre presente B.
- A è *condizione sufficiente* per B se mancando B, A non può essere presente (attenzione però: la mancanza di A non permette, in generale, di escludere la presenza di B).

In sostanza si può dire che quando A è condizione sufficiente per B, in presenza di B è utile andare a controllare la presenza di A, in quanto A non è detto che sia presente.

Condizione necessaria e sufficiente. A è condizione *necessaria e sufficiente* per B se A implica B e contemporaneamente B implica A, ovvero se A e B sono equivalenti.

Si vedano i seguenti esempi, volutamente elementari.

- *Condizione necessaria perché un quadrilatero sia un quadrato è che abbia quattro angoli retti.* Se in un quadrilatero almeno uno dei quattro angoli non è retto, il quadrilatero non può essere un quadrato; però anche se i quattro angoli sono retti non è detto che il quadrilatero sia un quadrato.

— *Condizione sufficiente perché un numero sia pari è che termini con la cifra* 2. È chiaro che se un numero ha come ultima cifra 2 è pari, ma può essere pari anche se non ha come ultima cifra 2.
— *Condizione necessaria e sufficiente perché un triangolo sia isoscele è che abbia due angoli congruenti.* Se un triangolo ha due angoli congruenti allora è isoscele (cioè ha due lati congruenti); viceversa se è isoscele ha due angoli congruenti. In sostanza nella definizione di triangolo isoscele si può sostituire la proprietà di "avere due lati congruenti" con quella di "avere due angoli congruenti".

Dimostrazioni per assurdo

Le dimostrazioni per assurdo sono molto frequenti in matematica ed è bene avere chiaro come si procede e su che basi poggia questo tipo di dimostrazione. Come già detto, dimostrare un teorema significa provare, a partire dall'ipotesi e utilizzando altre proprietà note o dimostrate, che una certa affermazione, detta tesi, è vera. Poiché nella nostra logica la tesi può essere solo vera o falsa, se proviamo che non può essere falsa, avremo provato che è vera. Per dimostrare che la tesi non può essere falsa si assume vero il contrario della tesi, cioè si nega la tesi, e si dimostra che questo porta a una contraddizione, dopodiché si conclude che la tesi è vera.

In questo testo si trovano parecchie dimostrazioni per assurdo, anche negli esercizi: anche se di norma non proponiamo le dimostrazioni dei teoremi via via enunciati, spesso lo faremo proprio per le dimostrazioni per assurdo, considerata la loro importanza e le difficoltà che normalmente si incontrano.

Lemmi

Alcune volte la dimostrazione di un teorema può essere estremamente complessa e può richiedere un grande numero di implicazioni logiche intermedie. In questi casi si usa suddividere il teorema in parti, anteponendo una o più proposizioni, dette *lemmi*, la cui tesi viene utilizzata quasi esclusivamente nel teorema stesso.

Corollari

Alcune volte succede che da un determinato teorema ne seguano subito alcuni altri, la cui dimostrazione è talmente immediata da poter essere appena accennata. Questi teoremi sono detti *corollari*. Si noti che in molti casi le tesi espresse nei corollari possono essere più importanti di quelle espresse nel teorema principale.

Definizioni

Definire significa descrivere, con parole che si suppongono note, un oggetto. In sostanza si può dire che le definizioni servono a sostituire espressioni complesse con un'unica parola. Le definizioni servono a introdurre nuovi concetti, dopo quelli primitivi.

Bisogna prestare attenzione a capire la differenza che c'è tra una definizione e un ente primitivo: la definizione utilizza gli enti primitivi o teoremi già dimostrati, gli enti primitivi sono oggetti su cui non si fa alcuna affermazione.

Una chiarificazione completa del concetto di definizione e del suo uso in matematica sarebbe in realtà molto complessa, e ci porterebbe troppo lontano.

1.5. Insiemi

Assumiamo la nozione di *insieme* come primitiva, fidandoci della nostra intuizione. Volendo si potrebbero usare delle circonlocuzioni, del tipo "un insieme è una *collezione* di oggetti, detti *elementi*", ma in realtà non avremmo detto nulla di significativo: è come dire "un insieme è un insieme". Abitualmente, ma non sempre, indicheremo gli insiemi con le lettere maiuscole corsive: A, B, ...

La scrittura

$$x \in A \tag{1.4}$$

sta ad indicare che l'oggetto x è un elemento dell'insieme A e si legge "x appartiene ad A". La (1.4) si può scrivere anche $A \ni x$. La negazione della (1.4) si scrive

$$x \notin A, \tag{1.5}$$

che si legge, naturalmente, "x non appartiene ad A". La (1.5) si può scrivere anche $A \not\ni x$.

Due insiemi sono uguali se e solo se hanno gli stessi elementi. Questo si può scrivere, usando i simboli logici sopra introdotti,

$$A = B \Leftrightarrow (\forall x \; x \in A \Leftrightarrow x \in B). \tag{1.6}$$

È conveniente introdurre uno speciale insieme, detto *insieme vuoto* e indicato con $\emptyset$, privo di elementi. Poiché due insiemi possono essere diversi se e solo differiscono per qualche loro elemento, dovremo ritenere che di insiemi vuoti ce ne sia uno solo.

Per assegnare un insieme possiamo usare due metodi.

1. Rappresentazione estensiva: consiste nell'elencare tutti gli elementi dell'insieme, per esempio $A = \left\{ 0, \pi, \sqrt{2}, \text{Roma} \right\}$.
2. Rappresentazione intensiva: consiste nell'assegnare gli elementi indicando una proprietà che li contraddistingue, per esempio $A = \{ x \mid x \text{ è un numero naturale pari} \}$. La proprietà caratteristica non è altro che un predicato $\mathscr{P}(x)$, e l'insieme contiene i valori di x per cui il predicato è vero.

La seconda possibilità è soprattutto indicata per insiemi che contengano infiniti elementi e in particolare per sottoinsiemi di altri insiemi. Anche gli insiemi infiniti però potranno, se non sono possibili equivoci, essere descritti per elencazione. Potremo, a volte, scrivere $A = \{ 3, 6, 9, 12, \dots \}$ per indicare l'insieme dei numeri naturali multipli di 3, ma occorre prestare la massima attenzione. Per esempio se scrivessimo

$$A = \{ 2, 3, \dots \}$$

non sarebbe assolutamente possibile dedurre se intendiamo riferirci ai numeri naturali maggiori o uguali a 2, oppure ai numeri primi.

Occorre poi prestare ancora più attenzione alle rappresentazioni intensive. Esse non presentano problemi quando si usano per assegnare sottoinsiemi di un insieme noto[4], possono invece produrre gravi difficoltà logiche in altri casi, come mostra il famoso *Paradosso del barbiere*[5]: se il barbiere è definito come colui che fa la barba a chi non se la fa da solo, non è possibile stabilire se il barbiere si faccia o no la barba.

[4] È la stessa osservazione fatta a proposito dei predicati, quando abbiamo detto che si deve precisare in quale ambiente varia la x che compare nel predicato.

[5] In internet si possono trovare migliaia di pagine che trattano questo problema.

È da segnalare il fatto che, se per assegnare un insieme dobbiamo necessariamente avere un criterio per decidere quali sono i suoi elementi, a volte la verifica esplicita se un elemento sta o no in un insieme può essere estremamente complessa. L'esempio classico di questa situazione è quello dell'insieme, $\mathbb{P}$, dei numeri primi. Mentre è immediato che, per esempio, $31 \in \mathbb{P}$, è molto più difficile verificare che anche $15\,485\,863 \in \mathbb{P}$, e per verificare che $2^{43\,112\,609} - 1 \in \mathbb{P}$ (uno dei più grandi primi conosciuti a metà del 2012, con ben 12 978 189 cifre) ci vogliono lunghissimi tempi di calcolo anche su un elaboratore molto potente.

Dati due insiemi A e B, se ogni elemento di A è anche elemento di B, diremo che A è un *sottoinsieme* di B, o che è *contenuto* in B, o anche che B è un *soprainsieme* di A, o che *contiene* A, e scriveremo

$$A \subseteq B \quad , \quad B \supseteq A . \tag{1.7}$$

Osserviamo esplicitamente che, con questa notazione, per ogni insieme A si ha $A \subseteq A$, cioè ogni insieme è contenuto in se stesso. Per indicare che $A \subseteq B$, ma che esiste qualche elemento di B che non è contenuto in A useremo la scrittura

$$A \subset B, \text{ oppure } B \supset A \tag{1.8}$$

e parleremo di sottoinsieme (o soprainsieme) *proprio*.

Alcuni usano la notazione (1.8) per indicare i sottoinsiemi, anche non propri. In questo caso per indicare i sottoinsiemi propri si usa una delle notazioni seguenti:

$$A \subsetneq B \quad \text{oppure} \quad A \underset{\neq}{\subseteq} B.$$

Tra i vari sottoinsiemi di un insieme possiamo sempre annoverare anche l'insieme vuoto: $\emptyset \subseteq A$, $\forall A$. Ci potranno interessare anche sottoinsiemi costituiti da un solo elemento: se $a \in A$, allora $\{\, a \,\} \subseteq A$. Si noti la radicale differenza che c'è tra i due simboli $\in$ e $\subset$ (o $\subseteq$): il primo mette in relazione oggetti diversi (elementi e insiemi), il secondo mette in relazione oggetti dello stesso tipo (insiemi).

Dato un insieme A ammettiamo di poter considerare l'insieme di tutti i suoi sottoinsiemi, detto *insieme delle parti* e indicato con $\mathscr{P}(A)$. Per esempio, se $A = \{\, a,\ b \,\}$, allora

$$\mathscr{P}(A) = \{\, \emptyset, \{\, a \,\}, \{\, b \,\}, A \,\} .$$

Usando l'insieme delle parti si possono costruire, "sul vuoto", insiemi molto complessi. Si vedano gli esempi che seguono.

$$\begin{aligned} \mathscr{P}(\emptyset) &= \{\, \emptyset \,\} \\ \mathscr{P}(\mathscr{P}(\emptyset)) &= \{\, \emptyset, \{\, \emptyset \,\} \,\} \\ \mathscr{P}(\mathscr{P}(\mathscr{P}(\emptyset))) &= \{\, \emptyset, \{\, \emptyset \,\}, \{\, \{\, \emptyset \,\} \,\}, \{\, \emptyset, \{\, \emptyset \,\} \,\} \,\} \end{aligned}$$

Mentre l'insieme vuoto non ha elementi, gli insiemi qui sopra proposti hanno, nell'ordine, 1, 2, 4 elementi. Torneremo in seguito sul problema del conteggio degli elementi di un insieme e dell'insieme delle sue parti.

1.6. Operazioni tra insiemi

Definizione 1.3 (Unione di insiemi). *Dati due insiemi A e B, si chiama loro* unione, *e si indica con $A \cup B$, l'insieme formato dagli elementi che appartengono ad A, a B o a entrambi.*

$$A \cup B \stackrel{\text{def}}{=} \{ x \mid x \in A \vee x \in B \} . \tag{1.9}$$

Esempio 1.2. Se $A = \{ 0, 1, 2, 3 \}$ e $B = \{ 2, 3, 4 \}$, allora $A \cup B = \{ 0, 1, 2, 3, 4 \}$.

Definizione 1.4 (Intersezione di insiemi). *Dati due insiemi A e B, si chiama loro* intersezione, *e si indica con $A \cap B$, l'insieme formato dagli elementi che appartengono contemporaneamente ad A e a B.*

$$A \cap B \stackrel{\text{def}}{=} \{ x \mid x \in A \wedge x \in B \} . \tag{1.10}$$

Esempio 1.3. Se A e B sono come nell'esempio precedente, allora $A \cap B = \{ 2, 3 \}$.

Due insiemi la cui intersezione sia vuota si dicono *disgiunti*. L'insieme vuoto è sempre disgiunto da ogni altro insieme.

Le operazioni di unione e intersezione sono ovviamente associative e dunque si potrà scrivere l'unione o intersezione di più insiemi senza usare alcuna parentesi.

Saremo anche interessati a considerare *famiglie di insiemi*: se a ogni elemento di un dato insieme $\mathcal{A} \neq \emptyset$ corrisponde un insieme A, la famiglia di insiemi sarà denotata con

$$\{ A_\alpha \mid \alpha \in \mathcal{A} \} .$$

Quando non si darà adito a equivoci potremo anche denotare una famiglia di insiemi semplicemente con $\{ A_\alpha \}$, senza precisare l'insieme $\mathcal{A}$ di variabilità degli indici. In molte situazioni l'insieme $\mathcal{A}$ sarà l'insieme dei numeri naturali o un suo sottoinsieme. Per le unioni di tutti gli insiemi di una famiglia useremo scritture del tipo

$$\bigcup_{\alpha \in \mathcal{A}} A_\alpha \quad , \quad \bigcup_{n=0}^{+\infty} A_n \quad , \quad \bigcup_{n=0}^{10} A_n \quad , \quad \bigcup_{n \in \{1,2,\dots,k\}} A_n \quad , \quad \text{ecc.,}$$

e analoghe per le intersezioni[6]. Potremo anche usare notazioni abbreviate come

$$\bigcup_{\alpha} A_\alpha$$

se l'insieme di variabilità degli indici è chiaro dal contesto.

Le seguenti sono alcune proprietà di uso comune dell'unione e dell'intersezione.

$$A \cup A = A; \quad A \cap A = A; \tag{1.11a}$$

$$A \cup B = B \cup A; \quad A \cap B = B \cap A; \tag{1.11b}$$

$$A \cup \emptyset = A; \quad A \cap \emptyset = \emptyset; \tag{1.11c}$$

$$A \cup B \supseteq A; \quad A \cap B \subseteq A; \tag{1.11d}$$

$$A \cup B = A \Leftrightarrow A \supseteq B; \quad A \cap B = A \Leftrightarrow A \subseteq B. \tag{1.11e}$$

[6] Si noti l'analogia con la notazione usata per il simbolo di sommatoria.

Valgono anche le proprietà distributive di un'operazione rispetto all'altra:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad , \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C). \tag{1.12}$$

Si noti come le proprietà espresse dalle (1.11) e (1.12) evidenzino il fatto che le operazioni di unione e intersezione tra insiemi sono *diverse* da quelle di somma e prodotto di numeri (anche se esistono alcune analogie). Per esempio mentre $A \cup A = A, \forall A$, $a + a = a$ solo se $a = 0$; inoltre mentre $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $a + (bc) \neq (a + b)(a + c)$.

Definizione 1.5 (Differenza di insiemi)**.** *Dati due insiemi A e B, si chiama loro* differenza*, e si indica con $A \setminus B$*[7]*, l'insieme formato dagli elementi che appartengono ad A ma non a B.*

$$A \setminus B \overset{\text{def}}{=} \{\, x \mid x \in A \wedge x \notin B \,\}. \tag{1.13}$$

Esempio 1.4. Se A e B sono come nell'esempio già considerato per l'unione, allora $A \setminus B = \{\, 0, 1 \,\}$.

Nel caso che $B \subseteq A$, l'insieme $A \setminus B$ si chiama anche *complementare di B rispetto ad A* e si indica con $\complement_A B$, o semplicemente con $\complement B$ se l'insieme A è precisato una volta per tutte. In molte situazioni si conviene di fissare un insieme, detto *universo*, di cui tutti gli insiemi della teoria sono sottoinsiemi: questo evita di avere problemi tipo quelli del paradosso del barbiere. In questo caso quando si parla di complementare senza ulteriori precisazioni si intende sempre il complementare rispetto all'universo.

Definizione 1.6 (Differenza simmetrica)**.** *Dati due insiemi A e B, si chiama loro* differenza simmetrica*, e si indica con $A \triangle B$, l'insieme formato dagli elementi che appartengono a uno e uno solo dei due insiemi A e B. In formule*

$$A \triangle B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B). \tag{1.14}$$

Esempio 1.5. Se $A = \{\, 1, 2, 3, 5, 7 \,\}$ e $B = \{\, 2, 3, 4, 5, 6 \,\}$, si ha $A \triangle B = \{\, 1, 4, 6, 7 \,\}$.

Utilizzando il concetto di complementare di un insieme (rispetto all'universo) è possibile scrivere le operazioni di differenza e differenza simmetrica di due insiemi mediante le sole operazioni di unione e di intersezione e questo spesso facilita la dimostrazione di certe proprietà degli insiemi. Precisamente si ha:

$$A \setminus B = A \cap \complement B, \quad A \triangle B = (A \cap \complement B) \cup (B \cap \complement A), \tag{1.15}$$

come è facile verificare.

In analogia con quanto fatto nella logica delle proposizioni si dimostrano facilmente le seguenti formule, dette *formule di de Morgan*:

$$\complement(A \cap B) = \complement A \cup \complement B \quad \text{e} \quad \complement(A \cup B) = \complement A \cap \complement B. \tag{1.16}$$

Assumiamo anche un altro concetto primitivo, che utilizzeremo continuamente, e precisamente quello di *coppia ordinata*, che indicheremo con (x, y), dove è importante il posto occupato dagli elementi x e y:

$$(x, y) = (x_1, y_1) \Leftrightarrow x = x_1 \wedge y = y_1.$$

Conviene osservare esplicitamente che, in generale,

$$\{\, a, b \,\} = \{\, b, a \,\} \quad \text{mentre} \quad (a, b) \neq (b, a).$$

[7] Alcuni testi usano la scrittura $A - B$ al posto di $A \setminus B$, ma questa notazione è sconsigliata dalle norme ISO.

Definizione 1.7 (Prodotto cartesiano). *Dati due insiemi A e B si chiama loro* prodotto cartesiano, *o semplicemente* prodotto, *l'insieme, indicato con $A \times B$, delle coppie ordinate il cui primo elemento appartiene ad A e il secondo a B:*

$$A \times B \stackrel{\text{def}}{=} \{ (a,b) \mid (a \in A) \wedge (b \in B) \} .$$

È una conseguenza immediata della definizione che $A \times B \neq B \times A$. Nel caso particolare che $A = B$ si scrive anche A^2 in luogo di $A \times A$.

Si possono considerare anche prodotti cartesiani di più di due insiemi (attenzione all'ordine!) e, nel caso del prodotto cartesiano di un insieme per se stesso n volte si scriverà A^n in luogo di $A \times A \times \cdots \times A$.

Definizione 1.8 (Partizione). *Se A è un insieme non vuoto, una famiglia $\{ A_\alpha \}$ di sottoinsiemi di A si dirà una* partizione *o* ripartizione *di A se:*

1. $A_\alpha \neq \emptyset \, \forall \, \alpha$;
2. $A_\alpha \cap A_\beta = \emptyset$, *se* $\alpha \neq \beta$;
3. $\bigcup\limits_\alpha A_\alpha = A$.

In sostanza una partizione di un insieme è una suddivisione dell'insieme stesso in sottoinsiemi non vuoti e a due a due disgiunti, in modo che l'unione dei sottoinsiemi riproduca l'insieme originale. Gli elementi della famiglia si diranno anche *classi*.

Esempio 1.6. Sia A l'insieme dei punti di un piano e r una retta dello stesso piano. La famiglia di tutte le rette parallele a r costituisce una ripartizione di A.

Esempio 1.7. Sia $\mathbb{Z}$ l'insieme degli interi: la famiglia di insiemi $\{ Z^-, \{0\}, Z^+ \}$ costituisce una ripartizione di $\mathbb{Z}$.

Esempio 1.8. Sia A l'insieme di tutte le rette di un piano e $\mathcal{A}$ l'insieme di tutte le rette del fascio avente centro in un punto O qualsiasi. Considerato l'insieme

$$A_\alpha = \{ r \mid r \text{ è una retta parallela a una retta } r_\alpha \text{ di } \mathcal{A} \},$$

la famiglia $\{ A_\alpha \}$ è una ripartizione di A.

Esempio 1.9. Sia $\mathbb{N}^*$ l'insieme dei naturali positivi (cioè maggiori di 0) e consideriamo i sottoinsiemi $A =$ insieme dei pari, $B =$ insieme dei primi, $C =$ insieme dei dispari non primi. Allora la famiglia $\{ A, B, C \}$ non costituisce una ripartizione di $\mathbb{N}^*$, in quanto l'unione degli elementi della famiglia riproduce $\mathbb{N}^*$, ma l'intersezione a due a due non è vuota perché $2 \in A \wedge 2 \in B$.

1.7. Diagrammi di Eulero-Venn e altre rappresentazioni grafiche

In molte situazioni è utile servirsi dei cosiddetti *diagrammi di Eulero-Venn* per rappresentare gli insiemi e verificare le proprietà delle operazioni tra insiemi. In questo tipo di diagrammi gli insiemi sono individuati da regioni del piano delimitate da una curva chiusa. In certi casi si conviene di evidenziare esplicitamente alcuni elementi di un insieme mediante punti: la cosa è particolarmente utile nel caso di insiemi finiti, quando si possono anche evidenziare tutti gli elementi degli insiemi stessi. Pur essendo questo tipo di rappresentazione grafica molto significativa, non bisogna abusarne ed è opportuno prestare la massima attenzione.

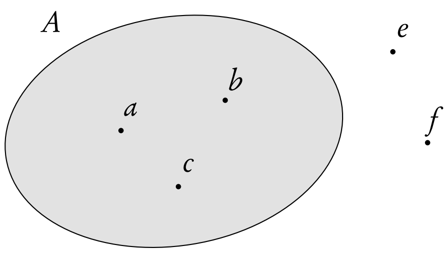


Figura 1.1.: *Un insieme con alcuni elementi che gli appartengono ed altri che non gli appartengono*

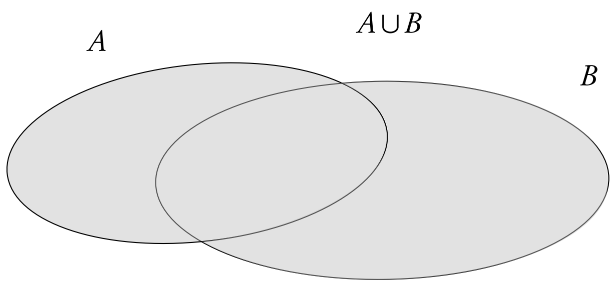


Figura 1.2.: *Due insiemi A e B e la loro unione*

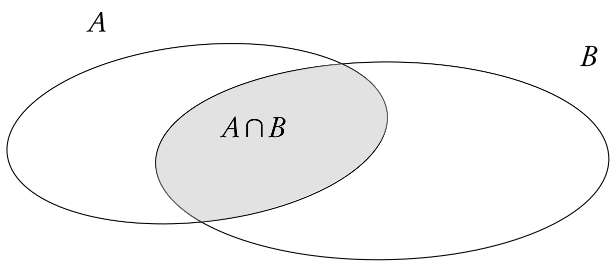


Figura 1.3.: *Due insiemi A e B e la loro intersezione*

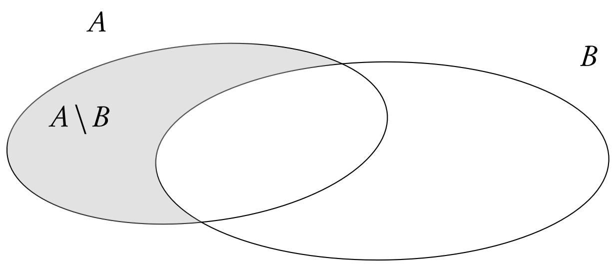


Figura 1.4.: *Due insiemi A e B e la differenza* $A \setminus B$

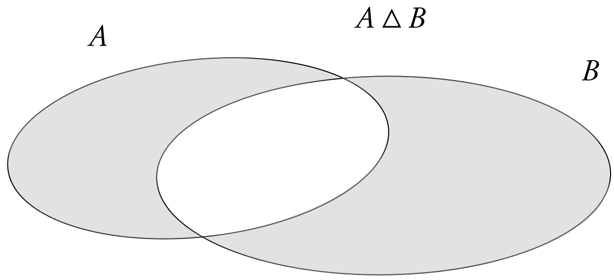


Figura 1.5.: *Due insiemi A e B e la loro differenza simmetrica*

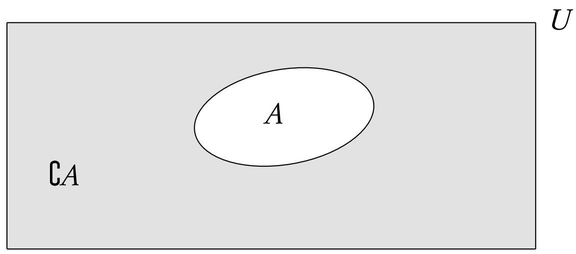


Figura 1.6.: *Complementare di un insieme rispetto all'universo U*

Nella figura 1.1 è rappresentato un insieme A con evidenziati tre elementi, a, b, c che gli appartengono e due elementi e, f, che non gli appartengono.

Nel caso di sottoinsiemi dell'insieme dei numeri reali (tra cui anche $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$) spesso è conveniente la rappresentazione su una retta sulla quale si sia fissato un sistema di ascisse. In casi come questo è utile rappresentare su una riga l'intera retta con l'indicazione dell'origine e dell'orientamento e su una riga sottostante l'insieme in considerazione; se gli insiemi sono più d'uno si possono usare diverse righe, una per ogni insieme: questo sistema rende facile individuare intersezioni, unioni, differenze, ecc.

Vediamo il tutto su un esempio. Si considerino gli insiemi $A = \{\, x \in \mathbb{R} \mid 2 < x \leq 5 \,\} =]2,5]$ e $B = \{\, 2,3,4,6 \,\}$. Una rappresentazione grafica molto conveniente è quella della figura 1.7.

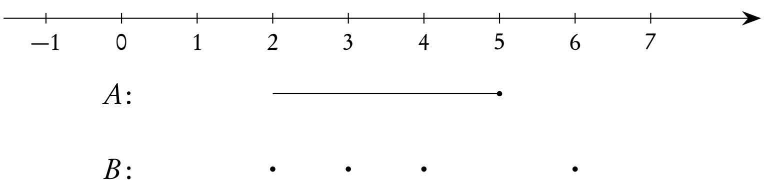


Figura 1.7.: *Rappresentazione grafica per sottoinsiemi di* $\mathbb{R}$

Si noti, nella figura 1.7, che abbiamo evidenziato il fatto che il punto 5 appartiene all'insieme A, mentre il punto 2 non vi appartiene: è opportuno essere molto precisi in questo, anche in vista delle tecniche da usare nella risoluzione di disequazioni e sistemi di disequazioni.

Dalla figura 1.7 si deduce facilmente che

— $A \cup B = [2,5] \cup \{ 6 \}$;
— $A \cap B = \{ 3,4 \}$;
— $A \setminus B =]2,3[\cup]3,4[\cup]4,5]$;
— $B \setminus A = \{ 2,6 \}$;
— ....

Ricordiamo, infine, che per rappresentare sottoinsiemi di $\mathbb{R} \times \mathbb{R}$ si usa normalmente un piano in cui si sia introdotto un riferimento cartesiano (di solito ortogonale e molto spesso monometrico) Oxy.

1.8. Relazioni binarie

1.8.1. Definizioni

Definizione 1.9 (Relazione)**.** *Siano A e B insiemi non vuoti. Chiamiamo* relazione (binaria) *tra A e B un qualunque sottoinsieme, $\mathcal{R}$, del prodotto cartesiano $A \times B$. Se $A = B$ la relazione è detta relazione in A.*

Se $\mathcal{R}$ è una relazione e $(a,b) \in \mathcal{R}$, scriveremo $a\mathcal{R}b$ e diremo che a è in relazione con b. Trattandosi di un insieme, la relazione potrà essere data sia specificando l'insieme che assegnando una proprietà caratteristica che individui gli elementi della relazione, cioè le coppie (a,b) tali che a sia in relazione con b. In sostanza una relazione è un "legame" tra due insiemi (o all'interno di uno stesso insieme) che collega, mette in relazione appunto, gli elementi a due a due.

Esempio 1.10. La relazione "$\le$" è una relazione in $\mathbb{R}$, cioè è un sottoinsieme di $\mathbb{R} \times \mathbb{R}$. Poiché $\mathbb{R} \times \mathbb{R}$ può essere rappresentato in un piano in cui si sia introdotto un sistema di coordinate cartesiane ortogonali (monometrico), potremo rappresentare questa relazione con un sottoinsieme del piano stesso. Si veda la figura 1.8.

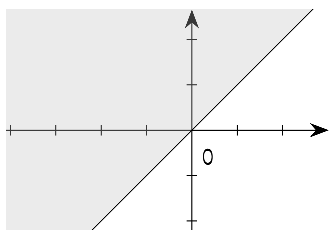


Figura 1.8.: *La relazione "$\le$" in* $\mathbb{R}$

Esempio 1.11. "Essere parallelo" è una relazione nell'insieme di tutte le rette dello spazio.

Esempio 1.12. "$x\mathcal{R}y$ se $x+y$ è pari" è una relazione nell'insieme $\mathbb{Z}$ degli interi.

Esempio 1.13. "Essere nati nello stesso giorno dell'anno" è una relazione che si può stabilire in un certo gruppo di individui.

Una rappresentazione grafica significativa si può fare per le relazioni tra insiemi finiti, in maniera simile alla rappresentazione cartesiana sopra considerata. Per esempio siano $A = \{ a, b, c \}$ e $B = \{ *, +, \#, - \}$ e consideriamo la seguente *tabella a doppia entrata*, in cui abbiamo messo gli elementi di A sull'orizzontale e quelli di B sulla verticale, in maniera simile a quanto si fa con il piano cartesiano.

$-$			
#			
$+$			
$*$			
	a	b	c

Basterà, a questo punto, segnare nei quadratini se c'è o no relazione tra un elemento di A e uno di B. Per esempio la relazione costituita dalle coppie $(a,+)$, $(a,-)$, $(b,\#)$, $(c,*)$ sarà rappresentata dalla tabella seguente.

$-$	√		
#		√	
$+$	√		
$*$			√
	a	b	c

Le relazioni su un insieme A possono godere di alcune proprietà: elenchiamo quelle che più ci interessano.

Proprietà riflessiva $\forall x \in A \;\; x\mathcal{R}x$, ovvero "ogni elemento è in relazione con se stesso". Un esempio è fornito dalla relazione di parallelismo tra le rette dello spazio: $r \parallel r$ è vera per ogni retta r dello spazio.

Proprietà simmetrica $\forall x, y \in A \;\; x\mathcal{R}y \Rightarrow y\mathcal{R}x$, ovvero "se x è in relazione con y, anche y è in relazione con x". Un esempio è ancora fornito dalla relazione di parallelismo prima citata: se $r \parallel s$, allora $s \parallel r$.

Proprietà antisimmetrica $\forall x, y \in A \;\; x\mathcal{R}y \wedge y\mathcal{R}x \Rightarrow x = y$, ovvero "se x è in relazione con y e y è in relazione con x, x e y sono uguali". Un esempio è fornito dalla relazione $\leq$: se $x \leq y \wedge y \leq x \Rightarrow x = y$.

Proprietà transitiva $\forall x, y, z \in A \;\; x\mathcal{R}y \wedge y\mathcal{R}z \Rightarrow x\mathcal{R}z$, ovvero se x è in relazione con y e y è in relazione con z, allora x è in relazione con z. Esempi sono forniti dalle relazioni di parallelismo e $\leq$ prima citate.

Per le relazioni su un insieme finito A, la rappresentazione *tabulare* sopra considerata è particolarmente utile in quanto permette di verificare subito alcune proprietà della relazione, per esempio in maniera molto facile la riflessività e la simmetria. Nelle quattro relazioni rappresentate dalle tabelle seguenti sull'insieme $A = \{a, b, c\}$, la prima è simmetrica e transitiva, la seconda riflessiva, simmetrica e transitiva, la terza è riflessiva, antisimmetrica e transitiva, la quarta è transitiva.

1:

c			
b	√	√	
a	√	√	
	a	b	c

, 2:

c		√	√
b		√	√
a	√		
	a	b	c

, 3:

c			√
b		√	√
a	√	√	√
	a	b	c

, 4:

c		√	
b			
a	√	√	√
	a	b	c

1.8.2. Relazioni di equivalenza

Un primo tipo di relazioni molto importante nelle applicazioni è quello delle relazioni di equivalenza.

Definizione 1.10 (Relazione di equivalenza)**.** *Sia A un insieme non vuoto e $\mathcal{R}$ una relazione in A. $\mathcal{R}$ è detta una* relazione di equivalenza *se gode delle proprietà*

- *riflessiva,*
- *simmetrica,*
- *transitiva.*

Spesso le relazioni di equivalenza sono indicate con il simbolo $\sim$.

Esempio 1.14. La relazione di parallelismo tra rette dello spazio è di equivalenza.

Se in un insieme A è definita una relazione di equivalenza $\mathcal{R}$, è possibile costruire una speciale famiglia di sottoinsiemi di A. Precisamente si dà la definizione seguente.

Definizione 1.11 (Classe di equivalenza)**.** *Sia $\mathcal{R}$ una relazione di equivalenza in un insieme A e sia $x \in A$. L'insieme di tutti gli y che sono in relazione con x si indica con $[x]$ e si chiama la* classe di equivalenza *individuata da x, cioé si pone*

$$[x] \stackrel{\text{def}}{=} \{\, y \in A \mid x\mathcal{R}y \,\}. \tag{1.17}$$

Ogni elemento $x \in [x]$ si chiama un rappresentante *della classe $[x]$.*

Non è difficile provare (e lo si può fare come semplice ed utile esercizio) che la famiglia delle classi di equivalenza costituisce una partizione di A stesso. Questa famiglia prende il nome di *insieme quoziente* di A rispetto alla relazione di equivalenza $\mathcal{R}$. In molte situazioni queste classi di equivalenza prendono nomi speciali, come risulta dagli esempi che seguono.

- Nell'insieme delle rette dello spazio dove si è introdotta la relazione di parallelismo, le classi di equivalenza prendono il nome di *direzioni*.
- Nell'insieme dei piani dello spazio dove si è introdotta la relazione di parallelismo, le classi di equivalenza prendono il nome di *giaciture*.

Un esempio molto interessante[8] è costituito dalla seguente relazione nell'insieme $\mathbb{N} \times \mathbb{N}$ delle coppie di naturali:

$$(m,n) \sim (p,q) \quad \text{se} \quad n+q = m+p\,.$$

Se rappresentiamo l'insieme $\mathbb{N} \times \mathbb{N}$ sul piano cartesiano, non è difficile verificare che le classi di equivalenza sono costituite da tutte quelle coppie di naturali che stanno sulle rette parallele alla bisettrice $y = x$, come evidenziato nella figura 1.9, da cui si evince altresì che l'insieme delle classi di equivalenza è una partizione di $\mathbb{N} \times \mathbb{N}$.

[8]Si tratta della strategia che porta alla costruzione dei numeri interi relativi a partire dai naturali, ma non intendiamo insistere su questo fatto.

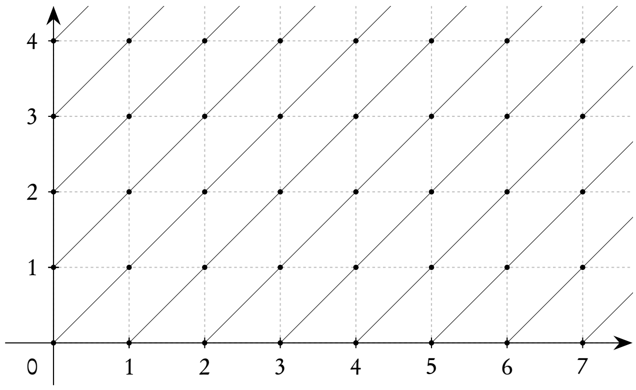


Figura 1.9.: *Classi di equivalenza della relazione "*$(m,n) \sim (p,q) \quad \text{se} \quad n+q = m+p$*" in* $\mathbb{N}\times\mathbb{N}$

1.8.3. Relazioni d'ordine

Un secondo tipo di relazioni, anche questo importante nelle applicazioni, è quello delle relazioni d'ordine.

Definizione 1.12 (Relazione d'ordine)**.** *Sia A un insieme non vuoto e $\mathcal{R}$ una relazione in A. $\mathcal{R}$ è detta una* relazione d'ordine *se gode delle proprietà*

— *riflessiva,*
— *antisimmetrica,*
— *transitiva.*

Esempio 1.15. Se $\mathscr{P}(A)$ è l'insieme delle parti di un insieme A, la relazione $\subseteq$ è una relazione d'ordine.

Esempio 1.16. Negli insiemi $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ la relazione $\leq$ è d'ordine.

Le relazioni d'ordine sono di solito indicate con simboli del tipo $\preceq$: la scrittura $a \preceq b$ si legge *a precede b* o anche *b segue a*. In una relazione d'ordine indicata con $\preceq$, per indicare che b segue a si scrive anche $b \succeq a$.

È evidente che se A è un insieme ordinato con la relazione $\preceq$, ogni suo sottoinsieme è ancora ordinato con la stessa relazione "ristretta" al sottoinsieme. Useremo continuamente questo fatto.

Una relazione d'ordine si dice *totale* se dati due elementi a e b, accade sempre che $a \preceq b \vee b \preceq a$: in altri termini una relazione d'ordine è totale se due elementi qualunque sono sempre confrontabili.

Esempio 1.17. Se A è un insieme con un solo elemento, la relazione $\subseteq$ in $\mathscr{P}(A)$ è d'ordine totale.

Esempio 1.18. Se A è un insieme con più di un elemento, la relazione $\subseteq$ in $\mathscr{P}(A)$ non è d'ordine totale. Per esempio se $A = \{\, a, b \,\}$ si ha $\mathscr{P}(A) = \{\, \emptyset, \{a\}, \{b\}, \{\, a, b \,\} \,\}$ ed è chiaro che $\{a\}$ non è contenuto né contiene $\{b\}$.

Esempio 1.19. La relazione $\leq$ negli insiemi numerici sopra citati è d'ordine totale.

Una relazione d'ordine che non sia d'ordine totale si dice d'ordine *parziale*.

Definizione 1.13 (Minimo e massimo di un insieme)**.** *Sia A un insieme ordinato con la relazione $\preceq$. Se esiste un elemento $m \in A$ tale che $m \preceq x \ \forall x \in A$, allora m si dice il* minimo *di A. Analogamente se esiste un elemento $M \in A$ tale che $x \preceq M \ \forall x \in A$, allora M si dice il* massimo *di A.*

L'articolo determinativo usato sia davanti al minimo che al massimo è giustificato dal seguente teorema che ne garantisce l'unicità.

Teorema 1.14. *Se un insieme ordinato ha minimo e/o massimo, essi sono unici.*

Dimostrazione. Se m_1 ed m_2 sono due minimi si deve avere

$$(m_1 \preceq m_2) \wedge (m_2 \preceq m_1),$$

da cui $m_1 = m_2$, per la proprietà antisimmetrica. Analogo discorso per il massimo. □

Esempio 1.20. Nell'insieme delle parti di un dato insieme A, l'insieme vuoto è sempre il minimo e l'insieme A stesso è sempre il massimo rispetto alla relazione di inclusione tra insiemi $\subseteq$.

Esempio 1.21. L'insieme dei numeri naturali $\mathbb{N}$ ha lo zero come minimo e non ha massimo, rispetto all'ordine usuale.

Esempio 1.22. L'insieme degli interi non ha né massimo né minimo rispetto all'ordine usuale.

Come mostrano gli esempi appena considerati un insieme può tranquillamente non avere massimo e/o minimo. Per questo motivo si "inventa" un surrogato del massimo e del minimo che gioca un ruolo simile, ma ovviamente non identico. Purtroppo nemmeno questo surrogato esiste sempre, come vedremo.

Definizione 1.15 (Minoranti e maggioranti)**.** *Sia A un insieme ordinato con la relazione $\preceq$ e B un sottoinsieme di A (eventualmente coincidente con A stesso). Se esiste un elemento $l \in A$ tale che $l \preceq x$ $\forall x \in B$, l si dice un* minorante *o una* limitazione inferiore *per B.*

Analogamente se esiste un elemento $L \in A$ tale che $x \preceq L$ $\forall x \in B$, L si dice un maggiorante *o una* limitazione superiore *per B.*

Un sottoinsieme B che abbia maggioranti si dice superiormente limitato*, uno che abbia minoranti si dice* inferiormente limitato*. Se il sottoinsieme ha sia maggioranti che minoranti si dice semplicemente* limitato*.*

È ovvio, considerando il caso $B = A$, che l'insieme A stesso può avere come maggioranti solo l'eventuale massimo e come minoranti solo l'eventuale minimo.

Per un generico sottoinsieme B possiamo considerare l'insieme dei maggioranti e l'insieme dei minoranti (che potrebbero anche essere vuoti). Ha interesse sapere se questi due insiemi hanno, oppure no, minimo (il primo) e massimo (il secondo).

Definizione 1.16 (Estremo superiore e inferiore)**.** *Sia A un insieme ordinato con la relazione $\preceq$ e B un sottoinsieme di A. Se l'insieme dei maggioranti di B ha un minimo, esso si chiama l'*estremo superiore *di B e si indica con $\sup B$; se l'insieme dei minoranti ha un massimo esso si chiama l'*estremo inferiore *di B e si indica con $\inf B$.*

Come già fatto con il massimo e il minimo si può osservare che ci può essere al più un estremo superiore (o inferiore).

Esempio 1.23. Sia $B = \left\{ x \in \mathbb{Q} \mid x^2 < 4 \right\}$. Allora $\sup B = 2$ e $\inf B = -2$.

Esempio 1.24. Sia $B = \{\, x \in \mathbb{Q} \mid x < 0 \vee (x \geq 0 \wedge x^2 < 2) \,\}$. Allora $\inf B$ non può esistere perché B non è inferiormente limitato, e non esiste nemmeno $\sup B$, nonostante B abbia ovviamente dei maggioranti. L'esistenza di sottoinsiemi dei razionali che, pur essendo superiormente limitati, non hanno estremo superiore è un grave "handicap" dell'insieme dei razionali, ed è per questo che si è costretti a introdurre l'insieme dei reali, dove le cose vanno in maniera completamente diversa[9].

A partire da una relazione d'ordine del tipo che abbiamo introdotto, si può introdurre una nuova relazione, che possiamo indicare con $\prec$, ponendo

$$a \prec b \quad \text{se e solo se} \quad (a \preceq b) \wedge (a \neq b).$$

Al posto della proprietà riflessiva questa relazione gode della proprietà *antiriflessiva*: nessun elemento è in relazione con se stesso. Una relazione siffatta si chiama *relazione d'ordine in forma antirifilessiva*.

È anche evidente che da una relazione d'ordine in forma antiriflessiva si può ottenere una relazione d'ordine usuale: i due tipi di relazioni d'ordine hanno dunque lo stesso contenuto di informazione.

1.9. Funzioni o applicazioni

Richiamiamo qui il concetto di funzione con le proprietà essenziali, riservandoci di ritornarci sopra in dettaglio nel capitolo 4, vista l'importanza del concetto.

Tra tutte le relazioni che possono intercorrere tra due insiemi, hanno un interesse cruciale le *funzioni* o *applicazioni*[10]; in questo testo in particolare considereremo le funzioni che collegano tra di loro insiemi di numeri reali. La seguente definizione riassume quelle che sono le proprietà che ci interesseranno.

Definizione 1.17 (Funzione)**.** *Dati due insiemi A e B, si dice* funzione *di A in B una qualunque legge che faccia corrispondere* a ogni *elemento di A* uno ed un solo elemento *di B.*

L'insieme A è detto dominio *della funzione, l'insieme B è detto* codominio*. Se x è un elemento dell'insieme A e y è l'*unico *elemento di B che corrisponde ad A, si dice che y è* funzione *di x e si scrive $y = f(x)$ (leggi: "y uguale a effe di x"). L'elemento y di B è anche detto* immagine *di x tramite f.*

La notazione più completa per le funzioni è la seguente:

$$f : A \to B,\ x \mapsto f(x),$$

ma spesso si scrive solo

$$x \mapsto f(x),$$

se gli insiemi A e B sono già stati precisati o sono chiari dal contesto. Si può anche dire semplicemente *la funzione $y = f(x)$*, anche se questo tipo di scrittura può dar luogo ad equivoci e per questo fatto può far storcere il naso ai puristi.

Esempio 1.25. Se A e B sono l'insieme dei numeri reali, si può considerare la funzione che a ogni numero reale x fa corrispondere il suo quadrato. In questo caso si dovrebbe scrivere

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$$

[9] Ritorneremo brevemente su queste considerazioni trattando i numeri razionali e reali.

[10] I due termini sono qui considerati sinonimi, anche se alcuni riservano il nome di funzioni al caso in cui il dominio e il codominio sono insiemi di numeri reali.

ma si può scrivere anche semplicemente

$$x \mapsto x^2$$

oppure (e spesso questa è la notazione usata nei testi di scuola secondaria)

$$y = x^2 .$$

Si noti, nella definizione di funzione, che è *obbligatorio* che ad ogni punto dell'insieme A corrisponda esattamente un punto dell'insieme B, cioè che ogni punto dell'insieme A sia in relazione con esattamente un punto dell'insieme B, mentre non è affatto richiesto che ogni punto dell'insieme B sia in relazione con qualche punto dell'insieme A. In ogni caso l'insieme dei punti di B che è in relazione con qualche punto di A è molto importante e si dà la seguente definizione.

Definizione 1.18 (Insieme immagine). *Se $f : A \to B$ è una funzione, l'insieme degli $y \in B$ che siano immagine di almeno un $x \in A$ si chiama* insieme immagine *o semplicemente* immagine *di A tramite f e si indica con* $\mathrm{Im}(f)$ *o anche con* $f(A)$*. In formule:*

$$\mathrm{Im}(f) = \{\, y \in B \mid \exists x \in A,\ y = f(x) \,\} . \tag{1.18}$$

Hanno particolare interesse le funzioni che godono di alcune particolari proprietà aggiuntive, precisamente:

1. funzioni in cui punti diversi del dominio hanno immagini diverse;
2. funzioni in cui ciascun punto del codominio è immagine di almeno un punto del dominio;
3. funzioni in cui ciascun punto del codominio è immagine di un solo punto del dominio.

Diremo *iniettive* le funzioni che soddisfano al primo requisito, *suriettive* quelle che soddisfano al secondo, *biiettive* o *biunivoche* quelle che soddisfano al terzo. Segnaliamo, anche se la cosa è evidente, che il terzo requisito si ottiene quando sono verificati *contemporaneamente* il primo e il secondo. Nel caso di funzioni biunivoche si parla spesso anche di *corrispondenza biunivoca* tra l'insieme A e l'insieme B. Le funzioni biunivoche sono particolarmente importanti: come vedremo nel capitolo 4 in relazione ad esse si può introdurre il concetto di funzione inversa.

1.10. Cardinalità di un insieme

Questo paragrafo contiene solo un cenno molto sommario e informale a questo importante concetto: sarà compito specifico dei corsi universitari trattare in dettaglio l'argomento.

Cominciamo col dare la definizione di *equipotenza* tra insiemi.

Definizione 1.19 (Insiemi equipotenti). *Due insiemi A e B si dicono* equipotenti *se esiste una corrispondenza biunivoca tra i due insiemi.*

Esempio 1.26. Siano $A = \{\, 1, 2, 3 \,\}$ e $B = \{\, a, b, c \,\}$. La funzione definita da $f(1) = a$, $f(2) = b$, $f(3) = c$ è una corrispondenza biunivoca tra i due insiemi, che dunque sono equipotenti.

Esempio 1.27. Siano A l'insieme dei numeri naturali pari, $A = \{\, 0, 2, 4, 6, \dots \,\}$ e B l'insieme dei numeri naturali dispari, $B = \{\, 1, 3, 5, \dots \,\}$. La funzione che a ogni elemento n di A fa corrispondere il naturale successivo $n + 1$ è una corrispondenza biunivoca tra i due insiemi, che dunque sono equipotenti.

Esempio 1.28. Siano A l'insieme dei numeri naturali pari, $A = \{0, 2, 4, 6, \dots\}$ e $\mathbb{N}$ l'insieme di tutti i numeri naturali. La funzione che fa corrispondere ad ogni elemento di A la sua metà è una corrispondenza biunivoca tra i due insiemi, che dunque sono equipotenti.

L'ultimo degli esempi proposti è particolarmente significativo: si può notare infatti che esso prova che è possibile stabilire una corrispondenza biunivoca anche tra un insieme e un suo sottoinsieme proprio. La cosa non è possibile per tutti gli insiemi e a questo proposito si dà la seguente definizione.

Definizione 1.20 (Insiemi finiti e infiniti). *Un insieme si dice* finito *se non è equipotente a nessun suo sottoinsieme proprio, cioè se non è possibile stabilire una corrispondenza biunivoca tra l'insieme stesso e alcun suo sottoinsieme proprio. Un insieme si dice* infinito *se è equipotente a un suo sottoinsieme proprio.*

L'insieme $A = \{1, 2, 3, \dots, 100\}$ è un insieme finito, l'insieme $\mathbb{N}$ è un insieme infinito.

Sussiste la seguente importante proprietà degli insiemi finiti: ogni insieme finito A si può mettere in corrispondenza biunivoca con un sottoinsieme di $\mathbb{N}$ del tipo $\{1, 2, 3, \dots, n\}$, per un unico valore di n; il numero n si chiama la *cardinalità* dell'insieme A e si scrive $|A| = n$. Per gli insiemi infiniti si dice che hanno una cardinalità *transfinita*. Due insiemi infiniti non sono necessariamente equipotenti: un esempio famoso è costituito dagli insiemi dei numeri naturali e dei numeri reali. La prova di questo fatto è un utile esercizio che proponiamo qui di seguito, anche perché è un bell'esempio di *dimostrazione per assurdo*.

Supponiamo dunque che sia possibile stabilire una corrispondenza biunivoca tra i numeri naturali e i numeri reali e rappresentiamo i numeri reali nella loro forma decimale: $\alpha.\beta_1\beta_2\dots$, dove α è un intero e β_1, β_2, $\dots$ sono le cifre decimali. Potremo così costruire una tabella come la seguente:

$$
\begin{aligned}
0 &\to \alpha_0.\beta_{01}\beta_{02}\beta_{03}\cdots \\
1 &\to \alpha_1.\beta_{11}\beta_{12}\beta_{13}\cdots \\
2 &\to \alpha_2.\beta_{21}\beta_{22}\beta_{23}\cdots \\
3 &\to \alpha_3.\beta_{31}\beta_{32}\beta_{33}\cdots \\
&\cdots \\
n &\to \alpha_n.\beta_{n1}\beta_{n2}\beta_{n3}\cdots \\
&\cdots
\end{aligned}
$$

Ora consideriamo il numero reale $\gamma.\delta_1\delta_2\delta_3\dots$ costruito come segue: se $\alpha_0 = 0$ allora $\gamma = 1$, altrimenti $\gamma = 0$; se $\beta_{01} = 0$ allora $\delta_1 = 1$, altrimenti $\delta_1 = 0$; se $\beta_{12} = 0$ allora $\delta_2 = 1$, altrimenti $\delta_2 = 0$; ecc. Questo numero non può essere in corrispondenza con nessun numero naturale, perché è diverso da tutti quelli che compaiono nella tabella precedente o per la parte prima della virgola o almeno per una cifra decimale. Questo significa che la corrispondenza in questione non può essere biunivoca.

Questa dimostrazione prova in realtà che l'insieme dei numeri reali è *più numeroso* che non quello dei numeri naturali, ma non insistiamo oltre su questo.

Accenniamo molto sommariamente anche al fatto che nell'insieme dei numeri cardinali si può stabilire un ordine: un numero cardinale α (cardinalità di un certo insieme A) è minore a un cardinale β (cardinalità di un certo insieme B) se A è equipotente a un sottoinsieme di B, ma non viceversa.

A proposito dei cardinali finiti può essere interessante osservare che, a partire solo dall'insieme vuoto, è possibile costruire insiemi di cardinalità n, per ogni valore di $n \in \mathbb{N}$:

$$
\begin{aligned}
&0 = |\emptyset| \\
&1 = |\{\,\emptyset\,\}| \\
&2 = |\{\,\emptyset, \{\,\emptyset\,\}\,\}| \\
&3 = |\{\,\emptyset, \{\,\emptyset\,\}, \{\,\emptyset, \{\,\emptyset\,\}\,\}\,\}| \\
&4 = |\{\,\emptyset, \{\,\emptyset\,\}, \{\,\emptyset, \{\,\emptyset\,\}\,\}, \{\,\emptyset, \{\,\emptyset\,\}, \{\,\emptyset, \{\,\emptyset\,\}\,\}\,\}\,\}| \\
&5 = \dots
\end{aligned}
$$

Come si può facilmente dedurre, ogni insieme della tabella è costituito facendo l'unione degli insiemi di tutte le righe precedenti. Poiché $\emptyset \subset \{\,\emptyset\,\} \subset \{\,\emptyset, \{\,\emptyset\,\}\,\} \subset \dots$, se ne ricava anche l'ordine "naturale" di questi cardinali finiti.

Per concludere questi pochi cenni relativi al concetto di cardinalità, ricordiamo che, se un insieme finito ha cardinalità n, l'insieme delle sue parti ha cardinalità 2^n, cosa che può essere provata in vari modi.

1.11. Esercizi

Esercizio 1.1. *Scrivere la negazione di "Qualche ape non ride".*

Risoluzione. Si potrebbe anche ragionare in termini elementari, ma è preferibile formalizzare il predicato assegnato usando i simboli della logica. Se consideriamo come universo l'insieme delle api e chiamiamo $\mathscr{P}(x)$ il predicato "x ride", la proposizione assegnata si può scrivere

$$\exists x \quad \text{tale che} \quad \neg\mathscr{P}(x).$$

La negazione è allora

$$\forall x \quad \text{vale} \quad \neg(\neg\mathscr{P}(x)),$$

ovvero

$$\forall x \quad \text{vale} \quad \mathscr{P}(x),$$

che si può tradurre a parole nella frase "Tutte le api ridono". □

Esercizio 1.2. *Dimostrare le proprietà distributive dell'unione rispetto all'intersezione e dell'intersezione rispetto all'unione (formule (1.12)):*

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad , \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Risoluzione. Il modo più semplice per risolvere esercizi di questo tipo è quello di costruire una tabella simile alle tavole di verità: si prende un x generico (appartenente all'insieme universo) e ci si chiede: $x \in A$?, $x \in B$?, ecc., per ognuno degli insiemi coinvolti, esaminando tutti i casi possibili. Successivamente si controlla che x appartiene al primo dei due insiemi dell'uguaglianza proposta se e solo se appartiene al secondo. In questi casi, essendoci tre insiemi in gioco, ci saranno 8 (cioè 2^3) possibilità. Si possono costruire le tabelle 1.3 e 1.4, che evidenziano come le due colonne segnate con $*$ abbiano gli stessi "valori di verità". □

Tabella 1.3.: *Tavola per la verifica di* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

A	B	C	$B \cap C$	$A \cup (B \cap C)$ *	$A \cup B$	$A \cup C$	$(A \cup B) \cap (A \cup C)$ *
SI	SI	SI	SI	SI	SI	SI	SI
SI	SI	NO	NO	SI	SI	SI	SI
SI	NO	SI	NO	SI	SI	SI	SI
SI	NO	NO	NO	SI	SI	SI	SI
NO	SI	SI	SI	SI	SI	SI	SI
NO	NO	SI	NO	NO	NO	SI	NO
NO	SI	NO	NO	NO	SI	NO	NO
NO	NO	NO	NO	NO	NO	NO	NO

Tabella 1.4.: *Tavola per la verifica di* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

A	B	C	$B \cup C$	$A \cap (B \cup C)$ *	$A \cap B$	$A \cap C$	$(A \cap B) \cup (A \cap C)$ *
SI	SI	SI	SI	SI	SI	SI	SI
SI	SI	NO	SI	SI	SI	NO	SI
SI	NO	SI	SI	SI	NO	SI	SI
SI	NO	NO	NO	NO	NO	NO	NO
NO	SI	SI	SI	NO	NO	NO	NO
NO	NO	SI	SI	NO	NO	NO	NO
NO	SI	NO	SI	NO	NO	NO	NO
NO	NO	NO	NO	NO	NO	NO	NO

Esercizio 1.3. *Dati tre insiemi* A, B, C, *si provi che*

$$(A \vartriangle B) \vartriangle C = A \vartriangle (B \vartriangle C),$$

ovvero che vale la proprietà associativa della differenza simmetrica.

Risoluzione. Si può procedere come indicato nell'esercizio 1.2. Si ottiene la tabella 1.5, da cui si evince che le colonne segnate con asterisco hanno gli stessi "valori di verità".

È interessante osservare che $A \vartriangle B \vartriangle C$ (avendo dimostrato la proprietà associativa possiamo eliminare le parentesi) è costituito dagli elementi che appartengono a 1 o a tutti 3 gli insiemi A, B, C; è questo un caso particolare di una proprietà generale della differenza simmetrica che si può dimostrare per induzione (argomento che esula dagli scopi di questo testo): $A_1 \vartriangle A_2 \vartriangle A_3 \cdots \vartriangle A_n$ è costituito dagli elementi che appartengono a un numero dispari degli insiemi dati. Per esercizio si può provare a dimostrarlo direttamente con 4 insiemi costruendo una tabella come la 1.5. □

Tabella 1.5.: *Tavola per la verifica di* $(A \triangle B) \triangle C = A \triangle (B \triangle C)$

				*		*
A	B	C	$A \triangle B$	$(A \triangle B) \triangle C$	$B \triangle C$	$A \triangle (B \triangle C)$
SI	SI	SI	NO	SI	NO	SI
SI	SI	NO	NO	NO	SI	NO
SI	NO	SI	SI	NO	SI	NO
SI	NO	NO	SI	SI	NO	SI
NO	SI	SI	SI	NO	NO	NO
NO	NO	SI	NO	SI	SI	SI
NO	SI	NO	SI	SI	SI	SI
NO	NO	NO	NO	NO	NO	NO

Esercizio 1.4. *Siano* A, B, C *tre insiemi tali che* $A \cap B = C$, $B \cap C = A$, $C \cap A = B$. *Provare che* $A = B = C$.

Risoluzione.

— Se $x \in A \Rightarrow x \in B \cap C \Rightarrow A \subseteq B$ e anche $A \subseteq C$;

— se $x \in B \Rightarrow x \in C \cap A \Rightarrow B \subseteq C$ e anche $B \subseteq A$;

— se $x \in C \Rightarrow x \in A \cap B \Rightarrow C \subseteq A$ e anche $C \subseteq B$.

Si conclude facilmente che $A = B = C$. □

Esercizio 1.5. *Dati tre insiemi* A, B, C, *si provi che*

$$A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C),$$

ovvero che l'intersezione tra insiemi è distributiva rispetto alla differenza simmetrica.

Risoluzione. Si può costruire, molto facilmente, una tabella come la 1.5 oppure dimostrare che un elemento x appartiene al primo insieme dell'uguaglianza se e solo se appartiene al secondo. In effetti,

— un elemento x appartiene al primo insieme se e solo se $x \in A$ e contemporaneamente a uno solo dei due insiemi B e C,

— un elemento x appartiene al secondo insieme se e solo se x appartiene contemporaneamente ad A e B oppure ad A e C, ma non a tutti e tre gli insiemi,

ed è evidente che queste due affermazioni sono equivalenti. □

Esercizio 1.6. *Date due proposizioni* $\mathscr{P}$ *e* $\mathscr{Q}$, *mostrare che*

$$\neg(\mathscr{P} \vee \neg\mathscr{Q}) \quad e \quad \neg\mathscr{P} \wedge \mathscr{Q}$$

sono equivalenti.

Tabella 1.6.: *Tavola per verificare l'equivalenza di* $\neg(\mathscr{P} \vee \neg\mathscr{Q})$ *e* $\neg\mathscr{P} \wedge \mathscr{Q}$

$\mathscr{P}$	$\mathscr{Q}$	$\neg\mathscr{Q}$	$\mathscr{P} \vee \neg\mathscr{Q}$	$\neg(\mathscr{P} \vee \neg\mathscr{Q})$ *	$\neg\mathscr{P}$	$\neg\mathscr{P} \wedge \mathscr{Q}$ *
V	V	F	V	F	F	F
V	F	V	V	F	F	F
F	V	F	F	V	V	V
F	F	V	V	F	V	F

Risoluzione. Basta costruire le tavole di verità delle due proposizioni e verificare che esse hanno gli stessi valori di verità. Si veda la tabella 1.6 □

Esercizio 1.7. *Date due proposizioni* $\mathscr{P}$ *e* $\mathscr{Q}$*, costruire la tavola di verità di*

$$\neg\mathscr{P} \vee (\mathscr{P} \vee \neg\mathscr{Q}).$$

Risoluzione. Con il metodo usuale si ottiene la tabella 1.7.

Come si vede la proposizione in oggetto è sempre vera, quali che siano i valori di verità delle due proposizioni componenti. Una proposizione sempre vera si chiama una *tautologia*. □

Tabella 1.7.: *Tavola di verità di* $\neg\mathscr{P} \vee (\mathscr{P} \vee \neg\mathscr{Q})$

$\mathscr{P}$	$\mathscr{Q}$	$\neg\mathscr{P}$	$\neg\mathscr{Q}$	$\mathscr{P} \vee \neg\mathscr{Q}$	$\neg\mathscr{P} \vee (\mathscr{P} \vee \neg\mathscr{Q})$
V	V	F	F	V	V
V	F	F	V	V	V
F	V	V	F	F	V
F	F	V	V	V	V

Esercizio 1.8. *Date due proposizioni* $\mathscr{P}$ *e* $\mathscr{Q}$*, costruire la tavola di verità di*

$$(\mathscr{P} \wedge \mathscr{Q}) \wedge \neg(\mathscr{P} \vee \mathscr{Q})$$

Risoluzione. Con il metodo usuale si ottiene la tabella 1.8.

Come si vede la proposizione in oggetto è sempre falsa, quali che siano i valori di verità delle due proposizioni componenti. Una proposizione sempre falsa si chiama una *contraddizione*. □

Esercizio 1.9. *Siano* A, B, C *tre insiemi. Provare che*

$$(A \cup B) \times C = (A \times C) \cup (B \times C).$$

Tabella 1.8.: *Tavola di verità di* $(\mathscr{P} \wedge \mathscr{Q}) \wedge \neg(\mathscr{P} \vee \mathscr{Q})$

$\mathscr{P}$	$\mathscr{Q}$	$\mathscr{P} \wedge \mathscr{Q}$	$\mathscr{P} \vee \mathscr{Q}$	$\neg(\mathscr{P} \vee \mathscr{Q})$	$(\mathscr{P} \wedge \mathscr{Q}) \wedge \neg(\mathscr{P} \vee \mathscr{Q})$
V	V	V	V	F	F
V	F	F	V	F	F
F	V	F	V	F	F
F	F	F	F	V	F

Risoluzione. Per risolvere questo tipo di problemi conviene provare che se un elemento x appartiene al primo insieme, appartiene anche al secondo e, viceversa, che se un x appartiene al secondo insieme appartiene anche al primo, ovvero che

$$(A \cup B) \times C \subseteq (A \times C) \cup (B \times C) \text{ e contemporaneamente } (A \cup B) \times C \supseteq (A \times C) \cup (B \times C).$$

Se $x \in (A \cup B) \times C$, allora x è una coppia del tipo (a, c) oppure del tipo (b, c), con $a \in A$, $b \in B$ e $c \in C$; quindi $x \in A \times C$ oppure $x \in B \times C$, e, infine, $x \in (A \times C) \cup (B \times C)$. Se, viceversa, $x \in (A \times C) \cup (B \times C)$, allora $x \in A \times C$ oppure $x \in B \times C$, dunque x è del tipo (a, c) oppure (b, c) con $a \in A$, $b \in B$ e $c \in C$; dunque $x \in (A \cup B) \times C$. □

Esercizio 1.10. *Sia* $A = \{ 3x \mid x \in \mathbb{N} \}$. *Trovare una partizione di* A *in due insiemi infiniti* B *e* C.

Risoluzione. Tra le possibili soluzioni c'è quella di considerare i multipli pari e dispari di 3: $B = \{ 3(2n) \mid n \in \mathbb{N} \}$, $C = \{ 3(2n + 1) \mid n \in \mathbb{N} \}$. □

Esercizio 1.11. *Siano* $A = \{ x \in \mathbb{R} \mid x^5 - 3x + 2 = 0 \}$ *e* $B = \{ x \in \mathbb{R} \mid x^6 - 3x^2 + 2x = 0 \}$. *Senza risolvere le equazioni (cosa del resto non elementarmente possibile) provare che* $A \subset B$.

Risoluzione. Poiché $x^6 - 3x^2 + 2x = x(x^5 - 3x + 2)$, se ne deduce che $B = \{ 0 \} \cup A$, cioè che A è un sottoinsieme di B. Poiché inoltre $0 \neq A$, se ne deduce che A è un sottoinsieme proprio di B. □

Esercizio 1.12. *Sia* A *e* B *due insiemi costituiti rispettivamente da* 17 *e* 22 *elementi. Se* $A \cup B$ *ha* 30 *elementi, quanti elementi ha* $A \cap B$*?*

Risoluzione. Se A e B fossero disgiunti, la loro unione avrebbe 39 elementi. Se invece $A \cup B$ ne ha solo 30, significa che 9 elementi sono comuni tra A e B, dunque $A \cap B$ ha 9 elementi. □

Esercizio 1.13. *Siano* A, B, C *tre insiemi. Provare che* $A \subseteq B \cap C \Leftrightarrow A \subseteq B \wedge A \subseteq C$.

Risoluzione. Supponiamo che $A \subseteq B \cap C$, allora ogni $x \in A$ sta in $B \cap C$ e quindi sta sia in B che in C; ciò prova che $A \subseteq B \cap C \Rightarrow A \subseteq B \wedge A \subseteq C$. Viceversa se ogni elemento di A sta sia in B che in C, sta anche in $B \cap C$ e quindi $A \subseteq B \wedge A \subseteq C \Rightarrow A \subseteq B \cap C$. □

Esercizio 1.14. *Siano* A *e* B *due sottoinsiemi dell'universo* U. *Provare che*

1. $A \subseteq B \Leftrightarrow \complement B \subseteq \complement A$;
2. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$;

3. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$;
4. $(A \cup B) \times C = (A \times C) \cup (B \times C)$;
5. $(A \cap B) \times C = (A \times C) \cap (B \times C)$;

Risoluzione. La prima proprietà è di facile verifica, con una tecnica simile a quella dell'esercizio 1.13. Per le altre si possono costruire delle tavole di verità come quelle considerate nell'esercizio 1.2. □

Esercizio 1.15. *In $\mathbb{Z}$ si consideri la relazione $x\mathcal{R}y$ se $x+y$ è pari. Verificare che è di equivalenza e dire quali sono le classi di equivalenza.*

Risoluzione. Intanto si ha $x\mathcal{R}x$, perché $x+x=2x$ e $2x$ è pari, dunque le relazione è riflessiva. Ovviamente poi se $x\mathcal{R}y$ anche $y\mathcal{R}x$ perché $x+y=y+x$. Infine se $x\mathcal{R}y$ e $y\mathcal{R}z$ significa che x ed y sono entrambi pari o entrambi dispari; lo stesso dicasi per y e z e quindi per x e z, dunque $x\mathcal{R}z$ e la relazione è transitiva. Sia ha poi $[0]=$ insieme degli interi pari, in quanto $0+x$ è pari se e solo se x è pari; inoltre $[1]=$ insieme degli interi dispari, in quanto $1+x$ è pari se e solo se x è dispari. Poiché queste due classi di equivalenza (tra di loro disgiunte, come è ovvio per le classi di equivalenza) costituiscono una partizione di $\mathbb{Z}$, esse devono essere le uniche due classi di equivalenza con questa relazione. □

Esercizio 1.16. *Si consideri, nell'insieme $\mathbb{Q}$, la relazione $x\mathcal{R}y$ se $x^2 \geq y^2$. Si dica se è riflessiva, simmetrica, transitiva.*

Risoluzione. Per ogni $x \in \mathbb{Q}$ si ha $x^2 \geq x^2$, dunque la relazione è riflessiva. Si ha $2\mathcal{R}1$ perché $2^2 \geq 1^2$, ma non $1\mathcal{R}2$ perché $1^2 \not\geq 2^2$, dunque la relazione non è simmetrica. Se $x\mathcal{R}y$ e $y\mathcal{R}z$ si ha $x^2 \geq y^2$ e $y^2 \geq z^2$ dunque anche $x^2 \geq z^2$ e quindi $x\mathcal{R}z$ e la relazione è transitiva. □

Esercizio 1.17. *Un esempio di* falsa dimostrazione*: è molto importante prestare la massima attenzione per non incorrere in errori logici, spesso difficili da scoprire, nelle dimostrazioni. Si consideri la proposizione:* dimostrare che se una relazione $\mathcal{R}$ in un insieme A gode delle proprietà simmetrica e transitiva, gode anche di quella riflessiva, *e la relativa dimostrazione:* siano a, $b \in A$ tali che $a\mathcal{R}b$. Poiché $\mathcal{R}$ è simmetrica si ha anche $b\mathcal{R}a$. Per la transitività si ha allora che $a\mathcal{R}b \wedge b\mathcal{R}a$ implica $a\mathcal{R}a$, cioè che la relazione è riflessiva*. Si chiede di trovare l'errore nella dimostrazione.*

Risoluzione. L'errore consiste nel supporre che, dato $a \in A$, esista sempre $b \in A$ tale che $a\mathcal{R}b$. La cosa non è vera per una generica relazione: possono esistere relazioni in cui qualche elemento non è in relazione con nessun altro elemento. Per questi la dimostrazione proposta non vale e non si può concludere che sono in relazione con se stessi. Il teorema sarebbe vero se, nelle ipotesi, si aggiungesse il fatto che ogni elemento deve essere in relazione con qualche elemento (magari anche solo con se stesso). □

Esercizio 1.18. *Dimostrare che la relazione $x \sim y$, se $y-x$ è multiplo di 7, è una relazione di equivalenza in $\mathbb{Z}$.*

Risoluzione.

— Per ogni $x \in \mathbb{Z}$ si ha $x-x=0=0\times 7$, dunque $x \sim x$.
— Se $x \sim y$, si ha $y-x=7k$, con k intero; poiché $x-y=7(-k)$, si ha anche $y \sim x$.

— Se $x \sim y \wedge y \sim z$, si ha $y - x = 7h \wedge z - y = 7k$, con h, k interi; dunque $z - x = (z - y) + (y - x) = 7(k + h)$ e quindi $x \sim z$.

Se ne deduce che la relazione è di equivalenza. □

Esercizio 1.19. *Siano A e B due insiemi finiti con cardinalità $|A|$ e $|B|$. Provare che l'insieme delle funzioni $f: A \to B$ ha cardinalità $|B|^{|A|}$.*

Risoluzione. Poniamo $m = |A|$ e $n = |B|$. Per definire una funzione f da A in B dobbiamo decidere quale degli n elementi di B dobbiamo associare a ciascuno degli m elementi di A. Ovviamente per ciascuno degli m elementi di A abbiamo n scelte e le scelte sono indipendenti: n scelte per il primo, n scelte per il secondo, …, n scelte per l'm-esimo, dunque in totale n^m scelte.

Da questo fatto ha origine la tradizione di indicare con B^A l'insieme di tutte le funzioni $f: A \to B$, notazione mantenuta anche quando A e B non sono entrambi finiti. □

Esercizio 1.20. *Dire se la relazione tra $\mathbb{Q}$ e $\mathbb{Z}$ che a $p/q \in \mathbb{Q}$ fa corrispondere $p - q \in \mathbb{Z}$ è o no una funzione.*

Risoluzione. Non si tratta di una funzione perché, mentre $1/2 = 2/4$, $1 - 2 \neq 2 - 4$, cioè ad uno stesso elemento di $\mathbb{Q}$ corrispondono diversi elementi di $\mathbb{Z}$. □

2. Numeri: dai naturali ai reali

In questo capitolo richiameremo le principali proprietà degli insiemi dei numeri naturali, interi, razionali e reali. Ci occuperemo delle proprietà più importanti di ciascun insieme e discuteremo anche, se pur brevemente, il percorso logico che porta ad estendere il concetto di numero, a partire dai naturali. Sui numeri naturali ci dilungheremo un po', per l'importanza di fissare le notazioni e perché troppo spesso le conoscenze relative sono superficiali. Trattando i numeri reali presenteremo, con un certo dettaglio, le proprietà dei radicali, vista la loro importanza nelle applicazioni. Per avere un'idea, almeno a livello intuitivo, di come si possano introdurre i numeri reali, presentiamo anche una breve trattazione della rappresentazione decimale dei numeri.

2.1. I numeri naturali

"Dio ha fatto i numeri naturali, tutto il resto è opera dell'uomo": in questa celebre frase di Leopold Kronecker (1823-1891) è contenuta l'idea che i numeri naturali sono un concetto primitivo, in un certo senso innato nell'uomo. Studi abbastanza recenti (vedi per esempio alcuni risultati in Keith Devlin, [5]) provano che anche i bambini neonati hanno la capacità di "contare" fino a due o tre, e che molti animali riescono a farlo, seppure sempre in maniera più approssimata man mano che si cresce oltre il tre.

In realtà il significato della affermazione di Kronecker è un po' più complesso, in quanto Kronecker, in contrasto con Cantor e rifacendosi alle idee pitagoriche, voleva che l'aritmetica e l'analisi venissero basate sui numeri interi: egli rifiutava la costruzione dei numeri reali comunemente usata dai matematici del tempo e invocava una rivoluzione aritmetica che avrebbe dovuto mettere al bando come inesistenti i numeri irrazionali.

Giuseppe Peano (1858-1932) per i suoi fondamenti dell'aritmetica sceglie tre concetti primitivi:

— zero;
— numero (intero non negativo);
— la relazione "essere il successivo di".

Questi concetti soddisfano i seguenti cinque postulati:

1. Zero è un numero.
2. Se a è un numero, il successivo di a è un numero.
3. Zero non è il successivo di nessun numero.
4. Due numeri, i cui successivi sono uguali, sono essi stessi uguali.
5. Se un insieme S di numeri contiene zero e contiene anche il successivo di ogni numero contenuto in S, allora ogni numero è contenuto in S (*assioma di induzione*).

Non saremo naturalmente interessati a discutere in dettaglio la costruzione di Peano dei naturali: l'abbiamo menzionata solamente per dare un'idea del fatto che una sistemazione assiomatica di questo fondamentale concetto non è per niente semplice. Segnaliamo solo che un sistema per costruire operativamente i numeri naturali potrebbe essere quello di considerare la cardinalità dei seguenti insiemi (vedi la discussione nella pagina 28 del Capitolo 1): $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots$

Per i nostri scopi, comunque, quello che interessa sono alcune delle proprietà dei numeri naturali e precisamente:

— il fatto che, come insieme, i naturali hanno una certa cardinalità;
— il fatto che l'insieme dei numeri naturali è un insieme ordinato con certe caratteristiche;
— il fatto che sull'insieme dei naturali si possano eseguire le operazioni di somma e prodotto le quali godono di certe proprietà;
— come conseguenza delle proprietà delle operazioni, la possibilità di risolvere certe equazioni e l'impossibilità di risolverne altre.

2.1.1. La cardinalità dei naturali

L'insieme dei numeri naturali, che indichiamo con $\mathbb{N}$, è un insieme infinito ed è anzi il "più piccolo" degli insiemi infiniti: ogni altro insieme infinito o è equipotente all'insieme dei naturali, o ha un sottoinsieme equipotente all'insieme dei naturali. La cardinalità dell'insieme dei naturali si indica con il simbolo $\aleph_0$. Dunque ogni insieme transfinito ha cardinalità maggiore o uguale ad $\aleph_0$. Un insieme che abbia cardinalità $\aleph_0$ si dice anche *numerabile* e il motivo di questo nome è palese.

2.1.2. L'ordine nei naturali

Nell'insieme $\mathbb{N}$ è definito un ordine totale, detto *ordine naturale* che, nella forma antiriflessiva denotata con "$<$" (più comoda sui naturali che non quello nella forma riflessiva), gode dunque delle seguenti proprietà:

1. *proprietà antiriflessiva*: $\forall n \in \mathbb{N},\ n \not< n$;
2. *proprietà antisimmetrica*: $\forall m, n \in \mathbb{N},\ (m < n \Rightarrow n \not< m)$;
3. *proprietà transitiva*: se $m, n, p \in \mathbb{N},\ (m < n) \wedge (n < p) \Rightarrow (m < p)$;
4. *proprietà di tricotomia*: dati due naturali m ed n si ha $m < n$, oppure $m = n$, oppure $n < m$.

L'ordine gode inoltre delle seguenti ulteriori proprietà:

5. ogni numero naturale n ha un immediato seguente, $n + 1$;
6. ogni numero naturale $n > 0$ ha un immediato precedente, $n - 1$;
7. ogni sottoinsieme non vuoto di $\mathbb{N}$ ha un minimo, in particolare $\mathbb{N}$ ha 0 come minimo;
8. ogni insieme non vuoto e superiormente limitato di $\mathbb{N}$ ha un massimo.

Sussiste poi il *Principio di induzione* di cui però in questo testo non ci occuperemo.

2.1.3. Le operazioni nei naturali

Cominciamo con il richiamare la nozione di *operazione (interna)* su un insieme.

Definizione 2.1 (Operazione interna). *Dato un insieme A, si chiama* operazione (interna) *in A ogni funzione $\varphi\colon A \times A \to A$. Le funzioni di questo tipo si sogliono, abitualmente, indicare con simboli come $\circ$, $+$, $\times$, $/$, ecc.*

Ricordiamo anche che, invece di scrivere $z = \varphi(x, y)$ per indicare l'immagine della coppia (x, y) tramite la funzione (operazione) φ, si scrive $z = x + y$, $z = x \times y$, ecc., ovvero il simbolo che rappresenta la funzione si interpone tra i due elementi della coppia su cui agisce. In molti casi non si utilizza alcun simbolo, come per esempio nella moltiplicazione dove i due elementi della coppia vengono semplicemente scritti uno di fianco all'altro: $z = xy$. Ci sono anche altri modi di indicare le operazioni, ormai entrati nell'uso comune. Un esempio importante è quello dell'elevazione a potenza, dove il simbolo usuale, a tutti noto, è x^y.

Nell'insieme dei naturali sono anzitutto definite le operazioni di *addizione* e *moltiplicazione*: nell'addizione il risultato dell'operazione si chiama *somma* dei due numeri, che prendono il nome di *addendi*; se gli addendi sono a e b, la somma di indica con $a + b$. Nella moltiplicazione il risultato della operazione si chiama *prodotto* dei due numeri, che prendono il nome di *fattori*; se i fattori sono a e b, il prodotto si indica con $a \times b$, oppure con $a \cdot b$, oppure più semplicemente con ab[1]. Queste operazioni godono delle seguenti importanti proprietà.

1. Proprietà *associativa*:
$$\forall a, b, c \in \mathbb{N}, \quad (a+b)+c = a+(b+c); \quad (ab)c = a(bc);$$
questa proprietà consente di evitare l'uso delle parentesi nelle addizioni o moltiplicazioni ripetute: hanno dunque senso le scritture $a + b + c$ e abc.
2. Proprietà *commutativa*:
$$\forall a, b \in \mathbb{N}, \quad a+b = b+a; \quad ab = ba.$$
3. Esistenza dell'*elemento neutro*[2]:
$$\forall a \in \mathbb{N}, \quad a+0 = a; \quad a1 = a.$$
4. Proprietà *distributiva della moltiplicazione rispetto all'addizione*:
$$\forall a, b, c \in \mathbb{N}, \quad a(b+c) = ab + ac.$$
5. Legge dell'*annullamento del prodotto*:
$$ab = 0 \Leftrightarrow (a = 0) \vee (b = 0).$$
6. Legge di *cancellazione della somma*:
$$a + c = b + c \Leftrightarrow a = b.$$

[1] Le operazioni che stiamo trattando si chiamano *addizione* e *moltiplicazione*, mentre *somma* e *prodotto* si riferiscono al risultato delle operazioni. Tuttavia spesso si usano i termini somma e prodotto anche per indicare le operazioni, anche se si tratta di un abuso di linguaggio. Anche noi ogni tanto faremo così.

[2] Non tutti concordano nel ritenere che l'insieme dei naturali comprenda lo zero. Il discorso che abbiamo fatto a proposito di una possibile costruzione dei naturali come cardinalità di certi insiemi, rende opportuna invece l'inclusione dello zero in $\mathbb{N}$. Tuttavia esistono altre considerazioni che porterebbero alla sua esclusione, e dunque il fare una scelta o l'altra è sostanzialmente questione di gusto: per questo è molto importante controllare, nell'elenco delle notazioni, qual è la convenzione seguita nel testo che si sta consultando.

7. Legge di *cancellazione del prodotto*:

$$(c \neq 0) \Rightarrow (ac = bc \Leftrightarrow a = b).$$

8. *Compatibilità* dell'ordine con l'addizione:

$$a < b \Leftrightarrow a + c < b + c$$

9. *Compatibilità* dell'ordine con la moltiplicazione:

$$(c > 0) \Rightarrow (a < b \Leftrightarrow ac < bc).$$

L'introduzione delle due operazioni di addizione e sottrazione rende l'insieme dei numeri naturali una *struttura algebrica*, molto semplice: si tratta precisamente di un *monoide* o *semigruppo abeliano* dotato di unità rispetto ad entrambe le operazioni. La necessità di avere a disposizione strutture algebriche sempre più ricche è uno dei motivi che portano ad ampliare l'insieme dei numeri in uso nella matematica.

2.1.4. Le prime equazioni nei naturali

L'introduzione delle operazioni di addizione e moltiplicazione ci conduce immediatamente a considerare problemi come i due seguenti:

1. dati due numeri naturali a e b determinare, se esiste, un numero naturale x tale che $x + b = a$;
2. dati due numeri naturali a e b determinare, se esiste, un numero naturale x tale che $x \cdot b = a$.

Si tratta dei primi esempi di *equazioni in una incognita*: esse conducono, rispettivamente, all'introduzione delle operazioni di *sottrazione* e di *divisione*, anche se il termine "operazione" è inappropriato, secondo la definizione 2.1 che abbiamo dato; infatti né la sottrazione, né la divisione sono definite in tutto $\mathbb{N} \times \mathbb{N}$. Si danno precisamente le due seguenti definizioni.

Definizione 2.2 (Sottrazione). *Dati i numeri naturali a e b, con $a \geq b$, si chiama* differenza *di a e b (nell'ordine!) l'unico numero x che risolve l'equazione $x + b = a$. Tale numero x si indica con $a - b$. Il numero a si chiama* minuendo, *il numero b si chiama* sottraendo.

Definizione 2.3 (Divisione). *Dati i numeri naturali a e b, con $b \neq 0$, si chiama* quoziente *di a e b (nell'ordine!) l'unico numero x,* se esiste, *che risolve l'equazione $x \cdot b = a$. Tale numero x si indica con a/b o con $a : b$. Il numero a si chiama* dividendo, *il numero b si chiama* divisore.

Si tengano ben presenti le seguenti importanti osservazioni.

1. La sottrazione è possibile solo se il minuendo è maggiore o uguale al sottraendo.
2. La divisione non è mai possibile se il divisore è zero, *nemmeno* se il dividendo è zero: non ha alcun senso la scrittura $a/0$, e men che mai la scrittura $0/0$[3].

[3] Spesso, soprattutto sui testi elementari, si trova scritto che $0/0$ è indeterminato. Una tale affermazione potrebbe anche avere un senso in quanto l'equazione $x \cdot 0 = 0$ è verificata qualunque sia il numero x. È però preferibile limitarsi a fornire la soluzione (se c'è!) dell'equazione $x \cdot b$ solo ed esclusivamente quando $b \neq 0$, e questo per una serie di motivi che qui non possiamo riportare.

3. La divisione (con divisore diverso da zero) non è sempre possibile: quando lo è, si dice che a è *multiplo* di b, o anche che o anche che b è un divisore di a, o ancora che b *divide* a e si scrive $b \mid a$.
4. La sottrazione e la divisione *non* godono della proprietà associativa:

$$(a-b)-c \neq (a-(b-c); \quad (a/b)/c \neq a/(b/c).$$

5. La sottrazione e la divisione *non* godono della proprietà commutativa:

$$a-b \neq b-a; \quad a/b \neq b/a.$$

Si noti che la mancanza della proprietà commutativa è evidenziata anche dalla nomenclatura utilizzata: mentre nella addizione e moltiplicazione i due "operandi" hanno lo stesso nome (addendi o fattori), nella sottrazione e divisione no, e questo proprio per il fatto che non vale la proprietà commutativa.

L'impossibilità di risolvere in ogni caso le equazioni qui considerate è una delle ragioni che spingono ad ampliare l'insieme di numeri che si usano in matematica.

2.1.5. L'elevazione a potenza

In $\mathbb{N}$ si definisce anche l'operazione di *elevamento a potenza*, parzialmente come semplice abbreviazione della scrittura di certi prodotti. Precisamente si dà la definizione che segue.

Definizione 2.4 (Potenza). *Dati a e $n \in \mathbb{N}$, con $n > 0$ si pone*

$$a^1 = a; \quad a^n = \underbrace{a \cdot a \cdot \cdots \cdot a}_{n \text{ fattori}}, \text{ se } n > 1. \tag{2.1}$$

Si pone inoltre

$$a^0 = 1, \forall a > 0; \quad 0^n = 0, \forall n > 0. \tag{2.2}$$

Nella scrittura a^n, il numero a si chiama base, *il numero n si chiama* esponente.

Le formule (2.1) e (2.2) potrebbero anche essere riscritte, in maniera ricorsiva, come segue:

$$\text{Se } a > 0, a^0 = 1, a^{n+1} = a \cdot a^n; \quad \text{se } n > 0, 0^n = 0. \tag{2.3}$$

Si noti che al simbolo 0^0 *non* è attribuito alcun significato. Si noti altresì che la potenza è solo parzialmente un "modo compatto" per scrivere un prodotto con fattori tutti uguali: i simboli a^1 ed a^0 non sono in alcun modo legati all'operazione di moltiplicazione. Torneremo nel capitolo 7, in cui richiameremo le proprietà delle potenze, sui motivi che giustificano queste definizioni.

Sussistono le seguenti proprietà formali delle potenze. Se a, b, m, n sono naturali maggiori di 0 si ha

$$a^n a^m = a^{n+m}; \quad (a^n)^m = a^{nm}; \quad a^n b^n = (ab)^n; \quad \text{se } n \geq m, \ \frac{a^n}{a^m} = a^{n-m}. \tag{2.4}$$

Quando hanno senso, alcune delle (2.4) si estendono anche ai casi in cui qualcuno dei numeri che vi compaiono sono zero.

2.1.6. La precedenza nelle operazioni successive

Nel caso di più operazioni da eseguirsi su numeri, la regola base è che esse si eseguono nell'ordine in cui si presentano, a meno che non intervengano delle parentesi ad alterare questo "ordine naturale".

Per esempio la scrittura $24 : 6 : 2$ dà come risultato 2, in quanto si esegue prima $24 : 6 = 4$ e successivamente $4 : 2 = 2$. Si otterrebbe lo stesso risultato scrivendo $(24 : 6) : 2$. La scrittura $24 : (6 : 2)$ dà invece come risultato 8, in quanto la presenza delle parentesi impone di eseguire per primo la divisione $6 : 2 = 3$ e successivamente la $24 : 3 = 8$.

Per evitare di usare troppe parentesi si è comunque convenuto di assegnare ad alcune operazioni una precedenza sulle altre. Per quanto riguarda le operazioni fin qui definite si ha precisamente:

1. l'elevazione a potenza deve essere eseguita per prima;
2. successivamente devono essere eseguite le moltiplicazioni e divisioni nell'ordine in cui si presentano;
3. infine devono essere eseguite le addizioni e sottrazioni.

Per esempio la scrittura $3 + 5 \times 2^3$ dà come risultato 43 in quanto si esegue dapprima $2^3 = 8$, poi $5 \times 8 = 40$ e successivamente $3 + 40 = 43$.

Al di fuori di queste regole le operazioni si eseguono nell'ordine in cui si presentano. Per esempio $12 : 6 \times 2$ dà come risultato 4, in quanto non esiste precedenza tra divisione e moltiplicazione.

C'è comunque una situazione che è opportuno segnalare perché fonte di ambiguità, e riguarda l'elevazione a potenza con scritture (molto comuni in tutti i testi!) del tipo

$$a^{b^c}.$$

Per capire meglio, ragioniamo su un esempio concreto. Le due scritture seguenti

$$(2^3)^2 \quad \text{e} \quad 2^{(3^2)}$$

non pongono alcun problema interpretativo e forniscono come risultato, rispettivamente, 64 e 512. Ma la scrittura

$$2^{3^2}$$

come va interpretata? Purtroppo *non* c'è una convenzione univoca su questo: per alcuni significa $(2^3)^2 = 64$, per altri $2^{(3^2)} = 512$, e questo anche sui software di calcolo simbolico e sulle calcolatrici tascabili. Basta provare a scrivere 2^3^2 su una calcolatrice tascabile (e di solito si ottiene 64) o su un software come Mathematica, Maple, Geogebra (dove si ottiene 512).

Noi adotteremo la convenzione più diffusa e cioè

$$a^{b^c} = a^{(b^c)}, \tag{2.5}$$

che equivale a dire che $a\hat{}b\hat{}c = a\hat{}(b\hat{}c)$, ovvero che l'elevazione a potenza è, come si usa dire, *associativa a destra*. Questa nomenclatura serve a richiamare la differenza con, per esempio, la divisione, per la quale, come già osservato, si ha

$$a/b/c = (a/b)/c,$$

situazione che si esprime a parole dicendo che la divisione è *associativa a sinistra*.

2.1.7. Divisibilità. Numeri primi

Nell'insieme dei numeri naturali è definita un'operazione di *divisione con resto*. Segnaliamo che, come già per la sottrazione e divisione, non si tratta di una operazione nel vero senso della parola, in quanto non è un'applicazione di $\mathbb{N} \times \mathbb{N}$ in $\mathbb{N}$. Essa si basa sul seguente teorema.

Teorema 2.5. *Dati due numeri naturali a e b, con $b > 0$, esiste una e una sola coppia (q, r) di numeri naturali tali che*

1. $a = qb + r$,
2. $(0 \leq)\, r < b$.

I numeri q ed r prendono rispettivamente il nome di quoziente *e* resto *della divisione di a per b.*

Osserviamo che se è $r = 0$, significa che la divisione tra a e b è possibile, nel senso della definizione 2.3, e si usano i termini di multiplo, divisibile e divisore, già allora introdotti.

Se a è divisibile per b si dice anche che b *divide* a e si usa la seguente scrittura

$$b \mid a\,. \tag{2.6}$$

La determinazione di q ed r si fa con il classico algoritmo della divisione, noto a tutti fin dai tempi della scuola primaria.

Definizione 2.6 (Numero primo)**.** *Un numero naturale p si dice* primo *se è maggiore di* 1 *ed è divisibile solo per se stesso e per* 1.

È immediato constatare che 2 è il più piccolo primo ed è anche l'unico primo pari. I successivi primi sono $3, 5, 7, \ldots$. I numeri primi giocano un'importanza cruciale in moltissime applicazioni: segnaliamo solo, a titolo d'esempio, il loro uso nella moderna crittografia.

I numeri primi possono essere considerati i "mattoni" con cui sono costruiti tutti i numeri naturali, nel senso precisato dal seguente teorema.

Teorema 2.7 (Teorema fondamentale dell'aritmetica)**.** *Ogni numero naturale maggiore di* 1 *o è primo o è esprimibile come prodotto di fattori primi e tale espressione è unica a meno dell'ordine in cui i fattori sono scritti.*

Questo teorema si può compendiare nella formula seguente

$$\forall n \in \mathbb{N},\, n > 1, \quad n = p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}\,, \tag{2.7}$$

dove abbiamo raggruppato, usando l'operazione di potenza, tutti i fattori primi tra di loro uguali. Si può naturalmente supporre che i fattori che compaiono nella formula (2.7) siano scritti in ordine crescente. Si ha, per esempio, $28\,693\,665 = 3^2 \cdot 5 \cdot 7^3 \cdot 11 \cdot 13^2$.

Il contenuto di questo teorema è tutt'altro che banale (e la dimostrazione dell'unicità della decomposizione è abbastanza difficile). La verifica concreta della decomponibilità di un numero in un prodotto di primi può richiedere anni di lavoro anche su potenti computer. A questo proposito è rimasto famoso un articolo di Martin Gardner pubblicato su *Scientific American* nell'agosto del 1977. In relazione a un problema di crittografia, il problema proposto nell'articolo era quello di scomporre in fattori primi un numero di 129 cifre, e precisamente

$$
\begin{aligned}
n = {} & 114381625757888867669235779976146612010218296721242362562561 8\\
& 42935706935245733897830597123563958705058989075147599290026879543541\,.
\end{aligned}
$$

Nell'aprile del 1994, solo[4] 17 anni dopo, quella scomposizione fu scoperta: i due numeri p e q, tali che $n = pq$ sono

$$
\begin{aligned}
p &= 3490529510847650949147849619903898133417764638493387843990820577,\\
q &= 32769132993266709549961988190834461413177642967992942539798288533,
\end{aligned}
$$

il primo di 64 e il secondo di 65 cifre. Oggi si ritiene che un numero di 500 cifre possa richiedere oltre un secolo di calcoli per decidere se è primo o no.

I numeri primi diventano sempre più radi man mano che si considerano naturali sempre più grandi. Per esempio ci sono

- 168 primi tra 1 e 1000;
- 135 primi tra 1001 e 2000;
- 127 primi tra 2001 e 3000;
- 120 primi tra 3001 e 4000;
- 119 primi tra 4001 e 5000.

Tuttavia i primi sono infiniti, come prova un celeberrimo teorema dovuto ad Euclide, di cui riportiamo la dimostrazione perché è considerata una delle più eleganti dimostrazioni in matematica.

Teorema 2.8 (Infinità dei primi)**.** *Esistono infiniti numeri primi.*

Dimostrazione. Supponiamo che esista solo un sistema finito di primi: $p_1, p_2, p_3, \ldots, p_k$, e consideriamo il naturale

$$n = p_1 p_2 p_3 \cdots p_k + 1\,.$$

Questo numero non è divisibile per nessuno dei primi $p_1, p_2, p_3, \ldots, p_k$, in quanto il resto della divisione di n per uno di questi primi dà 1. Dunque o è primo (e dunque è assurdo che i primi siano solo quelli elencati), o deve essere divisibile per qualche primo diverso da quelli elencati, per il teorema fondamentale dell'aritmetica, e quindi è ancora assurdo supporre che i primi siano solo quelli elencati. □

Si noti che questa dimostrazione non produce un metodo per costruire nuovi numeri primi a partire da un dato insieme di primi. Per esempio, dati i primi 2, 3, 5, 7, 11, 13 il numero che si costruisce a partire da essi con l'algoritmo di Euclide, e cioè

$$2 \times 3 \times 5 \times 7 \times 11 \times 13 + 1 = 30031\,,$$

non è primo, ma è il prodotto dei primi 59 e 509 che non stanno nella lista precedente. La ricerca di formule per la generazione di numeri primi ha occupato l'attenzione di molti matematici, specie nel

[4]In realtà quando l'articolo fu pubblicato si riteneva necessario un tempo molto più lungo per questo scopo. L'avvento di internet e la collaborazione di 600 volontari di una ventina di paesi sotto la guida di un gruppo di ricercatori permisero invece la risoluzione dell'enigma in un tempo relativamente breve (circa otto mesi di lavoro).

diciassettesimo secolo. Un famoso esempio, dovuto ad Eulero, è quello della formula $x^2 + x + 41$ che produce numeri primi per tutti i naturali $x \leq 39$, mentre non produce un primo per $x = 40$. Non possiamo però dilungarci oltre su questo argomento.

Definizione 2.9 (Numeri primi tra di loro)**.** *Due naturali non nulli si dicono* primi tra di loro *o* mutuamente primi *se* 1 *è l'unico loro divisore comune.*

Il numero 1 è naturalmente primo con ogni altro numero naturale. È ovvio che due naturali a e b sono primi tra di loro se e solo se non esiste alcun fattore primo comune tra la scomposizione di a e di b. Se due numeri a e b non sono primi tra di loro ha grande interesse la determinazione del loro *massimo comun divisore*, cioè del massimo (certamente esistente perché non maggiore né di a né di b) dei loro divisori comuni, a volte indicato con $\mathrm{MCD}(a, b)$. Se essi sono primi tra di loro, questo massimo comun divisore è chiaramente 1. L'algoritmo più elementare per trovare il massimo comun divisore è quello di scomporre i due numeri in fattori primi e poi di fare il prodotto dei fattori comuni con il minimo esponente. Tuttavia esiste un algoritmo molto efficiente, dovuto ad Euclide, per la determinazione del massimo comun divisore, algoritmo presentato nel teorema che segue.

Teorema 2.10 (Algoritmo euclideo delle divisioni successive)**.** *Sia a e b (in cui possiamo supporre, senza perdere di generalità, $a > b$) e si eseguano le seguenti successive divisioni con resto*

$$
\begin{aligned}
a &= d_1 b + r_1 && \text{con } 0 \leq r_1 < b \\
b &= d_2 r_1 + r_2 && \text{con } 0 \leq r_2 < r_1 \\
r_1 &= d_3 r_2 + r_3 && \text{con } 0 \leq r_3 < r_2 \\
&\vdots && \vdots
\end{aligned}
$$

fin quando si ottiene resto zero (dopo meno di b passi). Il massimo comun divisore tra a e b è l'ultimo resto non nullo.

Esempio 2.1. Siano dati i numeri 326059 e 164143. Procedendo come indicato nel teorema 2.10 si ottiene:

$$
\begin{aligned}
326059 &= 1 \times 164143 + 161916, \\
164143 &= 1 \times 161916 + 2227, \\
161916 &= 72 \times 2227 + 1572, \\
2227 &= 1 \times 1572 + 655, \\
1572 &= 2 \times 655 + 262, \\
655 &= 2 \times 262 + 131, \\
262 &= 2 \times 131 + 0,
\end{aligned}
$$

da cui si deduce che il massimo comun divisore tra i numeri dati è 131.

Due numeri naturali a e b non nulli hanno sempre un multiplo comune, per esempio il loro prodotto. Nelle applicazioni ha interesse il più piccolo di questi multipli comuni (certamente esistente per le specifiche proprietà dell'ordine nei naturali), detto il *minimo comune multiplo*, a volte indicato con $\mathrm{mcm}(a, b)$. L'algoritmo più elementare per trovare il minimo comune multiplo è quello di scomporre i due numeri in fattori primi e poi di fare il prodotto dei fattori comuni e non comuni con il massimo esponente. La ricerca si può fare però anche determinando il massimo comun divisore con l'algoritmo euclideo delle divisioni successive e osservando poi che, in base alla definizione, si ha facilmente che $ab = \mathrm{MCD}(a,b)\,\mathrm{mcm}(a,b)$.

Esempio 2.2. Il minimo comune multiplo dei due numeri 326059 e 164143 già considerati nell'esempio precedente è dato da

$$\mathrm{mcm}(326059, 164143) = \frac{326059 \times 164143}{131} = 408551927.$$

2.1.8. Qualche criterio di divisibilità

Valgono i seguenti criteri di divisibilità[5] la maggior parte dei quali dimostrabili in maniera elementare.

1. Un numero è divisibile per 2 se e solo se è pari.
2. Un numero è divisibile per 3 se e solo se è divisibile per 3 la somma delle sue cifre.
3. Un numero è divisibile per 4 se e solo se è divisibile per 4 il numero formato dalle ultime due cifre.
4. Un numero è divisibile per 5 se e solo se termina per 5 o per 0.
5. Un numero è divisibile per 8 se e solo se è divisibile per 8 il numero formato dalle sue ultime tre cifre.
6. Un numero è divisibile per 9 se e solo se è divisibile per 9 la somma delle sue cifre.
7. Un numero è divisibile per 10 se e solo se termina per 0.
8. Un numero è divisibile per 11 se e solo se è divisibile per 11 il numero intero dato dalla somma delle cifre di posto pari meno quello dato dalla somma delle cifre di posto dispari. Di solito si parte dall'ultima cifra (cifra delle unità), che è di posto "0", in quanto corrisponde alla potenza 0 di 10.

Esistono anche altri criteri di divisibilità (per esempio per 7) ma sono più complicati e di uso poco frequente.

2.2. Gli interi

2.2.1. Generalità

A partire dall'insieme dei numeri naturali si può costruire l'insieme degli interi, detti anche interi relativi, insieme che è dotato di una struttura algebrica più raffinata di quella dei naturali e dove almeno le equazioni del tipo $x + b = a$ possono essere risolte senza limitazioni di sorta.

Non discuteremo la tecnica con cui l'insieme degli interi viene costruito, ci limitiamo a segnalare che esso è costituito dai numeri

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

cioè dai numeri naturali a cui abbiamo aggiunto i negativi. È però opportuno segnalare che gli interi *non* sono l'insieme delle coppie costituite da un segno e da un numero naturale: se così fosse, il numero $+0$ e il numero -0 sarebbero diversi, cosa che invece non è, come ben sappiamo. Si noti che l'insieme

[5] In questi criteri di divisibilità naturalmente si sottintende che il numero sia scritto in forma decimale

dei numeri naturali può essere considerato un sottoinsieme dell'insieme degli interi, anche se questa affermazione andrebbe meglio precisata.

In $\mathbb{Z}$ è definita una relazione d'ordine totale e due operazioni come in $\mathbb{N}$, con le proprietà già citate, salvo le seguenti modifiche e aggiunte.

1. Per quanto riguarda l'ordine:
   a) ogni numero intero ha un immediato precedente e un immediato seguente;
   b) $\mathbb{Z}$ non ha né massimo né minimo, ma ogni sottoinsieme non vuoto e inferiormente limitato ha minimo, mentre ogni sottoinsieme non vuoto e superiormente limitato ha massimo.
2. Per quanto riguarda le operazioni:
   a) la compatibilità dell'ordine con la moltiplicazione deve essere modificata come segue
   $$(c>0) \Rightarrow (a<b \Leftrightarrow ac<bc); \quad (c<0) \Rightarrow (a<b \Leftrightarrow ac>bc);$$
   b) esiste l'*opposto* nell'addizione
   $$\forall x \in \mathbb{Z}, \exists y \in \mathbb{Z} \text{ tale che } x+y=0,$$
   e tale y si indica con $-x$. Inoltre invece di scrivere $x+(-x)$ si scrive $x-x$.

L'esistenza dell'opposto ha come conseguenze importanti che $\mathbb{Z}$ diventa un gruppo abeliano rispetto alla somma (mentre mantiene le stesse caratteristiche rispetto al prodotto) e che l'equazione $x+b=a$ ha sempre una e una sola soluzione $x=a+(-b)$, che si scrive $a-b$.

Per quanto riguarda la cardinalità nulla cambia rispetto all'insieme dei naturali: dato $n \in \mathbb{N}$, basta considerare la corrispondenza, chiaramente biunivoca, che ad associa ad ogni naturale pari $2n$ l'intero positivo che n, e ad ogni naturale dispari $2n+1$ l'intero $-n-1$.

In $\mathbb{Z}$ non c'è più la necessità di introdurre l'operazione di sottrazione che può essere ridotta alla somma: dati a e b in $\mathbb{Z}$, si può porre, per definizione, $a-b=a+(-b)$.

Anche negli interi relativi risultano invece definite una divisione e una divisione con resto con le stesse proprietà già viste in $\mathbb{N}$. In particolare vale il seguente teorema, quasi identico al teorema 2.5.

Teorema 2.11. *Dati due numeri interi a e b, con $b>0$, esiste una e una sola coppia (q,r) di numeri interi tali che*

1. $a=qb+r$,
2. $0 \le r < b$.

I numeri q ed r prendono gli stessi nomi di quoziente e resto già usati nel caso dei naturali.

L'unica cosa da segnalare è che, mentre nel teorema 2.5 il fatto che r fosse ≥ 0 era evidente (tutti i naturali sono ≥ 0), ora fa parte dell'enunciato del teorema. Si ha, per esempio,
$$-87=-10\times 9+3.$$
Si noti che è essenziale la richiesta che r sia ≥ 0 (oltreché $<b$) per l'unicità della coppia (q,r). Senza questa richiesta potrei anche avere
$$-87=-9\times 9-6.$$

Un'ultima osservazione prima di chiudere questo richiamo sugli interi è che l'insieme degli interi dispari è chiuso rispetto alla moltiplicazione (il prodotto di due dispari è dispari) e analogamente l'insieme degli interi pari (il prodotto di due pari è pari): si tratta di proprietà quasi ovvie, ma utili nella risoluzione di molti quesiti.

2.2.2. Potenze negli interi

Il concetto di potenza, già introdotto nei naturali, può essere esteso agli interi con la seguente definizione.

Definizione 2.12. *Se $a \in \mathbb{Z}$, $a \neq 0$ e $n \in \mathbb{Z}$, $n < 0$ si pone, per definizione,*

$$a^n = \frac{1}{a^{-n}}. \tag{2.8}$$

Questa definizione è data in modo da mantenere le proprietà formali delle potenze, già considerate nella (2.4), nella pagina 41: anzi, con questa definizione, l'ultima delle (2.4) diventa valida qualunque siano i valori di n ed m, senza dover più supporre $n \geq m$. Si noti che, in base a questa definizione, *non* si attribuisce alcun significato al simbolo 0^m, con $m \leq 0$.

2.3. I numeri razionali

2.3.1. Generalità

L'introduzione degli interi consente di risolvere tutte le equazioni $x + b = a$, mentre non si ha alcun sensibile miglioramento per quanto riguarda la soluzione delle equazioni del tipo $xb = a$. Per fare questo si introducono i numeri razionali, indicati con $\mathbb{Q}$, a partire dall'insieme di tutte le frazioni costituite da due interi, di cui il secondo (quello al denominatore) diverso da zero. In questo insieme delle frazioni si introduce la relazione di equivalenza

$$\frac{a}{b} \sim \frac{c}{d} \Leftrightarrow ad = bc \tag{2.9}$$

le cui classi di equivalenza sono, appunto, i numeri razionali. Per scrivere i numeri razionali si dovrebbe dunque usare la simbologia delle classi di equivalenza; per esempio non si dovrebbe dire: "si consideri il numero razionale $1/2$", ma: "si consideri il numero razionale $[1/2]$" e tenere conto che, in realtà,

$$\left[\frac{1}{2}\right] = \left\{\frac{1}{2}, \frac{2}{4}, \frac{-4}{-8}, \dots\right\}.$$

Nella pratica però tutti si riferiscono ai razionali pensando alle frazioni, tenendo però conto, tutte le volte che serve, che in realtà si tratta di classi di equivalenza. Questo fatto risulta evidente quando si definisce, per esempio, la somma di due razionali:

$$\frac{1}{2} + \frac{3}{5} = \frac{5}{10} + \frac{6}{10} = \frac{11}{10}.$$

L'insieme dei numeri razionali può essere visto come un soprainsieme degli interi: basta pensare agli interi come alle frazioni con denominatore 1.

In $\mathbb{Q}$ si introducono un ordine totale e le operazioni di addizione e moltiplicazione in cui valgono le proprietà già viste in $\mathbb{Z}$, con alcune importanti modifiche e aggiunte che riportiamo qui di seguito. Le modifiche per l'ordine, in particolare, sono decisamente significative.

1. Per quanto riguarda l'ordine:

a) gli elementi di $\mathbb{Q}$ non hanno né un immediato seguente né un immediato precedente;
b) non solo $\mathbb{Q}$ non ha né massimo né minimo, ma anche gli insiemi superiormente limitati di razionali possono non avere massimo e, quel che è ancora peggio, nemmeno estremo superiore; analogamente gli insiemi inferiormente limitati di numeri razionali possono non avere minimo e, quel che è ancora peggio, nemmeno estremo inferiore.

2. Per quanto riguarda le operazioni:
   a) esiste il reciproco nella moltiplicazione:

$$\forall x \in \mathbb{Q},\ x \neq 0,\ \exists y \in \mathbb{Q} \text{ tale che } xy = 1,$$

   e tale y si indica con $1/x$.

L'esistenza del reciproco nella moltiplicazione ha come conseguenza importante che l'insieme $\mathbb{Q} \setminus \{0\}$ è un gruppo abeliano rispetto alla moltiplicazione: la struttura che si ottiene si chiama ora un *corpo commutativo*, o anche *campo ordinato*. L'esistenza del reciproco rende anche possibile risolvere in $\mathbb{Q}$ tutte le equazioni del tipo $xb = a$, con $b \neq 0$: la sua unica soluzione è $x = a \cdot 1/b = a/b$.

In $\mathbb{Q}$ non c'è più la necessità di introdurre né un'operazione di divisione né un'operazione di divisione con resto: qualunque siano i numeri razionali a e b, con $b \neq 0$, si definisce il quoziente tra a e b come il prodotto di a per il reciproco di b.

Come già segnalato, le modifiche riguardanti l'ordine, rispetto ai naturali e agli interi, sono molto significative. In particolare vale il seguente teorema.

Teorema 2.13. *Il campo dei razionali è* denso*, cioè fra due razionali qualunque esiste sempre un altro razionale (anzi ne esistono infiniti).*

Dimostrazione. Dati a e b con $a < b$, basta considerare il numero

$$\frac{a+b}{2},$$

che risulta essere compreso tra a e b. □

Il concetto di potenza con esponente intero si potrà estendere ai razionali senza alcun cambiamento, né nella definizione, né nelle proprietà. È possibile estendere il concetto di potenza consentendo che anche l'esponente possa essere razionale, ma per fare questo, come vedremo, è necessario introdurre i numeri reali. In ogni caso ce ne occuperemo in dettaglio nel capitolo 7 appositamente dedicato alle potenze.

2.3.2. La cardinalità dei razionali

Nonostante il teorema 2.13 faccia pensare che l'insieme dei razionali possa contenere molti più elementi che non l'insieme degli interi, in realtà la cardinalità dei razionali è ancora $\aleph_0$ e la dimostrazione di questo fatto è particolarmente significativa utilizzando il famoso "procedimento diagonale" di Cantor. Proveremo addirittura che l'insieme delle frazioni con denominatore positivo ha cardinalità $\aleph_0$, da cui discenderà subito che la cosa è vera anche per i razionali. Si esamini la figura 2.1 nella quale, per questioni grafiche, abbiamo scritto le frazioni come coppia di interi (di cui il secondo maggiore di zero).

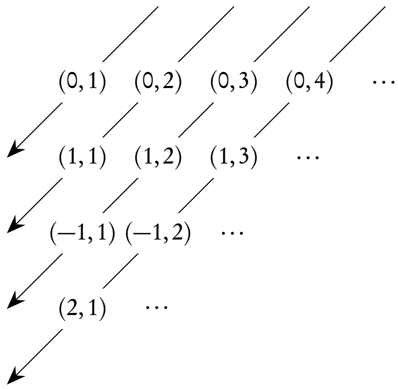


Figura 2.1.: *Numerabilità dei razionali*

La figura fa immediatamente capire quale sia la corrispondenza biunivoca che abbiamo in mente di realizzare tra i naturali e le frazioni:

$$\begin{aligned} 0 &\longleftrightarrow (0,1) \\ 1 &\longleftrightarrow (0,2) \\ 2 &\longleftrightarrow (1,1) \\ 3 &\longleftrightarrow (0,3) \\ 4 &\longleftrightarrow (1,2) \\ 5 &\longleftrightarrow (-1,1) \\ &\vdots \end{aligned}$$

Questo risultato è, a prima vista, abbastanza sorprendente: si pensi solo al fatto che tra due interi qualunque c'è sempre un numero finito di interi, mentre tra due razionali qualunque c'è sempre un numero infinito di razionali.

2.3.3. Perché i razionali non bastano?

Dal punto di vista della struttura algebrica e quindi della risolubilità delle equazioni che abbiamo considerato, l'insieme dei razionali è perfettamente soddisfacente. Che motivo c'è, allora di complicarsi ulteriormente la vita? La scoperta della "lacunosità" di questo insieme numerico è, storicamente, legata alla scoperta del famoso teorema di Pitagora: se un triangolo rettangolo ha i cateti di lunghezza 1, quanto vale la sua ipotenusa? Detto in altri termini: se ho un quadrato di lato 1, e quindi di area 1, quanto dovrà essere il lato di un quadrato che ha area doppia?

Usando il linguaggio delle equazioni questo problema è equivalente a quello di risolvere l'equazione $x^2 = 2$. Questa equazione non può avere soluzioni in $\mathbb{Q}$. Se infatti la frazione m/n fosse una soluzione, si dovrebbe anche avere $m^2 = 2n^2$. Ma questo non è possibile perché scomponendo m ed n in fattori primi avrei, in questa uguaglianza, un diverso numero di fattori uguali a 2 a sinistra e a destra dell'uguale (m^2 ed n^2 devono avere, essendo quadrati, un numero pari di fattori uguali a 2, magari nessuno, dunque a destra c'è comunque un numero dispari di fattori uguali a 2).

È molto importante segnalare che il problema in questione è squisitamente teorico: da un punto di vista pratico è possibile trovare una frazione che esprima l'ipotenusa del triangolo rettangolo di cateti uguali ad 1 con un'approssimazione "grande quanto si vuole". Per esempio, come è ampiamente noto, $1414/1000$ esprime la cercata ipotenusa con le prime tre cifre decimali esatte (il che è sufficiente per la maggior parte delle applicazioni).

Strettamente collegato a questo fatto (anche se a questo livello non è facile capire qual è il legame) è il fatto che un insieme superiormente limitato di razionali può non avere né massimo né estremo superiore.

La risoluzione corretta di questi problemi porta ad una nuova estensione dell'insieme di numeri che conduce all'insieme dei numeri reali. Questa estensione produrrà un insieme sostanzialmente più ricco e complesso rispetto a quelli finora considerati. Per avere un'idea almeno intuitiva di come si possa costruire questo insieme di numeri, trattiamo brevemente il problema, interessante di per sè, della rappresentazione decimale dei numeri finora introdotti.

2.4. La rappresentazione decimale dei razionali

In questa sezione vogliamo mostrare come si possa arrivare a scrivere tutti i razionali usando un numero finito di simboli, detti cifre, opportunamente disposti in fila: in sostanza tratteremo se pur brevemente la scrittura *posizionale* dei numeri. Scopo finale è quello di arrivare alla scrittura decimale, ma, almeno in via provvisoria e per non creare confusione, cominciamo con l'utilizzare la scrittura romana degli interi (soprattutto perché questa scrittura non è posizionale), con un'aggiunta obbligatoria e cioè un simbolo per lo zero che i romani non avevano: i naturali saranno dunque $0, \mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}$, ecc. Solo per questioni estetiche scriveremo da subito 1 al posto di I.

2.4.1. Il caso degli interi

Il risultato fondamentale che consente di arrivare alla scrittura posizionale dei naturali (per gli interi basterà poi eventualmente solo aggiungere il segno "$-$") è il seguente.

Teorema 2.14. *Se b è un naturale >1, per ogni naturale n esiste una ed una sola successione di naturali compresi tra 0 e $b-1$, detti* cifre*, che indichiamo con $c_k, c_{k-1}, \dots, c_1, c_0$, tali che*

$$n = c_k b^k + c_{k-1} b^{k-1} + \cdots + c_1 b^1 + c_0 b^0 , \tag{2.10}$$

e si scrive

$$n = (c_k c_{k-1} \dots c_1 c_0)_b , \tag{2.11}$$

cioè per rappresentare il numero n si scrivono le cifre affiancate e con l'indicazione del numero b che viene chiamato base della scrittura posizionale.

Osserviamo esplicitamente che il numero c_k della formula (2.10) è diverso da zero, mentre gli altri possono anche essere nulli: anche se nulli vanno indicati nella scrittura (2.10) e, quindi nella (2.11). Il segreto della notazione posizionale sta essenzialmente nella possibile presenza di zeri tra le cifre. In sostanza la scrittura usata evidenzia subito, in base alla *posizione*, a quale potenza della base ciascuna cifra si riferisce.

Per evitare confusione è opportuno usare per le cifre (cioè per i numeri naturali compresi tra 0 e $b-1$) simboli semplici e tali da non creare ambiguità quando sono scritti semplicemente uno di fianco all'altro. Il caso $b=\mathrm{X}$, che porta alla scrittura *decimale*, è di gran lunga il più comune, anche se l'avvento dei calcolatori ha portato in auge anche i casi $b=\mathrm{II}$ (scrittura *binaria*), $b=\mathrm{VIII}$ (scrittura *ottale*) e $b=\mathrm{XVI}$ (scrittura *esadecimale*). Si noti che b deve essere maggiore di 1, cioè un simbolo non è sufficiente.

È chiaro che se $b \le \mathrm{X}$ si usano come cifre le cifre arabe; nel caso $b=\mathrm{XVI}$ si aggiungono a queste le lettere $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}$. Siccome la base X è la più comune, in questo caso si omette l'indicazione della base. Si osservi che la base b della scrittura si scrive sempre $(10)_b$: infatti si ha

$$b = 1b^1 + 0b^0 .$$

Esempio 2.3.

$$\begin{aligned}
\mathrm{CLXXVII} &= 1\times \mathrm{X}^{\mathrm{II}} + 7\times \mathrm{X}^1 + 7\times \mathrm{X}^0 = (177)_{\mathrm{X}} = 177\\
&= 2\times \mathrm{III}^{\mathrm{IV}} + 0\times \mathrm{III}^{\mathrm{III}} + 1\times \mathrm{III}^{\mathrm{II}} + 2\times \mathrm{III}^1 + 0\times \mathrm{III}^0 = (20120)_{\mathrm{III}}\\
&= \mathrm{B}\times \mathrm{XVI}^1 + 1\times \mathrm{XVI}^0 = (\mathrm{B}1)_{\mathrm{XVI}}
\end{aligned}$$

Si noti anche che, quanto più grande è b, tanto più "corta" è la scrittura posizionale di un numero.

2.4.2. Cenno al cambiamento di base

Ci limiteremo solo ad illustrare brevemente la tecnica da applicare per passare da una base qualunque alla base 10 e viceversa.

Il passaggio da una base qualunque alla base 10 si può fare in maniera elementare, ricordando il significato della scrittura posizionale e scrivendo anche la base in forma decimale. L'esempio seguente illustra la tecnica e può essere applicato in ogni caso.

$$(241)_{\mathrm{V}} = 2\times \mathrm{V}^2 + 4\times \mathrm{V}^1 + 1\times \mathrm{V}^0 = 2\times 5^2 + 4\times 5^1 + 1\times 5^0 = 71 .$$

Per fare il passaggio inverso basterà eseguire, nel sistema in base 10, delle successive divisioni con resto del numero dato per la base b voluta fino ad ottenere quoziente 0: i resti successivi, scritti dall'ultimo al primo, daranno la scrittura nella base b del numero dato in base 10. Si vedano i seguenti esempi.

1. $71 : 5 = 14$ con resto di 1;
2. $14 : 5 = 2$ con resto di 4;
3. $2 : 5 = 0$ con resto di 2:

si ha dunque $(71)_{\mathrm{X}} = (241)_{\mathrm{V}}$.

1. $2014 : 12 = 167$ con resto di 10, cioè A;
2. $167 : 12 = 13$ con resto 11, cioè B;
3. $13 : 12 = 1$ con resto di 1;
4. $1 : 12 = 0$ con resto di 1:

si ha dunque $(2014)_{\mathrm{X}} = (11\mathrm{BA})_{\mathrm{XII}}$.

2.4.3. Il caso dei razionali

Nel caso dei razionali ci limiteremo a considerare solo la scrittura decimale, con virgola. Consideriamo un numero razionale m/n, e supponiamo che la frazione usata per rappresentarlo sia ridotta ai minimi termini, cioè senza divisori comuni tra il numeratore e il denominatore; potremo anche limitarci a considerare solo i razionali positivi, per gli altri basterà aggiungere il segno "—".

Per capire meglio il senso di quanto diremo, richiamiamo il classico procedimento della divisione con virgola tra due naturali, considerando alcuni esempi.

Esempio 2.4. Nell'eseguire la divisione 19 : 4, relativa alla frazione $19/4$, il processo di divisione ha termine perché ad un certo punto si ottiene come resto 0:

$$\begin{array}{r|l} 19 & 4 \\ \cline{2-2} 30 & 4.75 \\ 20 & \\ 0 & \end{array}$$

In questo caso si scrive, come è ben noto,

$$\frac{19}{4} = 4.75$$

e il significato di questa scrittura è il seguente:

$$\frac{19}{4} = 4 \times 10^0 + 7 \times 10^{-1} + 5 \times 10^{-2},$$

con una chiara estensione della scrittura già adottata per i naturali nella formula (2.10). Si è in particolare introdotto il punto[6] per separare le cifre che si riferiscono alle potenze di 10 con esponente maggiore o uguale a zero da quelle che si riferiscono alle potenze di 10 con esponente negativo.

Un numero decimale ottenuto da una divisione in cui ad un certo punto di ottiene come resto zero si chiama numero *decimale finito*.

Esempio 2.5. Consideriamo la divisione relativa alla frazione $19/6$:

$$\begin{array}{r|l} 19 & 6 \\ \cline{2-2} 10 & 3.16\ldots \\ \underline{4}0 & \\ \underline{4} & \end{array}$$

In questo caso il processo ha termine perché un resto si ripete: continuando la divisione si otterrebbe esattamente, all'infinito, la stessa cifra dopo la virgola. Come è ben noto si esprime questo fatto dicendo che la rappresentazione decimale è *periodica* e la cifra, o il gruppo di cifre, che si ripete (in questo caso 6) è detto *periodo*, mentre la cifra, o il gruppo di cifre, che precede il periodo, subito dopo la virgola, si

[6] Come dichiarato anche nell'elenco delle notazioni, abbiamo deciso di usare il punto come separatore decimale, anziché la virgola come prescriverebbero le regole. Continueremo comunque a paralre di "divisione con la virgola", oppure di "cifre dopo la virgola", ecc.

chiama *antiperiodo*. Se non c'è antiperiodo il decimale si chiama *periodico semplice*, altrimenti *periodico misto*. Si usa la ben nota scrittura

$$\frac{19}{6} = 3.1\overline{6},$$

e il significato di questa scrittura è il seguente:

$$\begin{aligned}\frac{19}{6} &= 3\times 10^0 + 1\times 10^{-1} + 6\times 10^{-2} + 6\times 10^{-3} + \cdots \\ &= 3\times 10^0 + 1\times 10^{-1} + 6\times 10^{-2}\times\left(1 + 10^{-1} + 10^{-2} + \cdots\right) \\ &= 3\times 10^0 + 1\times 10^{-1} + 6\times 10^{-2}\times\sum_{i=0}^{+\infty} 10^{-i}.\end{aligned}$$

La somma di infiniti addendi che compare nell'espressione precedente si chiama *serie*: la trattazione di questo argomento esula dagli scopi di questo testo, ma non è difficile provare che, nonostante gli infiniti addendi, tutti positivi, il suo valore è $^{10}/_{9}$. Per rendersene conto a livello elementare si può verificare direttamente che

$$\frac{19}{6} = 3\times 10^0 + 1\times 10^{-1} + 6\times 10^{-2}\times\frac{10}{9}.$$

Una verifica ancora più diretta si ha dal seguente calcolo formale:

$$\sum_{i=0}^{+\infty} 10^{-i} = 1 + 10^{-1} + 10^{-2} + \cdots = 1 + 0.1 + 0.01 + 0.001 + \cdots = 1.111\ldots = 1.\overline{1} = \frac{10}{9}.$$

Esempio 2.6. Come ulteriore esempio consideriamo la divisione $40:33$ relativa alla frazione $^{40}/_{33}$:

$$\begin{array}{r|l} 40 & 33 \\ \cline{2-2} \underline{7}0 & 1.21\ldots \\ 40 & \\ \underline{7} & \end{array}$$

In questo caso la scrittura che si adotta è

$$\frac{40}{33} = 1.\overline{21},$$

con il seguente significato:

$$\begin{aligned}\frac{40}{33} &= 1\times 10^0 + 2\times 10^{-1} + 1\times 10^{-2} + 2\times 10^{-3} + 1\times 10^{-4} + \cdots \\ &= 1\times 10^0 + \left(2\times 10^{-1} + 1\times 10^{-2}\right) + \left(2\times 10^{-3} + 1\times 10^{-4}\right) + \cdots \\ &= 1\times 10^0 + \frac{21}{100}\left(1 + 10^{-2} + 10^{-4} + \cdots\right) = 1\times 10^0 + \frac{21}{100}\times\sum_{i=0}^{+\infty} 10^{-2i},\end{aligned}$$

dove compare di nuovo una somma simile alla precedente il cui valore è, questa volta, $^{100}/_{99}$ (se ne può fare una verifica con la stessa tecnica già indicata prima).

Esempio 2.7. È chiaro che, dovendo essere il resto minore del denominatore, ci potrà essere solo un numero finito di passi prima che uno dei resti si ripeta: precisamente se il denominatore è n si avrà una ripetizione al massimo dopo n passi. Consideriamo come esempio la divisione relativa alla frazione $^{15}/_{7}$:

$$
\begin{array}{r|l}
15 & 7 \\ \cline{2-2}
\underline{1}0 & 2.142857\ldots \\
30 & \\
20 & \\
60 & \\
40 & \\
50 & \\
\underline{1} &
\end{array}
$$

In questo caso il processo termina esattamente al 7° passo perché il resto "1" si ripete: con denominatore 7 è il massimo consentito, e poiché tutte le 6 cifre ottenute fanno parte del periodo si parla di *decimale periodico massimale*. Si ha la scrittura

$$\frac{15}{7} = 2.\overline{142857}\,.$$

Si noti che il processo di divisione termina quando si ripete un resto, *non* quando si ripete una delle cifre dopo la virgola, come mostrano gli esempi che seguono.

Esempio 2.8. Consideriamo la divisione 2249 : 1125, relativa alla frazione $^{2249}/_{1125}$:

$$
\begin{array}{r|l}
2249 & 1125 \\ \cline{2-2}
11240 & 1.9991\ldots \\
11150 & \\
10250 & \\
\underline{125}0 & \\
\underline{125} &
\end{array}
$$

da cui la scrittura

$$\frac{2249}{1125} = 1.999\overline{1}\,.$$

Esempio 2.9. Consideriamo la divisione 1 : 9900 relativa alla frazione $^{1}/_{9900}$:

$$
\begin{array}{l|l}
1 & 9900 \\ \cline{2-2}
10 & 0.0001\ldots \\
\underline{100} & \\
1000 & \\
10000 & \\
9900 & \\
\underline{100} &
\end{array}
$$

da cui la scrittura

$$\frac{1}{9900} = 0.00\overline{01}\,.$$

Una scrittura del tipo di quelle considerate negli esempi si chiama un *allineamento decimale*, rispettivamente finito o periodico. Possiamo convenire, per questioni di uniformità, di scrivere un allineamento decimale finito come un allineamento decimale periodico con periodo 0.

Riguardo agli allineamenti decimali vale il seguente importante teorema.

Teorema 2.15. *Nell'eseguire la divisione con virgola tra due naturali non si può mai ottenere un allineamento decimale con periodo* 9.

Tenendo conto di questo teorema e delle osservazioni fatte negli esempi trattati possiamo concludere con il seguente teorema.

Teorema 2.16. *Ogni numero razionale può essere rappresentato con un unico allineamento decimale periodico, eventualmente con periodo* 0, *ma mai con periodo* 9.

Vale anche il viceversa: ad ogni allineamento decimale con periodo diverso da 9 corrisponde un unico numero razionale, detto *frazione generatrice*, che si può trovare con un ragionamento facilmente deducibile dal seguente esempio.

Esempio 2.10. Consideriamo l'allineamento decimale $3.21\overline{973}$ e proponiamoci di ricavare la frazione generatrice, che chiameremo x :

$$x = 3.21\overline{973}\,.$$

Si ha, successivamente,

$$100x = 321.\overline{973}$$
$$1000 \times 100x = 321973.\overline{973}\,.$$

Sottraendo la prima dalla seconda si ottiene:

$$99\,900x = 321\,973 - 321\,,$$

in quanto tutte le cifre dopo la virgola si eliminano nella sottrazione, essendo esattamente identiche nelle due righe. Da qui si ricava subito

$$x = \frac{321\,973 - 321}{99\,900} = \frac{321\,652}{99\,900} = \frac{80\,413}{24\,975}\,.$$

Se ne deduce la (famosa!) regola: "per ottenere la frazione generatrice di un decimale periodico si scrive al numeratore la differenza fra il numero ottenuto allineando tutte le cifre fino al primo periodo compreso e quello ottenuto allineando tutte le cifre fino al primo periodo escluso, e al denominatore il numero ottenuto allineando tanti 9 quante sono le cifre del periodo, seguiti da tanti 0 quante sono le cifre dell'antiperiodo". Questa regola funziona anche per i decimali finiti (cioè con periodo 0), ma in questo caso conviene più semplicemente scrivere una frazione che ha al numeratore l'allineamento di tutte le cifre del numero (escluso il periodo 0) e al denominatore la potenza 10^n, dove n è il numero di cifre che seguono la virgola (senza il periodo 0).

$$4.75 = \frac{475}{100} = \frac{19}{4}\,.$$

Per capire, almeno a livello elementare, il perché del teorema 2.16, cioè l'impossibilità di ottenere allineamenti decimali con periodo 9, osserviamo che la scrittura $0,\overline{9}$, sulla base delle convenzioni adottate, significa

$$9\times 10^{-1}+9\times 10^{-2}+\cdots=\frac{9}{10}\times\sum_{i=0}^{\infty}10^{-i}=\frac{9}{10}\times\frac{10}{9}=1\,.$$

Del resto questo risultato si ottiene anche applicando la regola per la frazione generatrice, più sopra richiamata. Gli allineamenti decimali con periodo 9 si chiamano anche *allineamenti decimali impropri*. Essi in realtà possono essere pensati come una seconda scrittura degli allineamenti decimali finiti. Si ha, per esempio, $23.75\overline{9}=23.76$, come è facile provare.

Per concludere questi richiami sulla rappresentazione decimale dei razionali ricordiamo i seguenti fatti relativi al tipo di allineamento decimale che corrisponde ad un dato numero razionale m/n, dove si suppone che la frazione sia ridotta ai minimi termini e dove $n>0$.

1. Se il denominatore è 1 oppure se, scomposto in fattori, contiene solo i numeri 2 e 5, il numero può essere rappresentato con due allineamenti decimali di cui uno finito e uno improprio, cioè con periodo 9. Il numero razionale si dice anche numero razionale decimale (in considerazione del fatto che si può ridurre ad una frazione equivalente avente per denominatore una potenza di dieci, eventualmente $10^0=1$, se il numero stesso è un intero).
2. Se il denominatore, scomposto in fattori, contiene solo fattori diversi da 2 e da 5, il numero può essere rappresentato con un unico allineamento decimale periodico semplice, cioè senza antiperiodo.
3. Se il denominatore, scomposto in fattori, contiene sia uno dei fattori 2 o 5, che altri fattori, il numero può essere rappresentato con un unico allineamento decimale periodico misto (cioè con antiperiodo).

2.4.4. Operazioni con i decimali

La rappresentazione dei razionali mediante allineamenti decimali è molto utile in diverse circostanze, in particolare in tutte le applicazioni tecnico-pratiche. In particolare i calcolatori utilizzano normalmente rappresentazioni decimali, addirittura anzi solo rappresentazioni decimali finite, arrotondando o troncando opportunamente quelle illimitate. Tra l'altro gli allineamenti decimali consentono di trattare molto facilmente tutti i problemi relativi all'ordine. Le difficoltà nascono quando si considerano le operazioni di somma e prodotto. Per gli allineamenti finiti si possono applicare le tecniche note fin dalla scuola primaria, ma esse sono di difficile realizzazione pratica nel caso in cui il numero delle cifre decimali è molto grande, e in ogni caso esse non sono applicabili per gli allineamenti periodici se non in alcuni semplici casi. Per capire il problema consideriamo alcuni esempi.

Esempio 2.11. Essendo $1/3=0.\overline{3}$ si ha, facilmente,

$$3\times 0.\overline{3}=0,\overline{9}=1\,,$$

come è giusto. In questo caso la moltiplicazione è stata facile, perché non esiste il problema dei riporti. Ma come potremmo calcolare $4\times 0.\overline{3}$? La tecnica della moltiplicazione richiede di iniziare "sull'estrema destra", ma in questo caso ciò non è possibile: "l'ultima cifra" del prodotto dovrebbe essere un 2 e poi tutte le altre dovrebbero diventare 3, per il riporto, ma qual è l'ultima cifra? Si può scommettere, in un caso così semplice, che il risultato sarà $1.\overline{3}$, come è giusto.

Esempio 2.12. Supponiamo ora di voler calcolare $0.\overline{3} \times 0.\overline{3}$. Passando attraverso le frazioni il calcolo è molto semplice:

$$0.\overline{3} \times 0.\overline{3} = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9} = 0.\overline{1}\,.$$

Lavorare direttamente con gli allineamenti decimali non è altrettanto facile (anche se, forse, con un po' di sforzo si potrebbe arrivare al risultato corretto, tenendo conto che si sa a priori che il risultato deve essere un decimale periodico).

Esempio 2.13. A proposito delle difficoltà di calcolo con gli allineamenti decimali è interessante un problema proposto nell'aprile 2001 sul sito dell'Institut de Recherche sur l'Enseignement des Mathématiques dell'Unversité Louis Pasteur di Strasbourg. La domanda era:

> *Qual è il periodo del numero decimale* $(1.\overline{001})^2$?

Il problema non è per niente semplice, nemmeno con un buon software di calcolo su computer e men che mai con una calcolatrice tascabile. Poiché

$$1.\overline{001} = \frac{1000}{999} = \frac{1000}{3^3 \times 37}\,,$$

se ne deduce che il decimale richiesto sarà periodico semplice (il denominatore non contiene né il fattore 2 né il fattore 5) e con non più di $998\,000 = 999^2 - 1$ cifre! Un software come *Mathematica* può fornire un valore approssimato della frazione data con un arbitrario numero di cifre decimali, ma con un periodo probabile così lungo non è facile fare un controllo. Si può, più proficuamente, utilizzare un software che ricerchi il periodo eseguendo la divisione e cercando la prima ripetizione: si trova il risultato seguente, con un periodo di 2997 cifre:

$$(1.\overline{001})^2 = 1.\overline{002\,003\ldots995\,996\,997\,999\,000\,001}.$$

Il risultato è abbastanza curioso: nello sviluppo compaiono tutti i numeri di tre cifre, da 002 a 997, più 999, 000, 001, cioè tutti i numeri di tre cifre escluso il 998, e questa stranezza giustifica il perché era stato proposto questo esercizio. Il risultato potrebbe essere ottenuto anche manualmente, ma con un bel po' di fantasia e organizzazione, come dimostrato nella soluzione proposta sul citato sito dell'università di Strasburgo.

Normalmente, come già citato, i computer fanno i calcoli con un allineamento decimale finito, ottenuto troncando o approssimando lo sviluppo decimale esatto, ma questo può comportare gravi problemi, come prova l'esempio, elementare, che segue.

Esempio 2.14. Poiché

$$\frac{1}{3} \times 4 - 1 = \frac{1}{3}\,,$$

ci si deve aspettare che moltiplicando nuovamente il risultato per 4 e sottraendo 1 si ottenga di nuovo, e all'infinito, sempre $1/3$. La cosa è certamente vera se operiamo con le frazioni. Supponiamo invece di approssimare la frazione $1/3$ con, per esempio, 0.3333 (quattro cifre decimali esatte: l'errore è più piccolo di un decimillesimo), e di ripetere il calcolo indicato dieci volte:

$$(((((((((0.3333 \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1) \times 4 - 1)\,.$$

Otteniamo come risultato -34.6192, cioè un valore completamente inaccettabile! Usando una migliore approssimazione avremmo ottenuto un risultato migliore, ma aumentando il numero di ripetizioni avremmo comunque ottenuto un valore sballato.

Si deve sempre tenere conto di questi fatti quando si opera con gli allineamenti decimali.

2.5. I numeri reali

2.5.1. Definizioni e proprietà

Come già osservato, l'insieme dei numeri razionali non è sufficiente per trattare problemi anche semplici, come quello della risoluzione dei triangoli rettangoli. Purtroppo l'introduzione rigorosa dell'insieme dei numeri reali, che amplia quello dei razionali e risolve i problemi accennati, non è semplice e ci limiteremo solo ad un introduzione intuitiva che renda l'idea di come si possa procedere e di quali siano le difficoltà: il resto sarà compito dei corsi universitari di matematica.

Fissiamo l'attenzione sul problema di risolvere l'equazione $x^2 = 2$, e andiamo alla ricerca delle migliori soluzioni, positive, approssimate per difetto, rispettivamente con una, due, tre, ecc. cifre decimali esatte. Si può facilmente costruire la tabella

	0	1	2	3	4	5	6	...
x	1	1.4	1.41	1.414	1.1442	1.41421	1.414213	...
x^2	1	1.96	1.9881	1.999396	1.99996164	1.9999899241	1.999998409369	...

Così facendo si viene a determinare una successione di allineamenti decimali: 1; 1.4; 1.41; 1.414; 1.4142; ..., costituita da numeri il cui quadrato è sempre minore di 2, ma costituenti la migliore approssimazione per difetto di una soluzione "esatta" dell'equazione $x^2 = 2$, a quel livello di cifre decimali esatte. Come sappiamo, non esiste alcuna soluzione di quest'equazione nell'insieme dei razionali, e quindi non c'è possibilità che l'allineamento decimale sia finito o che ad un certo punto diventi periodico: procedendo all'infinito potremo costruire un allineamento decimale illimitato e non periodico e assumere che proprio questo allineamento illimitato e non periodico sia la soluzione "esatta" dell'equazione in questione. Questo ci suggerisce di considerare l'insieme di *tutti gli allineamenti decimali*, periodici e non, e di vedere se è possibile introdurre in questo insieme un ordine e le operazioni di addizione e di moltiplicazione in modo da estendere le proprietà delle stesse operazioni valide in $\mathbb{Q}$. Anche se il processo non è semplice, come ben si capisce se si tiene conto delle difficoltà già accennate nell'eseguire le operazioni con gli allineamenti periodici, questa idea ha successo e si conclude con l'introduzione di un nuovo insieme numerico detto insieme dei *numeri reali*, e indicato con $\mathbb{R}$, di cui i razionali diventano un sottoinsieme: $\mathbb{Q} \subset \mathbb{R}$. Segnaliamo che non è questo l'unico modo per introdurre l'insieme dei reali (anzi forse non è nemmeno il più comune), ma qui ci è utile per queste brevi considerazioni introduttive.

Come già accennato, in questo insieme $\mathbb{R}$ si introduce un ordine (in maniera molto semplice: si tratta del cosiddetto *ordine lessicografico*, basato sulle stesse regole che sovraintendono all'ordine delle parole di una lingua) e due operazioni di addizione e moltiplicazione: l'insieme $\mathbb{R}$ con questa struttura diventa un campo ordinato e, dal punto di vista delle proprietà delle operazioni, nulla si modifica rispetto a quanto già detto in $\mathbb{Q}$. La novità veramente importante concerne invece l'ordine e precisamente vale la *Proprietà dell'estremo superiore*: ogni insieme non vuoto e superiormente limitato di numeri reali ha estremo superiore. Non possiamo ulteriormente insistere su questa cruciale proprietà (sarà compito dei corsi universitari farlo): rimarchiamo solo che è in conseguenza di essa, sostanzialmente, che l'equazione $x^2 = 2$ ha soluzioni in $\mathbb{R}$, mentre non ne aveva in $\mathbb{Q}$. Anzi, in $\mathbb{R}$ si ha che non solo la citata equazione di secondo grado ha soluzione, ma che ha soluzione anche ogni equazione del tipo

$$x^n = a, \quad n \in \mathbb{N}, n \geq 2, \quad a \in \mathbb{R}, a \geq 0. \tag{2.12}$$

L'insieme $\mathbb{R}$ ha anche un'altra importante proprietà, e precisamente quella di essere un corpo *archimedeo*; questo significa precisamente che, dati due reali $a > 0$ e $b > a$, esiste un multiplo di a che supera b: si tratta di una proprietà apparentemente intuitiva, ma che non tutti i corpi ordinati hanno. Anche su questo, comunque, non insistiamo oltre.

Per quanto riguarda la "quantità di elementi" appartenenti ad $\mathbb{R}$, ricordiamo che abbiamo già provato, nella pagina 28, che l'insieme dei reali è più numeroso dell'insieme dei naturali e quindi anche dell'insieme dei razionali, in quanto abbiamo anche provato, vedi il teorema 2.13 nella pagina 49, che i razionali hanno la stessa cardinalità dei naturali: l'insieme dei reali è dunque decisamente "più ricco" di quello dei razionali.

Avendo introdotto l'insieme dei reali come insieme di tutti gli allineamenti decimali (possiamo escludere quelli con periodo 9, per evitare duplicati), possiamo pensarlo come l'unione tra i decimali finiti o periodici (che corrispondono ai razionali) e i decimali illimitati non periodici: questi ultimi si chiamano *reali irrazionali* o semplicemente *irrazionali*. Se li indichiamo con $\mathbb{I}$, possiamo dire che $\mathbb{R} = \mathbb{Q} + \mathbb{I}$, con $\mathbb{Q} \cap \mathbb{I} = \emptyset$. L'insieme degli irrazionali è strutturalmente poco importante perché, come vedremo subito, la somma e il prodotto di due irrazionali può non essere irrazionale. Quanto abbiamo sopra osservato sulla cardinalità ha come conseguenza che gli irrazionali hanno cardinalità più grande di $\aleph_0$: la quasi totalità dei reali è irrazionale!

Anche con le poche cose finora dette possiamo provare le seguenti proprietà dei reali.

- La somma tra due irrazionali può benissimo essere razionale. Per un esempio basta considerare i numeri $a = 2.\alpha_1\alpha_2\dots$ e $b = -1.\alpha_1\alpha_2\dots$ (dove abbiamo supposto che gli allineamenti decimali siano illimitati non periodici e che le cifre dopo la virgola siano identiche). È evidente che, mentre $a, b \in \mathbb{I}$, $a + b = 1 \in \mathbb{Q}$.
- Il prodotto tra due irrazionali può essere razionale. Per un esempio basta considerare l'allineamento decimale illimitato non periodico originato dal risolvere l'equazione $x^2 = 2$: moltiplicandolo per se stesso si ottiene naturalmente 2, che è un razionale.
- La somma tra un razionale e un irrazionale è sicuramente irrazionale. Se infatti prendendo $r \in \mathbb{Q}$ e $i \in \mathbb{I}$ si avesse $r + i = s \in \mathbb{Q}$, si avrebbe $i = s - r \in \mathbb{Q}$ (perché $\mathbb{Q}$ è chiuso rispetto all'addizione) ma questo è contro l'ipotesi.
- Il prodotto tra un razionale e un irrazionale è sicuramente irrazionale. Se infatti prendendo $r \in \mathbb{Q}$ e $i \in \mathbb{I}$ si avesse $r \cdot i = s \in \mathbb{Q}$, si avrebbe $i = {}^s/_r \in \mathbb{Q}$ (perché $\mathbb{Q}$ è chiuso rispetto alla moltiplicazione) ma questo è contro l'ipotesi.

2.5.2. Ascisse sulla retta

Uno dei risultati più importanti relativi all'introduzione dei reali è il fatto di poter stabilire una corrispondenza biunivoca tra i reali stessi e i punti di una retta, stabilendo quello che si chiama un *sistema di ascisse* sulla retta stessa, che a questo punto viene chiamata *asse delle ascisse*. Sostanzialmente si procede nel seguente modo: su una retta r si fissa un punto O, detto *origine* e un punto U. Se si assume la lunghezza del segmento $\overline{OU}$ come unità di misura, ad esso si può far corrispondere il numero reale 1; ad ogni altro punto P della semiretta OU si può far corrispondere il numero reale positivo che rappresenta la lunghezza del segmento $\overline{OP}$; ad ogni punto P dell'altra semiretta di origine O si può far corrispondere l'opposto del numero che rappresenta la lunghezza del segmento $\overline{OP}$. In questo modo si viene ad associare ad ogni numero reale uno e un sol punto della retta su cui si sia fissato il sistema di ascisse, e viceversa, potendo addirittura identificare i numeri reali con i punti della retta stessa: in molti contesti si parla addirittura (e anche noi lo faremo spesso) di "punti" come sinonimo di "numeri reali". Si parla spesso di *retta reale* per intendere l'insieme dei numeri reali, che è comunemente immaginato proprio come una retta orientata.

È interessante osservare che l'introduzione dei numeri reali consente di risolvere quello che, a livello intuitivo, può sembrare un paradosso. Consideriamo infatti la costruzione della figura 2.2, dove OUAB è un quadrato.

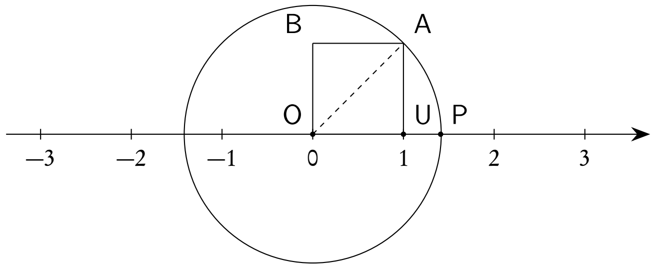


Figura 2.2.: *Intersezione tra retta e circonferenza*

Senza i numeri reali dovremmo concludere che non esiste intersezione tra l'asse delle ascisse e la circonferenza di centro O e passante per A: infatti il punto P (cioè uno dei punti "candidati" ad essere di intersezione tra retta e circonferenza) deve avere come ascissa un numero il cui quadrato è 2, e ben sappiamo che in $\mathbb{Q}$ tale numero non esiste.

2.5.3. Intervalli di numeri reali

Nella retta, segmenti e semirette sono sottoinsiemi importanti. Acquistano quindi importanza i sottoinsiemi di $\mathbb{R}$ che loro corrispondono nella rappresentazione dei reali sulla retta. Per semplificare le scritture (e per altri motivi che appariranno chiari trattando il concetto di limite), si usa aggiungere all'insieme dei numeri reali due altri oggetti: $-\infty$ (meno infinito) e $+\infty$ (più infinito). L'insieme così ottenuto si chiama *retta reale estesa* e si pone

$$\widetilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}. \tag{2.13}$$

Si può prolungare l'ordinamento di $\mathbb{R}$ ponendo

$$-\infty < x < +\infty,\ \forall x \in \mathbb{R} \quad \text{e} \quad -\infty < +\infty. \tag{2.14}$$

Non si estendono invece ad $\widetilde{\mathbb{R}}$ le operazioni di addizione e moltiplicazione: $-\infty$ e $+\infty$ *non* sono numeri e non possono essere trattati come tali. Nei corsi di analisi si vedrà che non è possibile estendere le operazioni di addizione e moltiplicazione a $\widetilde{\mathbb{R}}$ mantenendo le proprietà formali.

Definizione 2.17 (Intervalli)**.** *Si chiamano* intervalli *di* $\mathbb{R}$ *tutti i sottoinsiemi* I *di* $\mathbb{R}$ *che soddisfano alla seguente condizione: se* a, $b \in I$, *e* $a \leq b$ *allora ogni* $x \in \mathbb{R}$ *tale che* $a \leq x \leq b$ *appartiene ad* I. *A parole un sottoinsieme* I *di* $\mathbb{R}$ *è un intervallo se, non appena contiene due numeri reali, contiene anche tutti i numeri reali compresi tra questi due.*

Si può provare[7] che gli intervalli di $\mathbb{R}$ sono tutti e soli quelli di seguito elencati, dove a, $b \in I$, e $a \leq b$.

1. Intervalli *degeneri*.
   - $\emptyset$;
   - $\{a\}$, $\forall a \in \mathbb{R}$.
2. Intervalli *limitati*, corrispondenti ai segmenti, con o senza uno o entrambi gli estremi.
   - $[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$: intervallo *chiuso*, di estremi a e b;
   - $[a,b[= \{x \in \mathbb{R} \mid a \leq x < b\}$: intervallo *aperto a destra*, di estremi a e b;
   - $]a,b] = \{x \in \mathbb{R} \mid a < x \leq b\}$: intervallo *aperto a sinistra*, di estremi a e b;
   - $]a,b[= \{x \in \mathbb{R} \mid a < x < b\}$: intervallo *aperto*, di estremi a e b.
3. Intervalli *illimitati*, corrispondenti alle semirette o all'intera retta.
   - $]-\infty,a] = \{x \in \mathbb{R} \mid x \leq a\}$: intervallo chiuso inferiormente illimitato;
   - $]-\infty,a[= \{x \in \mathbb{R} \mid x < a\}$: intervallo aperto inferiormente illimitato;
   - $[a,+\infty[= \{x \in \mathbb{R} \mid x \geq a\}$: intervallo chiuso superiormente illimitato;
   - $]a,+\infty[= \{x \in \mathbb{R} \mid x > a\}$: intervallo aperto superiormente illimitato;
   - $]-\infty,+\infty[= \mathbb{R}$: intervallo illimitato.

2.5.4. Valore assoluto

Definizione 2.18 (Valore assoluto)**.** *Dato un numero reale* x, *si chiama suo* valore assoluto *o* modulo *il numero reale, indicato con* $\operatorname{abs} x$ *o con* $|x|$, *definito come segue*

$$\operatorname{abs} x = |x| = \begin{cases} x, & \text{se } x \geq 0; \\ -x, & \text{se } x < 0. \end{cases} \tag{2.15}$$

Se si fa riferimento alla rappresentazione dei reali sull'asse delle ascisse, $|x|$ rappresenta la distanza del punto x dall'origine. Valgono le seguenti proprietà del valore assoluto.

1. $|x+y| \leq |x| + |y|$ (detta anche *disuguaglianza triangolare*);
2. $|x-y| \geq \big||x| - |y|\big|$;
3. $|x \cdot y| = |x| \cdot |y|$;
4. $\left|\dfrac{x}{y}\right| = \dfrac{|x|}{|y|}$, con $y \neq 0$.

Esempio 2.15. $|-5| = +5$; $|+8| = +8$; $|x^2| = x^2$; $|x|^2 = x^2$.

[7] La cosa è intuitivamente evidente, ma è più profonda di quanto si pensi. Una dimostrazione si può vedere in [4].

2.5.5. I radicali

Abbiamo già detto che, in $\mathbb{R}$, ogni equazione del tipo

$$x^n = a, \quad n \in \mathbb{N}, n \geq 2, \quad a \in \mathbb{R}, a \geq 0 \tag{2.16}$$

ha soluzioni. Precisamente essa ha

— 2 soluzioni opposte, nel caso di n pari;
— 1 una sola soluzione positiva, nel caso di n dispari.

Si dà la seguente definizione.

Definizione 2.19 (Radice aritmetica)**.** *L'unica soluzione positiva dell'equazione* (2.16) *si chiama* radice n-esima aritmetica *o* radicale n-esimo aritmetico *del numero reale positivo a e si indica con*

$$\sqrt[n]{a}\,. \tag{2.17}$$

Il numero n si dice indice del radicale*, il numero a* radicando.

Si noti che la definizione appena data significa che se a è un reale ≥ 0 e n è un naturale ≥ 2, si ha:

$$\left(\sqrt[n]{a}\right)^n = a\,. \tag{2.18}$$

Se $n = 2$ si scrive semplicemente $\sqrt{a}$, senza l'indicazione dell'indice, e la radice si dice *radice quadrata*. Se $n = 3$ la radice si chiama *radice cubica*. Se nell'equazione (2.16) fosse $n = 1$ non ci sarebbe alcun problema nel risolverla, per qualunque valore del reale. Per questo motivo i radicali devono avere indice almeno 2.

Osservazione 2.20. Si noti che, parlando di radicale, non è escluso che il radicando sia un numero reale: nell'equazione (2.16) a può essere un qualunque numero positivo. Tuttavia, nelle applicazioni, hanno interesse solo i radicali con radicando razionale se non addirittura intero: per questo di solito si usa il nome di radicale solo in presenza di radicando razionale, e noi ci atterremo a questa consuetudine.

Anche l'equazione

$$x^n = a, \quad n \in \mathbb{N}, n \text{ dispari } \geq 3, \quad a \in \mathbb{R}, a < 0 \tag{2.19}$$

ha un'unica soluzione, negativa. Purtroppo per quest'unica soluzione negativa si adotta la stessa simbologia già adottata per l'unica soluzione positiva dell'equazione (2.16), cioè si usa ancora la scrittura

$$\sqrt[n]{a}\,,$$

solo che si parla di *radicale algebrico* anziché di radicale aritmetico. Il fatto che si usi lo stesso simbolo può essere fonte di equivoci e portare a grossolani errori. Bisogna prestare la massima attenzione.

Osservazione 2.21. Trattando l'equazione $x^2 = 2$, che è un caso particolare dell'equazione (2.16), abbiamo detto che la sua soluzione è costituita da un allineamento decimale illimitato e non periodico che inizia con $1.41421\ldots$. Perché allora abbiamo voluto introdurre un nuovo simbolo, cioè $\sqrt{2}$, per questa stessa soluzione? La scelta è dettata da vari fattori.

— L'introduzione dei radicali ci consente di avere un simbolo "universale", adatto a tutte le equazioni del tipo (2.16), per qualunque radicando e qualunque indice.
— L'uso dei radicali ci consente di capire immediatamente di quale equazione un certo radicale è soluzione.
— L'uso dei radicali consente di evitare, almeno entro certi limiti, i calcoli con gli allineamenti decimali. Il problema dei calcoli con gli allineamenti decimali, di cui abbiamo già messo in luce i limiti nel caso degli allineamenti finiti o periodici (vedi la sezione 2.4.4, nella pagina 57), diventa ancora più complesso nel caso degli allineamenti illimitati non periodici. Per capirlo basta considerare il semplice esempio che segue. Mentre

$$\left(\sqrt{2}\right)^{10}=\left(\left(\sqrt{2}\right)^2\right)^5=2^5=32,$$

se approssimiamo $\sqrt{2}$ con 1.4 (compiendo un errore di poco più di un centesimo), otteniamo invece

$$(1.4)^{10}=28.9254654976,$$

con un errore enorme. Torneremo ancora su questo problema con l'osservazione 2.22, nella pagina 69.

Per i *soli radicali aritmetici* valgono le proprietà seguenti, che in sostanza spiegano come si opera con i radicali. Alcune modifiche saranno necessarie per poter operare in maniera simile con radicali algebrici. È molto importante ricordare che le proprietà elencate valgono solo per i radicali aritmetici: lo evidenzieremo segnalando sempre che i radicandi dovranno essere reali ≥ 0.

Proprietà invariantiva dei radicali

Se $a \geq 0$ è un reale e n, m, k sono naturali maggiori o uguali a 2, si ha:

$$\sqrt[n]{a^m}=\sqrt[kn]{a^{km}}. \tag{2.20}$$

Questa proprietà consente di *ridurre più radicali allo stesso indice*. Per esempio i radicali

$$\sqrt[3]{5},\quad \sqrt[4]{7},\quad \sqrt[6]{11}$$

possono essere trasformati in radicali con indice il minimo comune multiplo degli indici:

$$\sqrt[12]{5^4},\quad \sqrt[12]{7^3},\quad \sqrt[12]{11^2}.$$

Letta da destra a sinistra questa stessa proprietà consente di *semplificare un radicale* per un divisore comune del suo indice e dell'esponente del radicando. Si ha, per esempio,

$$\sqrt[15]{7^{12}}=\sqrt[5]{7^4},\quad \sqrt[8]{64}=\sqrt[8]{2^6}=\sqrt[4]{2^3}.$$

Si noti, in particolare, che

$$\sqrt[n]{a^{mn}}=a^m. \tag{2.21}$$

Prodotto e quoziente di radicali con lo stesso indice

Se $a \geq 0$ e $b \geq 0$ sono reali, e n è un naturale ≥ 2, si ha

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{a \cdot b}. \tag{2.22}$$

Se $a \geq 0$ e $b > 0$ sono reali, e n è un naturale ≥ 2, si ha

$$\sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}. \tag{2.23}$$

Se i radicali non hanno lo stesso indice, il prodotto o il quoziente si possono eseguire dopo averli ridotti allo stesso indice. Si ha, per esempio,

$$\sqrt[3]{5} \cdot \sqrt[3]{7} = \sqrt[3]{35}, \quad \sqrt[3]{5} \cdot \sqrt[4]{7} \cdot \sqrt[6]{11} = \sqrt[12]{5^4} \cdot \sqrt[12]{7^3} \cdot \sqrt[12]{11^2} = \sqrt[12]{5^4 \times 7^3 \times 11^2}.$$

Potenza e radice di un radicale

Se a è un reale ≥ 0, n è un naturale ≥ 2, e m è un intero, si ha

$$\left(\sqrt[n]{a}\right)^m = \sqrt[n]{a^m}. \tag{2.24}$$

Se a è un reale ≥ 0, n ed m sono naturali ≥ 2, si ha

$$\sqrt[m]{\sqrt[n]{a}} = \sqrt[nm]{a}. \tag{2.25}$$

Portare dentro e fuori dal segno di radice

Utilizzando le proprietà precedenti si possono eseguire anche le seguenti due operazioni, di uso molto frequente, che consistono nel portare un fattore *positivo* dentro o fuori dal segno di radice.

$$\text{Se } a \geq 0 \quad \text{allora} \quad a\sqrt[n]{b} = \sqrt[n]{a^n \cdot b}, \quad \text{e inoltre} \quad \sqrt[n]{a^{mn} \cdot b} = a^m \sqrt[n]{b}. \tag{2.26}$$

Esempio 2.16. $3\sqrt[4]{5} = \sqrt[4]{3^4 \times 5} = \sqrt[4]{405}$.

Esempio 2.17. $-2\sqrt[3]{4} = -\sqrt[3]{2^3 \times 4} = -\sqrt[3]{32}$.

Esempio 2.18. $a\sqrt[n]{b} = \begin{cases} \sqrt[n]{a^n \cdot b}, & \text{se } a \geq 0; \\ -\sqrt[n]{a^n \cdot b}, & \text{se } a < 0. \end{cases}$

Esempio 2.19. $\sqrt[3]{20000} = \sqrt[3]{2^5 \times 5^4} = 2 \times 5 \times \sqrt[3]{2^2 \times 5} = 10\sqrt[3]{20}$.

Il problema della somma di radicali

Poiché, come abbiamo visto,

$$a\sqrt[n]{b} = \sqrt[n]{a^n \cdot b}, \quad \text{se } a \geq 0,$$

si estende la denominazione di *radicale* anche alle espressioni del tipo $a\sqrt[n]{b}$. A questo punto due radicali come $a\sqrt[n]{b}$ e $c\sqrt[n]{b}$ si dicono *simili*. È abbastanza evidente che la somma di due radicali simili è ancora un radicale, simile ai due dati:

$$a\sqrt[n]{b} + c\sqrt[n]{b} = (a+c)\sqrt[n]{b}.$$

Per convincersi basta applicare la proprietà distributiva della moltiplicazione rispetto all'addizione. Possiamo chiederci che cosa si può affermare della somma di due radicali non simili (e che non possano nemmeno essere resi simili mediante applicazioni delle proprietà). Per capire il problema consideriamo

$$\sqrt{2}+\sqrt{3}.$$

Cominciamo col provare che questo numero non è un razionale. Se infatti esistesse un $r \in \mathbb{Q}$ tale che

$$\sqrt{2}+\sqrt{3}=r,$$

si avrebbe, elevando al quadrato ambo i membri (positivi)

$$2+3+2\sqrt{6}=r^2 \Rightarrow \sqrt{6}=\frac{r^2-5}{2}$$

ovvero che $\sqrt{6}$ è razionale, in quanto ottenuto da razionali con le operazioni elementari. Con la stessa tecnica con cui abbiamo provato (vedi la pagina 50) che $\sqrt{2}$ è irrazionale, si può provare che anche $\sqrt{6}$ lo è, per cui si cade in un assurdo. In maniera simile si può provare che

$$\nexists r \in \mathbb{Q} \quad \text{tale che} \quad \sqrt{2}+\sqrt{3}=\sqrt{r}.$$

Più difficile è provare che, $\forall\, n \in \mathbb{N}$, $n>1$,

$$\nexists r \in \mathbb{Q} \quad \text{tale che} \quad \sqrt{2}+\sqrt{3}=\sqrt[n]{r}.$$

Dunque la somma proposta non è un radicale con radicando razionale[(8)].

Questa caratteristica della somma di radicali è uno dei fattori che rendono difficile il calcolo con i radicali: mentre la somma e il prodotto di due razionali è ancora un razionale, la somma e il prodotto di due radicali non è detto che sia un radicale, ovvero l'insieme dei radicali non ha nessuna struttura significativa rispetto alle operazioni di addizione e moltiplicazione

Avendo introdotto il concetto di radicale, le proprietà relative alla somma e prodotto di razionali e irrazionali, già considerate nella pagina 60, si possono verificare ancora più facilmente.

$$1-\sqrt{2}\in\mathbb{I}, \quad \sqrt{2}\in\mathbb{I}, \quad \text{ma} \quad \left(1-\sqrt{2}\right)+\sqrt{2}=1\in\mathbb{Q}.$$
$$\sqrt[3]{4}\in\mathbb{I}, \quad \sqrt[3]{16}\in\mathbb{I}, \quad \text{ma} \quad \sqrt[3]{4}\cdot\sqrt[3]{16}=\sqrt[3]{64}=4\in\mathbb{Q}.$$

Estensione delle proprietà ai radicali algebrici

Le proprietà dei radicali aritmetici possono essere estese, in alcuni casi, anche ai radicali algebrici. A nostro avviso non è opportuno fissare regole: conviene piuttosto esaminare alcuni esempi significativi che mettano in luce le possibili strategie.

$$\sqrt{(-3)^2}=\sqrt{3^2}=3=|-3|;$$

[8] È invece ovvio che

$$\sqrt{2}+\sqrt{3}=\sqrt{5+2\sqrt{6}}$$

cioè che la somma proposta è un radicale con radicando irrazionale, ma questo fatto è di poco interesse. Si veda in proposito l'osservazione 2.20 nella pagina 63.

$$\sqrt{(-3)(-4)} = \sqrt{3\times 4} = \sqrt{3}\sqrt{4} = \sqrt{|-3|}\sqrt{|-4|};$$
$$\sqrt{(-3)^2\times 5} = \sqrt{3^2\times 5} = 3\sqrt{5} = |-3|\sqrt{5};$$
$$\sqrt[6]{(-3)^4} = \sqrt[6]{3^4} = \sqrt[3]{3^2} = \sqrt[3]{(-3)^2};$$
$$\sqrt[6]{(-3)^2} = \sqrt[6]{3^2} = \sqrt[3]{3} = \sqrt[3]{|-3|}.$$

Una situazione che merita particolare attenzione è la seguente:

$$\sqrt{ab} = \sqrt{|a|}\sqrt{|b|}, \tag{2.27}$$

uguaglianza che è valida da sinistra a destra, ma non necessariamente da destra a sinistra. Si noti infatti che il primo membro ha senso quando a e b sono concordi, mentre il secondo ha senso qualunque siano a e b. Per esempio

$$\sqrt{(-2)(-3)} = \sqrt{|-2|}\sqrt{|-3|}, \quad \text{mentre} \quad \sqrt{|-2|}\sqrt{|3|} \neq \sqrt{(-2)\times 3}.$$

Un'altra situazione che si presenta spesso negli esercizi, e che bisogna ben tenere presente, è la seguente:

$$\sqrt{a^2} = |a|, \quad \forall a\in\mathbb{R}. \tag{2.28}$$

A proposito della formula (2.28) è opportuno segnalare che, in base alla definizione di radice (in questo caso quadrata), si ha

$$\left(\sqrt{a}\right)^2 = a, \quad \text{se} \quad a\geq 0. \tag{2.29}$$

Si noti dunque la differenza tra la (2.28), che vale per ogni reale a, e la (2.29), che invece vale solo per i reali ≥ 0: dunque

$$\sqrt{a^2} \neq \left(\sqrt{a}\right)^2, \tag{2.30}$$

mentre un'errata applicazione delle proprietà dei radicali, e in particolare della formula (2.24), porterebbe a concludere che i due membri della (2.30) sono uguali. Poiché è molto frequente un'errata estensione delle proprietà dei radicali aritmetici ai radicali algebrici, segnaliamo di nuovo che la formula (2.24), come le altre che si riferiscono alle proprietà dei radicali, esprime una proprietà dei radicali aritmetici e dunque è valida solo per $a\geq 0$: si ha infatti, correttamente,

$$\sqrt{a^2} = \left(\sqrt{a}\right)^2 \quad \text{se, e solo se,} \quad a\geq 0. \tag{2.31}$$

In generale, in presenza di radicali in cui la base del radicando sia negativa (oppure sia un parametro che può assumere sia valori positivi che negativi), bisogna controllare che l'applicazione delle proprietà dei radicali non alteri i segni, né modifichi le eventuali condizioni per l'esistenza.

Per concludere queste considerazioni sui radicali proponiamo due esempi molto interessanti che mettono in luce tutte le difficoltà insite nei calcoli relativi.

Esempio 2.20. Si ha $\sqrt{x^2(x-1)} \neq |x|\sqrt{x-1}$, in quanto il primo membro ha senso anche per $x=0$, mentre il secondo no.

Esempio 2.21. Si ha $\sqrt{x^2(x+1)} = |x|\sqrt{x+1}$, in quanto i due membri hanno senso per gli stessi valori di x, ovvero per $x\geq -1$.

Radicali doppi

Un'espressione del tipo

$$\sqrt{a \pm \sqrt{b}}$$

si chiama un *radicale doppio*, o *radicale quadratico doppio*. Vale la seguente identità:

$$\sqrt{a \pm \sqrt{b}} = \sqrt{\frac{a + \sqrt{a^2 - b}}{2}} \pm \sqrt{\frac{a - \sqrt{a^2 - b}}{2}}, \tag{2.32}$$

che si può provare elevando al quadrato ambo i membri. Se

$$\sqrt{a^2 - b}$$

è un numero razionale (cioè se $a^2 - b$ è un "quadrato perfetto"), allora la (2.32) consente di trasformare il radicale doppio nella somma di due radicali semplici. Per esempio si ha

$$\sqrt{4 - 2\sqrt{3}} = \sqrt{4 - \sqrt{12}} = \sqrt{3} - 1.$$

Razionalizzazioni

Spesso nelle applicazioni si presentano delle frazioni contenenti al numeratore e/o al denominatore uno o più radicali e, a volte, è più conveniente modificare la frazione in modo che il numeratore e/o il denominatore non contenga radicali: si dice che si è operata una *razionalizzazione*. A livello di calcolo algebrico di solito si è interessati a razionalizzare i denominatori (e di questo ci occuperemo negli esempi proposti); nel calcolo di limiti o integrali, per esempio, si è però anche spesso interessati anche alla razionalizzazione di numeratori. Il problema, in generale, è molto difficile se non irresolubile e, a nostro avviso, non vale la pena di cimentarsi in situazioni troppo complesse, che non capitano mai nelle applicazioni: ci limiteremo solo a fornire tre regole, che ci paiono quelle di più frequente uso.

Regola 1

Se il denominatore è del tipo

$$\sqrt[n]{a^m}, \quad \text{con} \quad n > m$$

la razionalizzazione si ottiene moltiplicando numeratore e denominatore per

$$\sqrt[n]{a^{n-m}}.$$

Si è supposto $n > m$, perché altrimenti il radicale si può semplificare, portando un opportuno termine fuori dalla radice.

Esempio 2.22. $\dfrac{1}{\sqrt{2}} = \dfrac{1}{\sqrt{2}} \dfrac{\sqrt{2}}{\sqrt{2}} = \dfrac{\sqrt{2}}{2}$.

Esempio 2.23. $\dfrac{1}{\sqrt[4]{3^5}} = \dfrac{1}{3\sqrt[4]{3}} = \dfrac{1}{3\sqrt[4]{3}} \dfrac{\sqrt[4]{3^3}}{\sqrt[4]{3^3}} = \dfrac{\sqrt[4]{3^3}}{9}$.

Regola 2

Se il denominatore è del tipo[9]

$$\sqrt{a} \pm \sqrt{b}$$

la razionalizzazione si ottiene moltiplicando numeratore e denominatore per

$$\sqrt{a} \mp \sqrt{b}$$

Esempio 2.24. $\dfrac{1}{\sqrt{5}+\sqrt{3}} = \dfrac{1}{\sqrt{5}+\sqrt{3}} \dfrac{\sqrt{5}-\sqrt{3}}{\sqrt{5}-\sqrt{3}} = \dfrac{\sqrt{5}-\sqrt{3}}{2}$.

Esempio 2.25. $\dfrac{1}{2-\sqrt{3}} = \dfrac{1}{2-\sqrt{3}} \dfrac{2+\sqrt{3}}{2+\sqrt{3}} = 2+\sqrt{3}$.

Regola 3

Se il denominatore è del tipo

$$\sqrt[3]{a} \pm \sqrt[3]{b}$$

la razionalizzazione si ottiene moltiplicando numeratore e denominatore per

$$\sqrt[3]{a^2} \mp \sqrt[3]{ab} + \sqrt[3]{b^2}.$$

Esempio 2.26. $\dfrac{1}{1-\sqrt[3]{2}} = \dfrac{1}{1-\sqrt[3]{2}} \dfrac{1+\sqrt[3]{2}+\sqrt[3]{4}}{1+\sqrt[3]{2}+\sqrt[3]{4}} = \dfrac{1+\sqrt[3]{2}+\sqrt[3]{4}}{1-2} = -\left(1+\sqrt[3]{2}+\sqrt[3]{4}\right)$.

Osservazione 2.22. Per giustificare l'importanza del calcolo con i radicali, di cui abbiamo già discusso nell'osservazione 2.21 nella pagina 63, proponiamo un interessante esempio delle difficoltà che si possono presentare eseguendo i calcoli con gli allineamenti decimali (che necessariamente devono essere approssimati quando intervengono gli irrazionali).

Si considerino le seguenti 7 espressioni:

$$1\colon \left(\frac{\sqrt{2}-1}{\sqrt{2}+1}\right)^3, \qquad 2\colon \left(\sqrt{2}-1\right)^6, \qquad 3\colon \left(3-2\sqrt{2}\right)^3, \qquad 4\colon \left(5\sqrt{2}-7\right)^2,$$

$$5\colon 99-70\sqrt{2}, \qquad 6\colon \frac{1}{\left(\sqrt{2}+1\right)^6}, \qquad \frac{1}{99+70\sqrt{2}},$$

tutte esprimenti lo stesso numero reale, come si può provare usando le proprietà dei radicali, e come si può "stimare" usando un buon software di calcolo (dove si ottiene, per ciascuna delle espressioni, $0.005050633883\dots$). Se approssimiamo $\sqrt{2}$ con $1,4$, compiendo un errore di poco più di un centesimo, si ottengono, nell'ordine, i seguenti risultati (arrotondati alla sesta cifra dopo la virgola):

$$0,00463; \quad 0,004096; \quad 0,008; \quad 0; \quad 1; \quad 0,005233; \quad 0,005076.$$

Come si vede alcune espressioni riducono addirittura l'errore sul dato, altre lo amplificano enormemente.

[9] Sia nella regola 2 che nella regola 3 si applicano alcuni "prodotti notevoli" di cui parleremo in seguito, ma che dovrebbero essere noti dalla scuola secondaria, precisamente quelli relativi alla differenza di quadrati o alla differenza di cubi.

2.5.6. Altri numeri reali

L'introduzione dei radicali e l'uso delle operazioni con essi ci permette di costruire molti numeri irrazionali. A questo punto potremmo chiederci: tutti i numeri irrazionali possono essere costruiti a partire dai radicali? La risposta è no! Anzi si potrebbe dimostrare che l'insieme di tutti i numeri che si possono costruire usando i radicali e le operazioni con essi (i numeri così costruiti fanno parte dell'insieme dei cosiddetti *irrazionali algebrici*) costituiscono ancora un insieme con cardinalità $\aleph_0$ (cioè un insieme numerabile): devono dunque esistere altri numeri reali, se sappiamo che l'insieme dei reali è non numerabile.

Non intendiamo esplorare ulteriormente questo terreno (minato!), e segnaliamo solo che, per esempio, π (cioè il rapporto tra una qualunque circonferenza e il suo diametro) così come il numero di Nepero, "e", *non* sono numeri irrazionali algebrici e fanno invece parte dell'insieme degli *irrazionali trascendenti* (che, sulla base di quanto detto, deve contenere un'infinità non numerabile di reali, cioè, "la quasi totalità" dei reali).

2.5.7. Potenze nei reali

Il concetto di potenza con esponente intero può essere esteso anche al caso di basi reali senza alcun cambiamento, né nella definizione, né nelle proprietà. L'introduzione dei radicali consente di estendere questo concetto anche al caso di *esponenti razionali*, con la seguente definizione:

Definizione 2.23. *Sia* $a \in \mathbb{R}$, $a > 0$ *e* $\frac{m}{n}$ *un numero razionale diverso da zero, con* $n > 1$, *allora*

$$a^{m/n} = \sqrt[n]{a^m}. \tag{2.33}$$

Se poi $r > 0$ *è un razionale, si pone* $0^r = 0$.

Si noti che la definizione (2.33) *non* si applica al caso di basi negative: il motivo è legato al fatto che l'estensione a basi negative comporterebbe difficoltà insormontabili. Citiamo qui solo un esempio di problema che potrebbe porsi, rinviando una trattazione più completa al capitolo 7 sulle potenze.

Esempio 2.27. Poiché $1/3 = 2/6$, se la definizione (2.33) si applicasse a basi negative, potrei avere la seguente catena di uguaglianze:

$$-1 = \sqrt[3]{-1} = (-1)^{1/3} = (-1)^{2/6} = \sqrt[6]{(-1)^2} = \sqrt[6]{1} = 1\,,$$

ovvero $-1 = 1$!

Il concetto di potenza si può poi estendere anche al caso di esponenti reali irrazionali: purtroppo la definizione è complessa e ne faremo solo un cenno nel capitolo 7 sulle potenze. Quello che ci preme segnalare fin da subito è che tutte le estensioni del concetto di potenza, in modo da comprendere esponenti via via più generali, è fatta con uno sguardo preciso alle proprietà (2.4): l'estensione viene fatta in modo da mantenere intatte proprio le proprietà formali delle potenze con esponente naturale.

2.5.8. Verso i numeri complessi

L'introduzione dei numeri reali risolve tutti i problemi che abbiamo via via incontrato e mette a disposizione un insieme numerico dotato di notevoli proprietà. Rimane ancora un problema: l'equazione

$$x^2 + 1 = 0$$

non si lascia proprio trattare, cioè non ha soluzioni nemmeno in $\mathbb{R}$. In realtà la richiesta che essa abbia soluzioni è proprio strana, in quanto richiede di trovare un numero il cui quadrato sia negativo e questo è un problema decisamente più complesso rispetto a quelli finora incontrati. La ricerca però di un insieme numerico che sia un'estensione dei reali e in cui un'equazione come questa abbia soluzioni ha successo e anzi permette di costruire un insieme che ha molte straordinarie proprietà: l'insieme dei numeri *complessi*. Ci occuperemo brevemente di questo insieme numerico solo nella parte seconda, in quanto normalmente non compreso nei programmi dei corsi di scuola secondaria.

2.6. Esercizi

Occorre avere una buona manualità nella risoluzione di esercizi con i radicali. Qui proponiamo solo alcuni spunti, che evidenziano le principali conoscenze richieste, invitando i lettori a riprendere in mano i testi della scuola secondaria per ulteriori problemi.

Esercizio 2.1. *Semplificare, fin dove possibile, le seguenti espressioni con radicali.*

1. $\sqrt[4]{(1+\sqrt{5})^2}$. Poiché $1+\sqrt{5}$ è positivo, si possono applicare le proprietà dei radicali e si ottiene $\sqrt{1+\sqrt{5}}$.
2. $\sqrt[6]{(2-\sqrt{5})^2}$. Poiché $2-\sqrt{5}$ è negativo, non si possono applicare direttamente le proprietà dei radicali. Avendosi però $\sqrt[6]{(2-\sqrt{5})^2} = \sqrt[6]{(\sqrt{5}-2)^2}$ si ottiene, con una semplificazione ora possibile, $\sqrt[3]{\sqrt{5}-2}$.
3. $\sqrt[9]{(2-\sqrt{5})^3}$. In questo caso $2-\sqrt{5}<0$ e anche $(2-\sqrt{5})^3<0$. Il radicale però è di indice dispari, e quindi ha senso. Si ottiene, per semplificazione, $\sqrt[3]{2-\sqrt{5}}$. Si noti che sia il radicale dato che quello ottenuto per semplificazione sono negativi.
4. $\sqrt[6]{(3-m)^2}$. Si osservi che $3-m$ è positivo per $m<3$, negativo per $m>3$. Si ha però: $(3-m)^2 = |3-m|^2$. Tenendo conto che $|3-m|\geq 0$ si ottiene $\sqrt[3]{|3-m|}$. Si noti che, se non si fosse preso il valore assoluto, il radicale di partenza sarebbe stato sempre positivo, quello di arrivo sarebbe stato negativo per $m>3$, pur continuando ad essere definito, in quanto radicale dispari.
5. $\sqrt{a+\sqrt{2a-1}}$. Perché il radicale quadratico più interno abbia senso deve essere $a\geq \frac{1}{2}$. Si ha inoltre $a^2-(2a-1)=(a-1)^2=|a-1|^2$. Ne segue

$$\sqrt{a+\sqrt{2a-1}} = \sqrt{\frac{a+\sqrt{|a-1|^2}}{2}} + \sqrt{\frac{a-\sqrt{|a-1|^2}}{2}} = \sqrt{\frac{a+|a-1|}{2}} + \sqrt{\frac{a-|a-1|}{2}} =$$

$$= \begin{cases} \sqrt{\frac{1}{2}} - \sqrt{\frac{2a-1}{2}}, & \text{se } \frac{1}{2}\leq a<1; \\ \sqrt{\frac{2a-1}{2}} - \sqrt{\frac{1}{2}}, & \text{se } a\geq 1. \end{cases}$$

6. $\sqrt{(a-2)^2(a-1)^3}$. Perché il radicale quadratico abbia senso, occorre che sia $a \geq 1$. Essendo poi $(a-2)^2 = |a-2|^2$, si ottiene $\sqrt{(a-2)^2(a-1)^3} = \sqrt{|a-2|^2(a-1)^2(a-1)} = |a-2|(a-1)\sqrt{a-1}$. Anche mettendo $|a-1|$ al posto di $(a-1)$ fuori dalla radice non cambia nulla, in quanto, come già osservato, deve essere $a \geq 1$.

7. $\sqrt[3]{81} - \sqrt[3]{24} + 10\sqrt[3]{3} - \sqrt[3]{\dfrac{3}{8}}$. L'espressione data equivale a $\sqrt[3]{3^3 \times 3} - \sqrt[3]{2^3 \times 3} + 10\sqrt[3]{3} - \sqrt[3]{\dfrac{3}{2^3}} = 3\sqrt[3]{3} - 2\sqrt[3]{3} + 10\sqrt[3]{3} - \dfrac{\sqrt[3]{3}}{2} = \left(3 - 2 + 10 - \dfrac{1}{2}\right)\sqrt[3]{3} = \dfrac{21}{2}\sqrt[3]{3}$.

Esercizio 2.2. *Ridurre ad un'unica frazione l'espressione che segue:*

$$\frac{a+\sqrt{b}}{a-\sqrt{b}} + \frac{\sqrt{b}-a}{a+\sqrt{b}} - \frac{2a\sqrt{b}}{a^2-b}.$$

Risoluzione. Conviene razionalizzare i denominatori delle prime due frazioni e poi operare. Si ottiene

$$\begin{aligned}\frac{a+\sqrt{b}}{a-\sqrt{b}}\,\frac{a+\sqrt{b}}{a+\sqrt{b}} + \frac{\sqrt{b}-a}{a+\sqrt{b}}\,\frac{a-\sqrt{b}}{a-\sqrt{b}} - \frac{2a\sqrt{b}}{a^2-b} &= \frac{\left(a+\sqrt{b}\right)^2}{a^2-b} + \frac{(\sqrt{b}-a)(a-\sqrt{b})}{a^2-b} - \frac{2a\sqrt{b}}{a^2-b} = \\ &= \frac{a^2+2a\sqrt{b}+b+a\sqrt{b}-b-a^2+a\sqrt{b}-2a\sqrt{b}}{a^2-b} = \frac{2a\sqrt{b}}{a^2-b}.\end{aligned}$$

Qualche calcolo poteva anche essere velocizzato, ma in un esercizio come questo forse non ne vale la pena. □

Esercizio 2.3. *Per quale valore della cifra a il numero* $3572a14$ *è divisibile per* 11*?*

Risoluzione. Il numero proposto ha 7 cifre. La somma delle cifre di posto pari è $4+a+7+3 = 14+a$, quella delle cifre di posto dispari è $1+2+5 = 8$. Si deve dunque imporre che $6+a$ sia divisibile per 11: $a = 5$. □

Esercizio 2.4. *Si provi che se un numero a è irrazionale, anche* $1/a$ *lo è.*

Risoluzione. Se $1/a$ fosse un razionale, diciamolo r, si avrebbe $1/r = a$, cioè $1/r$ sarebbe irrazionale, cosa palesemente falsa. □

Esercizio 2.5. *Trovare, usando l'algoritmo di Euclide, il massimo comun divisore dei tre numeri* 304920, 25725, 614922*. Ritrovare il risultato effettuando la scomposizione in fattori primi dei tre numeri.*

Risoluzione. Si può procedere trovando il MCD dei primi due numeri e, successivamente, il MCD tra questo risultato e il terzo numero. Si ha:

$$\begin{aligned}&304920 = 11 \times 25725 + 21945; \quad 25725 = 1 \times 21945 + 3780; \quad 21945 = 5 \times 3780 + 3045;\\ &\qquad 3780 = 1 \times 3045 + 735; \quad 3045 = 4 \times 735 + 105; \quad 735 = 7 \times 105 + 0,\end{aligned}$$

da cui $\mathrm{MCD}(304920, 25725) = 105$. Si ha poi:

$$614922 = 5856 \times 105 + 42; \quad 105 = 2 \times 42 + 21; \quad 42 = 2 \times 21 + 0,$$

da cui si conclude che il massimo comun divisore cercato è 21.

La scomposizione in fattori dei tre numeri porge:

$$\begin{aligned} 304920 &= 2^3 \times 3^2 \times 5 \times 7 \times 11^2, \\ 25725 &= 3 \times 5^2 \times 7^3, \\ 614922 &= 2 \times 3 \times 7 \times 11^4, \end{aligned}$$

da cui si trae la stessa conclusione di prima. □

Esercizio 2.6. *Quali proprietà dell'addizione servono per provare che, in* $\mathbb{N}$,

$$n_1 + n_2 + n_3 = n_3 + n_2 + n_1 \,?$$

Risoluzione. Abitualmente si pensa che la formula precedente (scambiano l'ordine dei termini in un'addizione qualunque il risultato non cambia) sia una semplice conseguenza della proprietà commutativa dell'addizione: essa richiede invece l'uso, e più volte ripetuto, delle proprietà commutativa e associativa. I passaggi sono i seguenti

$$n_1 + n_2 + n_3 \overset{(1)}{=} n_1 + (n_2 + n_3) \overset{(2)}{=} (n_2 + n_3) + n_1 \overset{(3)}{=} (n_3 + n_2) + n_1 \overset{(4)}{=} n_3 + n_2 + n_1,$$

di cui la (1) e la (4) hanno richiesto la proprietà associativa, la (2)e la (3) la proprietà commutativa. □

Esercizio 2.7. *Dimostrare che, per ogni* $n \in \mathbb{Z}$, $n^2 - 9n + 20$ *è un numero pari* ≥ 0.

Risoluzione. Si ha $n^2 - 9n + 20 = (n-4)(n-5)$. Dunque se $n = 4 \vee n = 5$ il numero proposto è 0, se $n < 4$ è il prodotto di due negativi, se $n > 5$ è il prodotto di due positivi. Inoltre per ogni n, almeno uno dei due numeri $n-4$ e $n-5$ è pari, dunque il numero dato è pari. □

Esercizio 2.8. *Si provi che il numero* $n(n^2+8)$ *è sempre divisibile per* 3.

Risoluzione. Se n è divisibile per 3 la cosa è ovvia, altrimenti n è del tipo $3m+1$ o $3m+2$. Nel primo caso si ha

$$\begin{aligned} n(n^2+8) = (3m+1)\big((3m+1)^2+8\big) &= (3m+1)(9m^2+6m+1+8) = \\ &= (3m+1)(9m^2+6m+9) = 3(3m+1)(3m^2+2m+3), \end{aligned}$$

e il numero è divisibile per 3. In modo analogo si procede se $n = 3m+2$. □

Esercizio 2.9. *Provare che,* $\forall n \in \mathbb{N}$ *il numero* $n^5 - n$ *termina con almeno uno* 0, *cioè è divisibile per* 10.

Risoluzione. Si ha $n^5 - n = n(n^4-1) = n(n^2-1)(n^2+1) = (n-1)n(n+1)(n^2+1)$. Dunque per ogni n, almeno uno dei numeri n e $n-1$ è pari, cioè il numero dato è divisibile per 2. Esaminiamo poi l'ultima cifra decimale del numero n e mostriamo che, in ogni caso, uno dei quattro numeri della scomposizione è divisibile per 5.

- Se l'ultima cifra è 0 non ci sono problemi;
- se l'ultima cifra è 1, allora $n-1$ termina per 0;
- se l'ultima cifra è 2, allora n^2+1 termina per 5;
- se l'ultima cifra è 3, allora n^2+1 termina per 0;
- se l'ultima cifra è 4, allora $n+1$ termina per 5;
- se l'ultima cifra è 5 non ci sono problemi;
- se l'ultima cifra è 6, allora $n-1$ termina per 5;
- se l'ultima cifra è 7, allora n^2+1 termina per 0;
- se l'ultima cifra è 8, allora n^2+1 termina per 5;
- se l'ultima cifra è 9, allora n^2+1 termina per 0.

Si può notare che questo problema implica che n ed n^5 terminano sempre con la stessa cifra. □

Esercizio 2.10. *Provare che, per ogni coppia di reali a e b non nulli,*

$$\frac{1}{a+b} \neq \frac{1}{a} + \frac{1}{b}.$$

Risoluzione. Supponiamo per assurdo che, nella formula precedente, valga il segno di uguale. Si avrebbe allora:

$$\frac{1}{a+b} = \frac{1}{a} + \frac{1}{b} \Rightarrow \frac{1}{a+b} = \frac{b+a}{ab} \Rightarrow ab = (a+b)^2,$$

da cui si dedurrebbe, intanto, che a e b devono essere concordi (perché il loro prodotto è positivo). Si avrebbe poi $a^2+b^2=-ab$ che è assurdo perché il primo membro è positivo mentre il secondo deve essere negativo. □

Esercizio 2.11. *Si dimostri che il prodotto di tre naturali consecutivi, maggiori di* 1*, non può essere il cubo di un naturale.*

Risoluzione. Sia $n \in \mathbb{N}$ e supponiamo per assurdo che $\exists m \in \mathbb{N}$ tale che

$$n(n+1)(n+2) = m^3.$$

Da qui si trae $m^3 = n^3+3n^2+2n > n^3$, da cui $m > n$. Inoltre $m^3 = n^3+3n^2+2n = n^3+3n^2+3n+1-n-1 = (n+1)^3-(n+1) < (n+1)^3$, da cui $m < n+1$. Ma questo è assurdo perché non esiste nessun naturale compreso tra n e $n+1$. □

Esercizio 2.12. *Provare che*

$$a = \sqrt[3]{2+\sqrt{5}} + \sqrt[3]{2-\sqrt{5}} = 1.$$

Risoluzione. Calcoliamo a^3. Si ha

$$\begin{aligned} a^3 &= 2+\sqrt{5}+3\sqrt[3]{(2+\sqrt{5})^2(2-\sqrt{5})}+3\sqrt[3]{(2+\sqrt{5})(2-\sqrt{5})^2}+2-\sqrt{5} = \\ &= 4+3\sqrt[3]{(2+\sqrt{5})(2+\sqrt{5})(2-\sqrt{5})}+3\sqrt[3]{(2+\sqrt{5})(2-\sqrt{5})(2-\sqrt{5})} = \\ &= 4+3\sqrt[3]{(2+\sqrt{5})(-1)}+\sqrt[3]{(-1)(2-\sqrt{5})} = 4-3\left(\sqrt[3]{2+\sqrt{5}}+\sqrt[3]{2-\sqrt{5}}\right) = 4-3a. \end{aligned}$$

Dunque il numero a soddisfa la condizione $a^3+3a-4=0$, equivalente a (scomposizione con Ruffini) $(a-1)(a^2+a+4)=0$. Da qui si ricava che può essere solo $a=1$. □

Esercizio 2.13. *Semplificare fin dove possibile l'espressione*

$$\sqrt[3]{4}+\frac{2-\sqrt[3]{-2}}{\sqrt[3]{4}}\sqrt[3]{-2}\,.$$

Risoluzione. Si ha

$$\sqrt[3]{4}+\frac{2-\sqrt[3]{-2}}{\sqrt[3]{4}}\sqrt[3]{-2}=\frac{\sqrt[3]{16}+2\sqrt[3]{-2}-\sqrt[3]{4}}{\sqrt[3]{4}}=\frac{2\sqrt[3]{2}+2\sqrt[3]{-2}-\sqrt[3]{4}}{\sqrt[3]{4}}=-1.$$ □

Esercizio 2.14. *Si verifichi che i numeri*

$$\sqrt[3]{4}+\sqrt[3]{6}+\sqrt[3]{9} \quad e \quad \sqrt[3]{3}+\sqrt[3]{-2}$$

sono reciproci.

Risoluzione. Basta provare che il prodotto vale 1. Si ha

$$\left(\sqrt[3]{4}+\sqrt[3]{6}+\sqrt[3]{9}\right)\left(\sqrt[3]{3}+\sqrt[3]{-2}\right)=\cancel{\sqrt[3]{12}}-\sqrt[3]{8}+\cancel{\sqrt[3]{18}}-\cancel{\sqrt[3]{12}}+\sqrt[3]{27}-\cancel{\sqrt[3]{18}}=-2+3=1.$$ □

Esercizio 2.15. *Si esprima il numero*

$$\sqrt{5}+\sqrt{3}$$

come radicale doppio.

Risoluzione. Si deve avere

$$\sqrt{5}+\sqrt{3}=\sqrt{a+\sqrt{b}}=\sqrt{\frac{a+\sqrt{a^2-b}}{2}}+\sqrt{\frac{a-\sqrt{a^2-b}}{2}}\,.$$

Basterà che

$$5=\frac{a+\sqrt{a^2-b}}{2} \quad \text{e} \quad 3=\frac{a-\sqrt{a^2-b}}{2}\,,$$

ovvero

$$\begin{cases} 10=a+\sqrt{a^2-b} \\ 6=a-\sqrt{a^2-b} \end{cases} \Rightarrow \begin{cases} \sqrt{a^2-b}=10-a \\ \sqrt{a^2-b}=a-6 \end{cases} \Rightarrow 10-a=a-6 \Rightarrow a=8 \Rightarrow b=60\,.$$ □

Esercizio 2.16. *Senza l'uso della calcolatrice si provi che i seguenti numeri irrazionali hanno la stessa parte decimale (sequenza di cifre dopo la virgola)*

$$\frac{\sqrt{5}+1}{2} \quad e \quad \frac{\sqrt{5}-1}{2}\,.$$

Risoluzione. I due numeri possono essere scritti nella forma:

$$\sqrt{\frac{5}{4}}+\frac{1}{2}=\sqrt{\frac{5}{4}}+0,5 \quad \text{e} \quad \sqrt{\frac{5}{4}}-\frac{1}{2}=\sqrt{\frac{5}{4}}-0,5.$$

Ora è chiaro che, qualunque sia l'espressione decimale del primo numero in questi addendi, le cifre dalla seconda in poi dopo la virgola non sono influenzate dalla somma o differenza. Rimane solo da valutare la prima cifra dopo la virgola. Si vede facilmente che, qualunque sia questa prima cifra per il primo addendo, sommandoci o sottraendoci 5 si ottiene, dopo la virgola, la stessa cifra. □

Esercizio 2.17. *Si provi che i numeri*

$$\frac{n}{37},$$

dove n non è un multiplo di 37*, hanno tutti il periodo costituito da sole tre cifre.*

Risoluzione. Il motivo è dovuto al fatto che

$$\frac{n}{37} = \frac{27n}{999},$$

e basta ricordare la regola per la frazione generatrice di un decimale periodico per concludere. Tra l'altro questo fatto implica la curiosa circostanza che

$$\frac{1}{27} = \frac{37}{999} = 0.\overline{037} \quad \text{e} \quad \frac{1}{37} = \frac{27}{999} = 0.\overline{027}. \qquad \square$$

Esercizio 2.18. *La scrittura*

$$0.777\overline{5777}$$

rappresenta un decimale periodico semplice o misto? E qual è il periodo del numero?

Risoluzione. La scrittura citata non rispetta le regole indicate per la scrittura dei decimali periodici. Espandendola si trova infatti

$$0.777\overline{5777} = 0.77757775777577757775777\dots,$$

da cui si capisce subito che la scrittura corretta è

$$0.\overline{7775},$$

e che il numero è decimale periodico semplice. In ogni caso applicando la regola per la frazione generatrice si ottiene, con entrambe le scritture,

$$\frac{7775}{9999},$$

da cui si capisce ancora meglio che si tratta di un decimale periodico semplice. $\square$

3. Algebra elementare

3.1. Monomi

3.1.1. Definizioni

Sappiamo che in Algebra, e non solo, si utilizzano lettere per rappresentare numeri. Ciò si fa, per esempio, quando si vogliano esprimere delle uguaglianze di carattere generale o quando si vogliano risolvere dei problemi.

Per esempio, l'area di un quadrato è espressa dall'uguaglianza $\mathcal{A} = l^2$, dove l indica la misura del lato. E ancora, il volume di una piramide a base rettangolare è dato dall'uguaglianza $\mathcal{V} = (1/3)abh$, dove a e b sono le misure dei lati del rettangolo di base e h è l'altezza della piramide. Espressioni come l^2 e $(1/3)abh$ si chiamano monomi. Più precisamente si dà la seguente definizione.

Definizione 3.1 (Monomio)**.** *Si chiama* monomio *il prodotto di fattori numerici per fattori letterali, i quali ultimi possono essere elevati a esponenti naturali. Il prodotto dei fattori numerici può essere ridotto ad un solo fattore numerico, che prende il nome di* coefficiente*; il prodotto dei fattori letterali prende il nome di* parte letterale*. Le lettere che figurano in un monomio si chiamano anche* indeterminate*.*

Sono monomi, per esempio, i seguenti:

— $3a^2bc^3$;

— $\dfrac{\sqrt{2}}{3}m^4n^2z$;

— $(\sqrt{2}+1)pq^2r$.

Alcuni autori chiamano monomi anche espressioni del tipo considerato nella definizione 3.1, in cui però le lettere possano avere anche esponenti negativi: in questo caso i monomi vengono detti frazionari, mentre nel caso (da noi considerato) di esponenti solo positivi per le lettere si usa l'espressione monomi interi. Poiché poi *tutti* usano il termine polinomio, che definiremo tra poco, per indicare la somma esclusivamente di monomi interi, e per l'importanza che hanno i polinomi in algebra, riteniamo preferibile attenerci alla definizione più restrittiva 3.1, scelta del resto adottata dalla stragrande maggioranza dei testi.

Naturalmente, se in un monomio non compare alcun fattore numerico, si può sempre pensare che il coefficiente sia 1. Se una lettera non compare esplicitamente, si può sempre pensare che essa compaia con esponente 0, e questo fatto è utile in molte circostanze. Se in un monomio una lettera compare più volte, si può riscrivere il monomio facendola comparire una sola volta usando le proprietà delle potenze. Per esempio a^2bac si può scrivere a^3bc. I monomi in cui compare un solo fattore numerico, che abitualmente viene scritto per primo, e in cui le singole lettere compaiono una sola volta si dice in *forma canonica* o *ridotta*: d'ora in poi supporremo sempre che i monomi siano scritti in forma canonica.

Di solito, inoltre, le lettere vengono scritte seguendo il loro ordine alfabetico. Due monomi con la stessa parte letterale si dicono *simili*.

Definizione 3.2 (Grado di un monomio). *In un monomio l'esponente di una lettera si chiama* grado del monomio rispetto a quella lettera*; la somma dei gradi di tutte le lettere si chiama semplicemente* grado del monomio*. I numeri reali si possono pensare come monomi di grado* 0*. Al polinomio nullo, cioè costituito solo dal numero reale* 0*,* non *si attribuisce alcun grado.*

Si tenga ben presente che, mentre al monomio costituito da un numero reale diverso da zero si attribuisce grado 0, al monomio nullo non si attribuisce alcun grado, e questo è legato alle proprietà del prodotto di monomi, di cui ora tratteremo.

3.1.2. Operazioni fra monomi

Tenendo conto del fatto che un monomio è un prodotto, e applicando le proprietà delle operazioni e delle potenze è possibile eseguire operazioni tra monomi, senza dover fornire nuove definizioni. Occorre però subito osservare che le usuali operazioni elementari tra due monomi non sempre producono monomi. Più precisamente si ha quanto segue.

Prodotto e potenza di monomi

Il prodotto di monomi e la potenza, con esponente naturale, di un monomio producono sempre un monomio. Per eseguirle basta applicare le proprietà della moltiplicazione e delle potenze. Proponiamo solo alcuni esempi.

- $\left(3a^2bc\right)\left(-2ab^4x\right) = -6a^3b^5cx$.
- $\left(2ab^2d^5\right)^3 = 8a^3b^6d^{15}$.
- $(-2a)\left(3a^2bx\right)\left(\frac{1}{9}ax^3\right) = -\frac{2}{3}a^4bx^4$.

Si noti che il grado del prodotto di due monomi non nulli è la somma dei gradi dei singoli monomi. Se uno dei monomi è il monomio nullo, il prodotto è ancora il monomio nullo: è proprio per rendere valida la regola sul grado del prodotto che *non* si attribuisce alcun grado al monomio nullo.

Quoziente di monomi

Il quoziente di due monomi può non produrre un monomio, in quanto applicando le usuali proprietà delle potenze e della moltiplicazione si possono ottenere lettere con esponente negativo[1]. Anche qui ci limitiamo solo a fornire alcuni esempi.

- $\dfrac{\sqrt{5}a^2bc^3}{3abc^2} = \dfrac{\sqrt{5}}{3}ac$: il risultato è un monomio.

[1]Secondo la nostra definizione in presenza di lettere con esponente negativo non si parla più di monomi. Se si fossero considerati anche i monomi frazionari, il quoziente di monomi avrebbe sempre fornito un monomio. Tuttavia la cosa è di poco interesse applicativo per cui, come già detto, non consideriamo monomi le espressioni in cui compaiono lettere con esponente negativo.

— $\dfrac{4a^2bc^3}{2ab^2c} = 2ab^{-1}c^2$: il risultato, che pure è un'espressione legittima e assume un ben determinato valore numerico se ad a, b e c si sostituiscono numeri ($b \neq 0$, naturalmente), non è però un monomio.

Somma di monomi

Anche la somma di monomi non fornisce generalmente un monomio. Solo la somma di monomi simili fornisce ancora un monomio, simile ai monomi addendi: basta applicare la proprietà distributiva. Nel caso di monomi non simili la somma …resta solo indicata. Come al solito forniamo qualche esempio.

— $3a^2bc^3 + 2a^2bc^3 - \sqrt{5}a^2bc^3 + \sqrt[3]{2}a^2bc^3 = \left(5 - \sqrt{5} + \sqrt[3]{2}\right)a^2bc^3$: la somma dei quattro monomi simili è ancora un monomio.
— $3abx^3 + 2abx + a^2b = 3abx^3 + 2abx + a^2b$: l'espressione non può essere scritta in maniera più semplice.

MCD e mcm di monomi

I monomi sono scritti come prodotti di più fattori elevati a certi esponenti; per questo la determinazione del MCD e del mcm si può fare esattamente come si fa con i naturali scomposti in fattori primi. Si conviene sempre, sia per il MCD che per il mcm, che il coefficiente sia 1. Per il MCD basterà poi prendere le lettere comuni con il minimo esponente, per il mcm le lettere comuni e non comuni con il massimo esponente. Per esempio:

$$\mathrm{MCD}\left(3abx^3, -2abx, \frac{3}{4}a^2b\right) = ab;$$
$$\mathrm{mcm}\left(3abx^3, -2abx, \frac{3}{4}a^2b\right) = a^2bx^3.$$

3.2. Polinomi

3.2.1. Generalità

Come già osservato, la somma di monomi non simili non è un monomio; tuttavia essa ha una grande importanza nelle applicazioni e per questo si dà la seguente definizione.

Definizione 3.3 (Polinomio)**.** *La somma algebrica di monomi non simili si chiama un* polinomio. *Per estensione si usa il termine polinomio anche per un singolo monomio. Si chiama* grado del polinomio *il massimo dei gradi dei monomi che lo compongono. Se tutti i monomi hanno lo stesso grado il polinomio si dice* omogeneo. *Se nel polinomio è presente un monomio di grado* 0 *(numero reale) esso si chiama anche* termine noto *del polinomio. Al polinomio nullo, cioè costituito solo dal numero reale* 0 *non si attribuisce alcun grado.*

La somma e il prodotto di polinomi si eseguono applicando opportunamente le note regole per la somma e prodotto di numeri reali; in particolare è qui molto importante la proprietà distributiva della moltiplicazione rispetto all'addizione. Il risultato di queste operazioni è sempre un polinomio: il grado della somma di polinomi di grado diverso è un polinomio con grado uguale al maggiore dei due gradi, il grado della somma di due polinomi dello stesso grado è minore o uguale al grado dei polinomi addendi; il grado del prodotto è uguale alla somma dei gradi dei polinomi fattori (e anche qui, come già nei monomi, è importante il fatto che al polinomio nullo non si attribuisce alcun grado).

Poiché la somma e il prodotto di polinomi non richiedono la scrittura di regole particolari, proponiamo solo alcuni esempi esplicativi.

— $(a+b)(c+d)(e+1)=(ac+ad+bc+bd)(e+1)=ace+ac+ade+ad+bce+bc+bde+bd$: qui abbiamo il prodotto di tre polinomi di grado 1 che produce un polinomio di grado 3.
— $(a^2b+b^3)(a-b)=a^3b-a^2b^2+ab^3-b^4$: qui abbiamo il prodotto di un polinomio di grado tre (omogeneo) per un polinomio di grado 1 (anch'esso omogeneo) che produce un polinomio di grado 4 (ancora omogeneo).
— $(a^2bc+ab-2a^2b)+(3a^2c^2+2ab^2+3a^2b)=a^2bc+ab+a^2b+3a^2c^2+2ab^2$: qui abbiamo la somma di due polinomi di quarto grado che produce ancora un polinomio di quarto grado.
— $(ab^3-2a^2b+ab)+(-ab^3+2a^2b+3ab^2)=ab+3ab^2$: qui abbiamo la somma di due polinomi di quarto grado che produce un polinomio di terzo grado.

3.2.2. La divisione tra polinomi

Come sappiamo, in un polinomio possono figurare più lettere, o indeterminate. In molti casi ha interesse fissare l'attenzione su alcune di queste, trattando le altre (se ci sono) alla stregua di coefficienti. Le indeterminate "vere e proprie" di solito si indicano con le ultime lettere dell'alfabeto, x, y, z, mentre le lettere che vanno trattate alla stregua di coefficienti si indicano con le prime lettere dell'alfabeto.

Esempio 3.1. Il polinomio nell'indeterminata x, ax^2+bx+c, è il più generale polinomio di secondo grado in una indeterminata.

Esempio 3.2. Il polinomio nelle indeterminate x e y, $ax^2+bxy+cy^2+dx+ey+f$, è il più generale polinomio in due indeterminate.

Abitualmente i polinomi si rappresentano con lettere, più spesso maiuscole, ma a volte anche minuscole: si dice il polinomio A, il polinomio B, ecc. Se si vuole mettere esplicitamente in evidenza come il polinomio dipende dalle indeterminate si usano scritture simili a quelle usate per le funzioni (e in realta, come vedremo, i polinomi danno origine alle funzioni polinomiali). Si scrive, per esempio, $A(x)=ax^2+bx+c$, $B(x,y)=ax^2+bxy+cy^2+dx+ey+f$. Per indicare il grado di un polinomio di usa la scrittura $\deg(P)$. Per esempio si ha $\deg(ax^2+bx+c)=2$.

Per i polinomi in una indeterminata useremo abitualmente la scrittura

$$P(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}+a_nx^n\,, \tag{3.1}$$

che mette in evidenza immediatamente a quale monomio nell'indeterminata x si riferisce ciascun coefficiente. Come già detto, a_0 si chiama *termine noto*, mentre a_n si chiama *primo coefficiente*.

Per i polinomi in una indeterminata sussiste il seguente teorema, analogo al 2.5 valido per la divisione tra naturali.

Teorema 3.4. *Dati due polinomi in una indeterminata N e D, con $\deg(N)\geq\deg(D)$, e D diverso dal polinomio nullo, esiste una ed una sola coppia di polinomi Q e R tali che*

1. $N = QD + R$,
2. $\deg(R) < \deg(D)$.

Inoltre $\deg(Q) = \deg(N) - \deg(D)$. *Il polinomio* Q *si chiama* quoziente*, il polinomio* R resto*. Se* R *è il polinomio nullo, si dice che* N *è* divisibile *per* D*, o che* D *divide* N *e si scrive, analogamente a quanto fatto per i numeri naturali,*

$$D \mid N \qquad \textit{o anche} \qquad D(x) \mid N(x) .$$

Nel teorema 3.4 si sarebbe potuto anche considerare il caso $\deg(N) < \deg(D)$, ma allora Q sarebbe il polinomio nullo e R coinciderebbe con N: è dunque un caso di nessun interesse.

Come nella divisione tra numeri, N si chiama *dividendo* e D *divisore*, ovvero *numeratore* e *denominatore*: proprio per questo abbiamo usato le lettere N e D.

La ricerca dei polinomi Q ed R si fa con un algoritmo del tutto simile a quello della divisione con resto tra numeri naturali e la illustreremo con alcuni esempi.

Cominciamo a considerare la divisione tra $N(x) = x^5 + x^2$ e $D(x) = x^3 - x^2$.Il primo passo consiste nello scrivere i monomi di N e D secondo l'ordine decrescente dei gradi. Successivamente si considera uno schema come nella divisione tra numeri, lasciando, nel polinomio N, uno spazio vuoto al posto dei monomi "mancanti" tra quello di grado più alto e quello di grado più basso.

$$\begin{array}{cccc|c}
x^5 & & & +x^2 & x^3-x^2 \\ \cline{5-5}
 & & & &
\end{array}$$

Poi si divide il primo monomio di N per il primo monomio di D, scrivendo il risultato (che possiamo chiamare primo quoziente parziale) sotto il divisore.

$$\begin{array}{cccc|c}
x^5 & & & +x^2 & x^3-x^2 \\ \cline{5-5}
 & & & & x^2
\end{array}$$

Si calcola il primo resto parziale sottraendo al dividendo il prodotto tra il divisore e il primo quoziente parziale.

$$\begin{array}{cccc|c}
x^5 & & & +x^2 & x^3-x^2 \\ \cline{5-5}
-x^5 & +x^4 & & & x^2 \\ \cline{1-4}
 & x^4 & & +x^2 &
\end{array}$$

Si noti che sotto al dividendo abbiamo scritto il prodotto tra il divisore e il primo quoziente parziale con i segni cambiati: questo ci consente di fare una addizione anziché una sottrazione, e ciò di solito aiuta ad evitare errori.

A questo punto il procedimento si ripete, fintanto che si ottiene un resto di grado minore di $\deg(D)$. Il passo successivo sarà:

$$\begin{array}{cccc|c}
x^5 & & & +x^2 & x^3-x^2 \\ \cline{5-5}
-x^5 & +x^4 & & & x^2+x \\ \cline{1-4}
 & x^4 & & +x^2 & \\
 & -x^4 & +x^3 & & \\ \cline{2-4}
 & & x^3 & +x^2 &
\end{array}$$

L'ultimo passo sarà:

$$
\begin{array}{rrrr|l}
x^5 & & & +x^2 & x^3-x^2 \\ \cline{5-5}
-x^5 & +x^4 & & & x^2+x+1 \\ \cline{1-4}
 & x^4 & & +x^2 & \\
 & -x^4 & +x^3 & & \\ \cline{2-4}
 & & x^3 & +x^2 & \\
 & & -x^3 & +x^2 & \\ \cline{3-4}
 & & & 2x^2 &
\end{array}
$$

Il quoziente è $Q(x)=x^2+x+1$, il resto $R(x)=2x^2$ e si ha

$$
x^5+x^2=(x^2+x+1)(x^3-x^2)+2x^2 .
$$

Come ulteriore esempio consideriamo la divisione tra $2x^4+x^3+x+2$ e x^2+3.

$$
\begin{array}{rrrrr|l}
2x^4 & +x^3 & & +x & +2 & x^2+3 \\ \cline{6-6}
-2x^4 & & -6x^2 & & & 2x^2+x-6 \\ \cline{1-5}
 & x^3 & -6x^2 & +x & +2 & \\
 & -x^3 & & -3x & & \\ \cline{2-5}
 & & -6x^2 & -2x & +2 & \\
 & & 6x^2 & & +18 & \\ \cline{3-5}
 & & & -2x & +20 &
\end{array}
$$

Si ha dunque $2x^4+x^3+x+2=(2x^2+x-6)(x^2+3)-2x+20$.

Si presti attenzione al fatto che il teorema 3.4 si applica ai polinomi in una indeterminata: nel caso di più indeterminate bisogna sceglierne una e trattare le altre come coefficienti, e la scelta di quale debba essere l'indeterminata può portare a risultati diversi. Lo vediamo nel caso dei polinomi a^3+ab-b^2 e $a-b$. Si può considerare[(2)] $N(a)=a^3+ba-b^2$ e $D(a)=a-b$, oppure $N(b)=-b^2+ab+a^3$ e $D(b)=-b+a$. Si ottiene, nei due casi:

$$
\begin{gathered}
Q(a)=a^2+ab+b^2+b \quad \text{e} \quad R(a)=b^3, \quad \text{da cui} \quad a^2+ab-b^2=(a^2+ab+b^2+b)(a-b), \\
Q(b)=b \quad \text{e} \quad R(b)=a^3, \quad \text{da cui} \quad -b^2+ab+a^3=b(-b+a)+a^3 .
\end{gathered}
$$

3.3. Prodotti notevoli e scomposizione di polinomi

In molti casi il calcolo del prodotto di polinomi si può fare utilizzando alcuni accorgimenti che riducono i tempi di esecuzione: si tratta dei cosiddetti *prodotti notevoli*.

[2]Si noti che in $N(a)$ abbiamo scritto la b prima della a, perché b è il coefficiente, mentre in $N(b)$ abbiamo scritto la a prima della b, perché ora a è il coefficiente.

Un'altra questione tecnica di grande importanza nelle applicazioni è quella della scomposizione di un polinomio nel prodotto di due o più polinomi: si parla di *scomposizione di un polinomio in fattori*. Gli stessi prodotti notevoli possono servire a questo scopo, come vedremo.

Come è noto, la scomposizione di un numero naturale in fattori primi ha, tra gli altri, i seguenti due scopi:

— trovare divisori comuni a due o più numeri, eventualmente il massimo comun divisore;
— trovare multipli comuni a due o più numeri, più piccoli che non il semplice prodotto dei numeri stessi, eventualmente il minimo comune multiplo.

Anche per i polinomi si pongono problemi analoghi e si introducono i concetti di divisore comune (eventualmente massimo comun divisore) e di multiplo comune (eventualmente minimo comune multiplo) e le definizioni sono le stesse date nel caso dei numeri interi: la difficoltà è tecnica, in quanto scomporre un polinomio è decisamente più complesso che non scomporre un numero intero e non si possono fornire regole generali[3]. Tratteremo alcune delle situazioni più comuni, avvertendo che avremo modo di tornare successivamente sull'argomento.

Prodotto di una somma per una differenza

Se A e B sono due polinomi (eventualmente costituiti da un solo monomio) si ha:

$$(A-B)(A+B)=A^2-B^2\,. \tag{3.2}$$

Letta da destra a sinistra questa formula permette di scomporre una differenza di quadrati in un prodotto.

Esempio 3.3. $\left(\frac{1}{3}a-2b\right)\left(\frac{1}{3}a+2b\right)=\frac{1}{9}a^2-4b^2\,.$

Esempio 3.4. $(3x+y-3)(3x+y+3)=\big((3x+y)-3\big)\big((3x+y)+3\big)=(3x+y)^2-9=\ldots=9x^2+6xy+y^2-9\,.$

Esempio 3.5. $2a^4-b^2=\left(\sqrt{2}a\right)^2-b^2=\left(\sqrt{2}a-b\right)\left(\sqrt{2}a+b\right).$

Esempio 3.6. $a^4-9=(a^2-3)(a^2+3)=\left(a-\sqrt{3}\right)\left(a+\sqrt{3}\right)(a^2+3)\,.$

Esempio 3.7. $(a-b)^2-4=(a-b-2)(a-b+2).$

Quadrato di un binomio

Se A e B sono due polinomi (eventualmente costituiti da un solo monomio) si ha:

$$(A\pm B)^2=A^2\pm 2AB+B^2\,. \tag{3.3}$$

Letta da destra a sinistra questa formula permette di trasformare certe somme di tre addendi in un quadrato: occorre individuare due quadrati e un doppio prodotto.

Esempio 3.8. $\left(-\frac{1}{4}xyz^2+\frac{2}{5}x^2y\right)^2=\left(-\frac{1}{4}xyz^2\right)^2+2\left(-\frac{1}{4}xyz^2\right)\left(\frac{2}{5}x^2y\right)+\left(\frac{2}{5}x^2y\right)^2=$
$=\frac{1}{16}x^2y^2z^4-\frac{1}{5}x^3y^2z^2+\frac{4}{25}x^4y^2\,.$

Esempio 3.9. $(3a-2b-c)^2=\big((3a-2b)-c\big)^2=(3a-2b)^2-2(3a-2b)c+c^2=9a^2-12ab+4b^2-6ac+4bc+c^2.$

[3]Ricordiamo comunque che anche la scomposizione di un naturale in un prodotto di primi può essere molto difficile: vedi per esempio quanto scritto nella pagina 43 relativamente a un problema posto da Martin Gardner.

Esempio 3.10. $(a-2b)^4=\left((a-2b)^2\right)^2=\left(a^2-4ab+4b^2\right)^2=\left((a^2-4ab)+4b^2\right)^2=$
$=(a^2-4ab)^2+2(a^2-4ab)4b^2+16b^4=a^4-8a^3b+16a^2b^2+8a^2b^2-32ab^3+16b^4=\dots$

Esempio 3.11. $4x^2+25y^2-20xy=(2x)^2+(5y)^2-2(2x)(5y)=(2x-5y)^2$.

Esempio 3.12. $\left(-\frac{1}{4}a^2+a-1\right)=-\left(\frac{1}{4}a^2-a+1\right)=\dots=-\left(\frac{1}{2}a-1\right)^2$.

Cubo di un binomio

Se A e B sono due polinomi (eventualmente costituiti da un solo monomio) si ha:

$$(A\pm B)^3=A^3\pm 3A^2B+3AB^2\pm B^3\,. \tag{3.4}$$

Letta da destra a sinistra questa formula permette di trasformare certe somme di quattro addendi in un cubo: occorre individuare due cubi e due opportuni tripli prodotti.

Esempio 3.13. $\left(\frac{1}{4}a-2b\right)^3=\left(\frac{1}{4}a\right)^3-3\left(\frac{1}{4}a\right)^2(2b)+3\left(\frac{1}{4}a\right)(2b)^2+(2b)^3=$
$=\frac{1}{64}a^3-3\left(\frac{1}{16}a^2\right)(2b)+3\left(\frac{1}{4}a\right)(4b^2)+8b^3=\frac{1}{64}a^3-\frac{3}{8}a^2b+3ab^2+8b^3\,.$

Esempio 3.14. $(a-b+c)^3=\left((a-b)+c\right)^3=(a-b)^3+3(a-b)^2c+3(a-b)c^2+c^3=$
$=a^3-3a^2b+3ab^2-b^3+3(a^2-2ab+b^2)c+3(a-b)c+c^3=$
$=a^3-3a^2b+3ab^2-b^3+3a^2c-6abc+3b^2c+3ac-3bc+c^3\,.$

Potenza n-esima di un binomio

Le formule per il quadrato e cubo di un binomio possono essere generalizzate ad un potenza n-esima, facendo uso del *triangolo aritmetico*, o di *Tartaglia*, o di *Pascal*. Le righe di questo triangolo contengono i coefficienti dei monomi del polinomio che si ottiene calcolando $(a+b)^n$, con $n=0$ (prima riga), $n=1$ (seconda riga), ecc. La parte letterale dei monomi è costruita prendendo i prodotti tra le potenze di a e quelle di b, quelle di a in ordine decrescente a partire da n fino a 0, quelle di b in ordine crescente a partire da 0 fino a n.

Tabella 3.1.: *Il triangolo aritmetico, o di Tartaglia, o di Pascal*

$n=0$							1						
$n=1$						1		1					
$n=2$					1		2		1				
$n=3$				1		3		3		1			
$n=4$			1		4		6		4		1		
$n=5$		1		5		10		10		5		1	
$n=\dots$	1		$\cdots$								$\cdots$		1

Le righe del triangolo, vedi la tabella 3.1, sono costruite iniziando e terminando con 1 e successivamente per ogni posizione facendo la somma dei due termini della riga precedente situati a destra e a sinistra della posizione in esame. Per esempio nella riga $n=5$ si ha $5=1+4$, $10=4+6$, $10=6+4$, $5=4+1$. È evidente che la costruzione di una riga richiede la costruzione di tutte le righe precedenti. Tuttavia è possibile usare anche la formula del binomio di Newton, di cui parleremo nel capitolo 12 sul calcolo combinatorio, e che è di uso più immediato.

Per calcolare, per esempio, $(a+b)^5$, i monomi dello sviluppo saranno 6 e avranno precisamente le seguenti parti letterali: $a^5b^0=a^5$, a^4b, a^3b^2, a^2b^3, ab^4, $a^0b^5=b^5$, mentre i coefficienti saranno quelli contenuti nella riga con $n=5$ del triangolo aritmetico. Si ottiene:

$$(a+b)^5=a^5+5a^4b+10a^3b^2+10a^2b^3+5ab^4+b^5\,.$$

Si noti che le formule per il quadrato e il cubo sono un caso particolare di questa formula più generale: solo la grande frequenza del loro uso ne giustifica una memorizzazione separata.

Nel caso di $(a-b)^n$ si potrà pensare di scrivere $\big(a+(-b)\big)^n$ oppure, equivalentemente, prendere i coefficienti con segno alterno, a partire dal segno "+".

Esempio 3.15. Calcoliamo nuovamente $(a-2b)^4$, già sviluppato più sopra, utilizzando il triangolo aritmetico. Si ha:

$$(a-2b)^4=a^4-4(a^3)(2b)+6a^2(2b)^2-4a(2b)^3+(2b)^4=a^4-8a^3b+24a^2b^2-32ab^3+16b^4\,.$$

Quadrato di un polinomio

Il quadrato di un polinomio si può calcolare riducendo il polinomio a una somma di due soli addendi, come abbiamo fatto in qualcuno degli esempi già trattati. Tuttavia è utile memorizzare una formula specifica che velocizza ulteriormente il calcolo.

$$(A+B+C+D+\dots)^2= \\ =A^2+B^2+C^2+D^2+\cdots+2AB+2AC+2AD+\cdots+2BC+2BD+\cdots+2CD+\dots\,. \tag{3.5}$$

Questa formula, a parole, si può esprimere così: *si fa la somma del quadrato di tutti gli addendi e del doppio prodotto di ogni addendo per tutti quelli che lo seguono*. Se c'è qualche sottrazione, conviene scriverla come addizione:

$$(a-b-c)^2=\big(a+(-b)+(-c)\big)^2=a^2+(-b)^2+(-c)^2+2a(-b)+2a(-c)+2(-b)(-c)= \\ =a^2+b^2+c^2-2ab-2ac+2bc\,.$$

A volte può essere utile leggere la formula (3.5) da destra a sinistra, per trasformare una somma in una potenza, ma la cosa non è sempre facile.

Raccoglimento a fattor comune

È questa la tecnica più comune per scomporre un polinomio in un prodotto, generalmente, ma non solo, di un monomio per un polinomio. Si tratta semplicemente di applicare la proprietà distributiva della moltiplicazione rispetto all'addizione, leggendola da destra a sinistra.

Esempio 3.16. $a^2+3ab+a^2x+3ay=a(a+3b+ax+3y)$.

Esempio 3.17. $(a+b)^2-a-b=(a+b)^2-(a+b)=(a+b)(a+b-1)$.

In molte situazioni questa tecnica si applica anche se non c'è un fattore comune. Vediamone l'applicazione su un esempio.

$$x^2+3x+5=x^2\left(1+\frac{3}{x}+\frac{5}{x^2}\right).$$

Naturalmente in questo caso occorrerà poter supporre $x \neq 0$. Inoltre non si ottiene qui la scomposizione di un polinomio nel prodotto di polinomi, tuttavia questa tecnica è spesso utile, per esempio, nel calcolo dei limiti di cui di avrà modo di parlare nei corsi di analisi.

Raccoglimenti successivi

Questa tecnica consiste sostanzialmente nella applicazione del raccoglimento a fattor comune anziché su tutto il polinomio su alcuni termini del polinomio stesso, per poter successivamente applicare il raccoglimento a fattor comune.

Esempio 3.18. $ax+ay+bx+by=a(x+y)+b(x+y)=(x+y)(a+b)$.

Esempio 3.19. $a^2x-a^2-x+1=a^2(x-1)-(x-1)=(x-1)(a^2-1)=(x-1)(a-1)(a+1)$.

Esempio 3.20. $(x+2y)^2-3x^2-6xy+(2x+4y)(a+4y)=(x+2y)^2-3x(x+2y)+2(x+2y)(a+4y)=$
$=(x+2y)(x+2y-3x+2a+8y)=(x+2y)(2y-2x+2a+8y)$.

Esempio 3.21. $n\geq 1$, $xy^{n+3}+x^2y^6-2y^{n-1}-2xy^2=y^{n-1}(xy^4-2)+xy^2(xy^4-2)=(xy^4-2)(y^{n-1}+xy^2)$.

Differenza di potenze n-esime

Abbiamo già esaminato il caso della differenza di due quadrati, come "regola inversa" del prodotto di una somma per una differenza. Esiste una formula generale anche per la differenza di due potenze n-esime.

$$A^n-B^n=(A-B)\big(A^{n-1}+A^{n-2}B+A^{n-3}B^2+\cdots+AB^{n-2}+B^{n-1}\big). \tag{3.6}$$

Di uso molto frequente il caso $n=3$, che vale la pena di memorizzare a parte:

$$a^3-b^3=(a-b)(a^2+ab+b^2).$$

Il secondo fattore nella decomposizione precedente si chiama anche un "falso quadrato", perché, rispetto al quadrato di un binomio, contiene il prodotto dei due fattori, anziché il doppio prodotto.

Differenza di potenze n-esime pari

Oltre alla formula (3.6) per la differenza di due potenze n-esime, nel caso di n pari, valgono anche le seguenti, conseguenze della (3.6) stessa e della regola per la differenza di quadrati.

$$n \text{ pari } \Rightarrow A^n-B^n=(A+B)\big(A^{n-1}-A^{n-2}B+A^{n-3}B^2-\cdots+AB^{n-2}-B^{n-1}\big). \tag{3.7}$$

$$n \text{ pari } \Rightarrow A^n-B^n=(A-B)(A+B)\big(A^{n-2}+A^{n-4}B^2+\cdots+A^2B^{n-4}+B^{n-2}\big). \tag{3.8}$$

Somma di potenze n-esime dispari

Una formula analoga alla (3.6) per la somma di potenze n-esime esiste solo per n dispari: si presti la massima attenzione a questo fatto.

$$n \text{ dispari } \Rightarrow A^n+B^n=(A+B)\big(A^{n-1}-A^{n-2}B+A^{n-3}B^2-\cdots AB^{n-2}+B^{n-1}\big). \tag{3.9}$$

Anche qui è di uso molto frequente il caso $n=3$, che vale la pena di memorizzare a parte:

$$a^3+b^3=(a+b)(a^2-ab+b^2).$$

Il secondo fattore nella decomposizione precedente si chiama ancora un "falso quadrato", perché, rispetto al quadrato di un binomio, contiene il prodotto dei due fattori, anziché il doppio prodotto.

Il fatto che la formula (3.9) valga solo nel caso di n dispari, non impedisce sempre di scomporre la somma di due potenze pari. Per esempio si ha

$$a^6+b^6=\left(a^2\right)^3+\left(b^2\right)^3=\left(a^2+b^2\right)\left(a^4-a^2b^2+b^4\right).$$

Altre volte bisogna usare qualche accorgimento particolare e ricondursi ad altri casi:

$$a^4+b^4=a^4+2a^2b^2+b^4-2a^2b^2=\left(a^2+b^2\right)^2-\left(\sqrt{2}ab\right)^2=\left(a^2+b^2+\sqrt{2}ab\right)\left(a^2+b^2-\sqrt{2}ab\right).$$

3.4. Divisibilità e scomposizione

Una conseguenza immediata del teorema 3.4 è che, se il polinomio N è divisibile per il polinomio D, allora

$$N=QD\,,$$

cioè il polinomio N risulta scomposto nel prodotto dei fattori Q e D. Per esempio, essendo

$$(x^7+3x^4-3x^3+6x^2-2x+3):(x^3-2x+3)=x^4+2x^2+1\,,$$

si ha

$$x^7+3x^4-3x^3+6x^2-2x+3=(x^4+2x^2+1)(x^3-2x+3)\,.$$

Il caso dei polinomi in una indeterminata è di grande importanza per quanto riguarda il problema della divisibilità e della scomposizione e lo tratteremo con maggiore dettaglio. Nel resto di questo paragrafo i polinomi saranno sempre in una sola indeterminata.

Definizione 3.5. *Un numero reale a si dice uno* zero del polinomio $P(x)$ *nella sola indeterminata x se* $P(a)=0$.

Il teorema fondamentale riguardante i polinomi in una indeterminata è il seguente

Teorema 3.6 (Teorema del resto o di Ruffini)**.** *Siano dati un polinomio $P(x)$, con $\deg(P)=n$, e un numero reale a. Esiste allora un polinomio $Q(x)$, di grado $n-1$, tale che*

$$P(x)=(x-a)Q(x)+P(a)\,. \tag{3.10}$$

Detto in altri termini e tenendo conto del teorema 3.4, $P(a)$ è il resto della divisione di $P(x)$ per $x-a$. Dunque $P(x)$ è divisibile per $x-a$ se e solo se $P(a)=0$. In questo caso il polinomio si scompone in

$$P(x)=(x-a)Q(x).$$

Non è affatto escluso che $Q(x)$ sia ancora divisibile per $x-a$: se questo succedesse si avrebbe

$$P(x)=(x-a)Q(x)=(x-a)(x-a)Q_1(x)=(x-a)^2Q_1(x)\,,$$

e la cosa potrebbe ancora proseguire. In proposito si dà la seguente definizione.

Definizione 3.7 (Molteplicità). *Se $P(x)$ è un polinomio di grado n, e a è uno zero del polinomio tale che*

$$P(x)=(x-a)^m Q(x), \quad m \leq n,$$

mentre a non è uno zero di $Q(x)$, a si chiama uno zero con molteplicità m. *Se $m=1$, a si dice uno* zero semplice.

Esempio 3.22. Essendo $x^3-3x^2+3x+1=(x-1)^3$, 1 è uno zero con molteplicità 3 o, come si usa dire, uno zero triplo del polinomio dato.

Esempio 3.23. Essendo $x^4-4x^3+5x^2-4x+4=(x-2)^2(x^2+1)$, 2 è uno zero con molteplicità 2, o doppio, per il polinomio dato; non esistono altri zeri (reali) per il polinomio dato.

Esempio 3.24. Essendo $x^3-x^2-8x+12=(x-2)^2(x+3)$, 2 è zero doppio, mentre -3 è zero semplice per il polinomio dato.

Il seguente teorema è una facile conseguenza del teorema 3.6.

Teorema 3.8. *Se un polinomio $P(x)$ è divisibile per $x-a$ e per $x-b$, con $b \neq a$, allora è anche divisibile per il prodotto $(x-a)(x-b)$.*

Dimostrazione. Se $P(x)$ è divisibile per $x-a$ e per $x-b$ si ha, contemporaneamente,

$$P(x)=(x-a)Q_1(x) \quad \text{e} \quad P(x)=(x-b)Q_2(x),$$

da cui $(x-a)Q_1(x)=(x-b)Q_2(x)$. Ponendo a al posto di x si trova $0=(a-b)Q_2(a)$, ovvero $Q_2(a)=0$. Ma allora $Q_2(x)$ è divisibile per $x-a$ e quindi $Q_2(x)=(x-a)Q_3(x)$. In conclusione

$$P(x)=(x-b)(x-a)Q_3(x). \qquad \square$$

Valgono anche i seguenti teoremi.

Teorema 3.9. *Un polinomio di grado n non può avere più di n zeri.*

Teorema 3.10. *Se un polinomio di grado minore o uguale a n ha $n+1$ zeri, allora è il polinomio identicamente nullo.*

Teorema 3.11 (Principio di identità dei polinomi). *Se due polinomi, di grado minore o uguale a n assumono valori uguali per $n+1$ valori distinti della indeterminata x, allora sono identicamente uguali.*

Importanti conseguenze di questi teoremi sono i due fatti seguenti (che nella sostanza sono due modi diversi di esprimere la stessa proprietà):

1. Un polinomio è identicamente nullo se e solo tutti i suoi coefficienti sono nulli.
2. Due polinomi sono identici se e solo se hanno gli stessi coefficienti.

Molto importante per la ricerca degli zeri di un polinomio è il teorema seguente.

Teorema 3.12 (Zeri razionali di un polinomio). *Un polinomio a coefficienti interi*

$$P(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}+a_nx^n, \quad \text{con } a_n \neq 0$$

ha uno zero razionale $\alpha = p/q$, con p e q interi e primi tra di loro, se e solo se p è un divisore del termine noto a_0 e q è un divisore del primo coefficiente a_n.

Si noti che se il polinomio ha $a_n = 1$, nel qual caso si dice *monico*, gli eventuali zeri razionali sono addirittura interi.

In sostanza per cercare gli eventuali zeri razionali di un polinomio si scrivono tutti i divisori del termine noto e tutti divisori del primo coefficiente; successivamente si scrivono tutte le frazioni con al numeratore un divisore del termine noto e al denominatore un divisore del primo coefficiente: poiché queste frazioni sono in numero finito, è possibile fare, anche manualmente, la ricerca degli eventuali zeri razionali del polinomio.

Esempio 3.25. Consideriamo il polinomio $P(x) = 15x^5 + 5x^4 + 3x^2 - 95x - 32$. I divisori del termine noto e del primo coefficiente sono, rispettivamente,

$$\{\pm 1, \pm 2, \pm 4, \pm 8, \pm 16, \pm 32\} \quad \text{e} \quad \{\pm 1, \pm 3, \pm 5, \pm 15\}.$$

I possibili zeri razionali sono allora:

$$\pm 1, \pm\frac{1}{3}, \pm\frac{1}{5}, \pm\frac{1}{15}, \pm 2, \pm\frac{2}{3}, \pm\frac{2}{5}, \pm\frac{2}{15}, \pm 4, \pm\frac{4}{3}, \pm\frac{4}{5}, \pm\frac{4}{15}, \pm 8, \pm\frac{8}{3}, \pm\frac{8}{5}, \pm\frac{8}{15},$$
$$\pm 16, \pm\frac{16}{3}, \pm\frac{16}{5}, \pm\frac{16}{15}, \pm 32, \pm\frac{32}{3}, \pm\frac{32}{5}, \pm\frac{32}{15}.$$

Con un po' di pazienza (!) si trova che solo $-1/3$ è uno zero: il polinomio potrà avere altri zeri (massimo altri quattro), ma essi, se ci sono, dovranno essere necessariamente irrazionali.

Dal punto di vista della scomposizione di un polinomio (in una indeterminata) in fattori è evidente l'importanza della determinazione degli eventuali zeri. Per la ricerca degli zeri razionali abbiamo a disposizione lo strumento fornito dal teorema 3.12. Nulla di simile, invece, per la ricerca degli eventuali zeri irrazionali. Esistono delle formule per i polinomi di grado minore o uguale a quattro, ma solamente quelle relative ai polinomi di primo e secondo grado sono alla nostra portata: le formule per i polinomi di terzo e quarto grado richiedono l'uso dei numeri complessi e quindi esulano dagli scopi di questo testo. Non esistono, né si possono trovare[4], formule per i polinomi di grado superiore al quarto.

3.5. Zeri di polinomi di primo e secondo grado

Per i polinomi di primo grado la determinazione dell'unico zero è immediata: se $P(x) = ax + b$, $a \neq 0$, $-b/a$ è palesemente uno zero ed è l'unico.

Occupiamoci dei polinomi di secondo grado. Si possono fare le seguenti manipolazioni, note come "tecnica del completamento dei quadrati": consigliamo vivamente di memorizzare i passaggi perché sono importanti tanto quanto il risultato che ne ricaveremo.

[4] Già Paolo Ruffini (nei primi anni del 1800) aveva fornito una dimostrazione parziale dell'impossibilità di fornire una formula generale che permetta, tramite operazioni di addizione, sottrazione, moltiplicazione, divisione e estrazioni di radici sui coefficienti, di trovare gli zeri di un polinomio di grado superiore al quarto. Questa dimostrazione fu completata nel 1824 da Niels Henrik Abel, con un teorema che prende il nome di Teorema di Ruffini-Abel.

Sia dato un polinomio di secondo grado, $P(x)=ax^2+bx+c$, $a\neq 0$. Si ha

$$ax^2+bx+c=a\left(x^2+\frac{b}{a}x+\frac{c}{a}\right)=a\left(x^2+\frac{b}{a}x+\frac{b^2}{4a^2}-\frac{b^2}{4a^2}+\frac{c}{a}\right)= \\ =a\left(\left(x+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a^2}\right). \tag{3.11}$$

Da qui si desume subito che se $b^2-4ac<0$ il polinomio non può avere zeri reali. Se $b^2-4ac=0$ il polinomio si riduce a

$$a\left(\left(x+\frac{b}{2a}\right)^2\right)$$

e ha un'unica radice reale, precisamente $-b/2a$. Se, infine, $b^2-4ac>0$ la quantità tra parentesi si può pensare come la differenza di due quadrati e se ne deduce che il polinomio ha due zeri reali, precisamente

$$\frac{-b}{2a}-\frac{\sqrt{b^2-4ac}}{2a} \quad \text{e} \quad \frac{-b}{2a}+\frac{\sqrt{b^2-4ac}}{2a},$$

che si scrivono, cumulativamente, nella ben nota forma

$$\frac{-b\pm\sqrt{b^2-4ac}}{2a}. \tag{3.12}$$

La quantità b^2-4ac si indica con "Δ" (leggi: *delta*) e si chiama *discriminante*, in quanto, come visto, "discrimina" con il suo segno il numero di zeri reali del polinomio ax^2+bx+c.

Nel caso in cui Δ sia zero si usa dire che il trinomio ha "due zeri reali coincidenti". Detti in ogni caso x_1 e x_2 i due zeri (eventualmente coincidenti) che il polinomio ha quando $\Delta\geq 0$, si ha la seguente scomposizione per il trinomio:

$$ax^2+bx+c=a(x-x_1)(x-x_2), \tag{3.13}$$

mentre se $\Delta<0$ il trinomio non è scomponibile (nei reali). Questi risultati sono un caso particolare del teorema fondamentale dell'algebra (vedi il teorema 3.13 nella pagina 92).

3.6. La regola di Ruffini

La divisione tra un polinomio $P(x)$ e un binomio monico del tipo $x-a$, dove a è un numero reale, si può fare, oltreché con la usuale regola della divisione tra polinomi, anche con una speciale tecnica, detta *regola di Ruffini* che può far risparmiare un po' di tempo. Segnaliamo che non è una necessità assoluta memorizzare questa tecnica, in quanto la tradizionale divisione funziona sempre, tuttavia la riportiamo per la sua notorietà. In particolare ci preme segnalare che il nome di Paolo Ruffini è più legato al teorema del resto, che non a questa tecnica per eseguire divisioni.

Ragioniamo su un esempio: dividere il polinomio $3x^5-5x^2+6$ per il binomio $x-1/3$. Si costruisce uno schema a tre righe e con le barre verticali e orizzontali come più sotto indicato; nella prima riga si riportano, nell'ordine delle potenze decrescenti, i coefficienti del dividendo, scrivendo 0 al posto dei coefficienti mancanti.

$$\begin{array}{c|ccccc|c} & 3 & 0 & 0 & -5 & 0 & +6 \\ & & & & & & \\ \hline & & & & & & \end{array}$$

Nella seconda riga si scrive, all'estrema sinistra, il termine noto del divisore cambiato di segno, mentre nella terza riga si "abbassa" il primo coefficiente del dividendo.

$$\begin{array}{c|ccccc|c} & 3 & 0 & 0 & -5 & 0 & +6 \\ \frac{1}{3} & & & & & & \\ \hline & 3 & & & & & \end{array}$$

Successivamente si moltiplica 3 per $1/3$, si scrive il risultato sotto al secondo coefficiente del dividendo e si somma (nella terza riga) questo risultato con il secondo coefficiente del dividendo.

$$\begin{array}{c|ccccc|c} & 3 & 0 & 0 & -5 & 0 & +6 \\ \frac{1}{3} & & 1 & & & & \\ \hline & 3 & 1 & & & & \end{array}$$

Si ripete il procedimento precedente fino al termine della prima riga.

$$\begin{array}{c|ccccc|c} & 3 & 0 & 0 & -5 & 0 & +6 \\ \frac{1}{3} & & 1 & \frac{1}{3} & \frac{1}{9} & -\frac{44}{27} & -\frac{44}{81} \\ \hline & 3 & 1 & \frac{1}{3} & -\frac{44}{9} & -\frac{44}{27} & \frac{442}{81} \end{array}$$

I numeri scritti nella terza riga, tra le due barre verticali, sono i coefficienti del polinomio quoziente, che ha esattamente un grado di meno rispetto al dividendo, ordinati secondo le potenze decrescenti dell'indeterminata; invece il numero scritto nella terza riga, dopo la barra verticale, è il resto (che deve avere grado zero, in quanto il divisore è di primo grado). Dunque

$$3x^5-5x^2+6=\left(3x^4+x^3+\frac{1}{3}x^2-\frac{44}{9}x-\frac{44}{27}\right)\left(x-\frac{1}{3}\right)+\frac{442}{81}.$$

3.7. Il teorema fondamentale dell'algebra

Il teorema fondamentale dell'algebra è uno dei teoremi più importanti riguardanti i polinomi; esso ha una formulazione particolarmente semplice ed elegante nell'insieme $\mathbb{C}$ dei numeri complessi. Poiché però noi qui ci limitiamo ai numeri reali, ne diamo la formulazione opportunamente adattata.

Teorema 3.13 (Teorema fondamentale dell'algebra). *Se $P(x)$ è un polinomio di grado n e α_1, α_2, …, α_r sono i suoi, eventuali, zeri reali, con molteplicità m_1, m_2, …, m_r, allora il polinomio si può sempre decomporre nel seguente modo*

$$P(x)=a_n(x-\alpha_1)^{m_1}(x-\alpha_2)^{m_2}\cdots(x-\alpha_r)^{m_r}(x^2+p_1x+q_1)^{s_1}\cdots(x^2+p_tx+q_t)^{s_t}\,, \tag{3.14}$$

con opportuni valori dei numeri reali p_1, q_1, …, p_t, q_t e dei naturali positivi s_1, …, s_t. Inoltre i trinomi di secondo grado che compaiono nella (3.14) hanno tutti discriminante negativo, cioè non sono ulteriormente scomponibili nei reali. Si ha anche, naturalmente,

$$n=m_1+m_2+\cdots+m_r+2s_1+\cdots+2s_t\,.$$

Si tenga ben presente che, se è ben vero che ogni polinomio in $\mathbb{R}$ si può fattorizzare nel modo detto, non è affatto detto che tale fattorizzazione possa essere effettivamente determinata, come si capisce se si tiene conto del teorema di Ruffini-Abel (vedi la nota nella pagina 89)[5].

Esempio 3.26. Si debba fattorizzare il polinomio $P(x)=x^7-2x^5+2x^4-2x^3+2x^2-3x+2$. Cerchiamo intanto eventuali zeri razionali, anzi interi visto che il polinomio è monico. La ricerca è abbastanza semplice, visto che i divisori del termine noto sono solamente ±1 e ±2. Si trova facilmente che 1 e -2 sono zeri. Procediamo ad eseguire la divisione di $P(x)$ prima per $x-1$ e poi per $x+2$. Si può costruire un unico "schema di Ruffini" come segue, dove la terza riga del primo schema diventa prima riga del secondo, senza dover riscrivere i coefficienti.

	1	0	−2	2	−2	2	−3	2
1		1	1	−1	1	−1	1	−2
	1	1	−1	1	−1	1	−2	0
−2		−2	2	−2	2	−2	2	
	1	−1	1	−1	1	−1	0	

Dunque

$$x^7-2x^5+2x^4-2x^3+2x^2-3x+2=(x-1)(x+2)(x^5-x^4+x^3-x^2+x-1)\,.$$

Il quoziente finale ottenuto ammette ancora 1 come zero. Eseguendo ancora una divisione con Ruffini si trova

$$x^7-2x^5+2x^4-2x^3+2x^2-3x+2=(x-1)^2(x+2)(x^4+x^2+1)\,.$$

La scomposizione di $x^5-x^4+x^3-x^2+x-1$ si poteva anche fare per raccoglimenti successivi:

$$x^5-x^4+x^3-x^2+x-1=x^4(x-1)+x^2(x-1)+(x-1)=(x-1)(x^4+x^2+1)\,,$$

[5] Un polinomio $P(x)$ che possa essere scomposto nel prodotto di due polinomi di grado maggiore o uguale a 1 si dice *riducibile*, altrimenti si dice *irriducibile*. Il fatto che un polinomio sia riducibile oppure no dipende dall'insieme numerico su cui si sta operando. Per esempio x^2-2 è irriducibile se operiamo sui razionali, mentre è scomponibile in $\left(x+\sqrt{2}\right)\left(x-\sqrt{2}\right)$ nei reali. Il teorema fondamentale dell'algebra implica che nei reali gli unici polinomi irriducibili sono quelli di grado 1 e quelli di grado 2 con discriminante negativo. Studiando i numeri complessi si troverà che gli unici polinomi irriducibili sono quelli di grado 1. Nella teoria dei polinomi, i polinomi irriducibili giocano un ruolo molto importante, per certi versi simile a quello dei numeri primi nei naturali.

ottenendo naturalmente lo stesso risultato. Si può osservare subito che l'ultimo quoziente, x^4+x^2+1 non può avere zeri reali (essendo la somma di due numeri maggiori o uguali a zero con il numero 1, sarà sempre maggiore o uguale a 1). La scomposizione ottenuta non è quella prevista, in quanto contiene un polinomio di quarto grado. Si ha però:

$$x^4+x^2+1=x^4+x^2+1+x^2-x^2=x^4+2x^2+1-x^2=\left(x^2+1\right)^2-x^2=(x^2+1-x)(x^2+1+x),$$

e i due trinomi di secondo grado hanno discriminante negativo, a conferma del fatto che questo polinomio non ha zeri reali. Dunque

$$x^7-2x^5+2x^4-2x^3+2x^2-3x+2=(x-1)^2(x+2)(x^2-x+1)(x^2+x+1).$$

Osservazione 3.14. Il teorema fondamentale dell'algebra ha come conseguenza importante il fatto che un polinomio di grado dispari deve avere almeno una radice reale, in quanto nella sua decomposizione non possono comparire solo trinomi di secondo grado con discriminante negativo.

3.8. Frazioni algebriche

Si chiamano *frazioni algebriche* le espressioni che si presentano come rapporti di polinomi, o eventualmente di monomi. Poiché si tratta di frazioni e poiché le lettere che compaiono nei polinomi rappresentano numeri reali, dovremo preoccuparci di controllare che i denominatori non assumano il valore 0. Per esempio, scrivendo $a-b$ le lettere a e b possono assumere qualunque valore, mentre se scriviamo $1/a-b$ dovremo escludere la possibilità che a sia uguale a b: si deve prestare molta attenzione a questo fatto per evitare grossolani errori. Proponiamo subito un esempio per chiarire la situazione.

Esempio 3.27. Sia data la frazione algebrica

$$\frac{a^2-1}{a+1}.$$

Si ha, facilmente,

$$\frac{a^2-1}{a+1}=\frac{(a-1)(a+1)}{a+1}.$$

Se ora procedessimo con una semplificazione, come faremmo con i numeri interi, del termine $a+1$, comune al numeratore e denominatore, otterremmo $a-1$. Il problema è che nella frazione iniziale il valore $a=-1$ non era ammesso, in quanto porta alla frazione[6] $0/0$, mentre nel polinomio finale $a+1$ il valore $a=-1$ è perfettamente lecito e porta al valore 0. La scrittura completa e corretta per questi passaggi è la seguente.

$$\frac{a^2-1}{a+1}=\frac{(a-1)\cancel{(a+1)}}{\cancel{a+1}}=a+1, \quad \text{se} \quad a\neq -1.$$

Si può operare con le frazioni algebriche esattamente come si opera con le frazioni numeriche, facendo addizioni, sottrazioni, moltiplicazioni, divisioni, e non sono necessarie tecniche nuove: occorrerà saper fare il prodotto e la somma di polinomi, saper trovare un multiplo comune a due o più polinomi,

[6] Si ricordi che una frazione con denominatore 0 *non* ha *mai* senso, nemmeno se il numeratore è 0: si veda la nota nella pagina 40.

eventualmente il minimo comune multiplo, ecc. Le difficoltà sono di natura tecnica: mentre di solito è facile per esempio "ridurre più frazioni numeriche al minimo comun denominatore" o "ridurre una frazione numerica ai minimi termini", tutto questo è molto più complesso nel caso di polinomi.

Tenendo conto di quanto sopra osservato segnaliamo che, prima di operare con le frazioni algebriche, occorre tenere conto di tutte le condizioni che devono essere verificate affinché le frazioni stesse abbiano senso: si tratta delle cosiddette *condizioni iniziali*.

A parte questa osservazione, non essendo necessaria la formulazione di nuove regole per operare con le frazioni algebriche, proponiamo solo alcuni esempi di semplificazione di espressioni.

Esempio 3.28.

$$\frac{a^2-2a+1}{b^3}\cdot\frac{ab^2+b^2}{a-1}\cdot\left(-\frac{b}{1-a}\right)=\frac{\cancel{(a-1)^2}}{\cancel{b^3}}\cdot\frac{(a+1)\cancel{b^2}}{\cancel{a-1}}\cdot\frac{\cancel{b}}{\cancel{a-1}}=a+1\,.$$

I calcoli hanno senso se $a\neq 1\wedge b\neq 0$.

Esempio 3.29.

$$\left(\frac{x+b}{2x-2b}-\frac{x-b}{2x+2b}-\frac{2b^2}{b^2-x^2}\right)\cdot\frac{x-b}{2b}=\left(\frac{x+b}{2(x-b)}-\frac{x-b}{2(x+b)}+\frac{2b^2}{(x-b)(x+b)}\right)\cdot\frac{x-b}{2b}=$$

$$=\frac{(x+b)^2-(x-b)^2+4b^2}{2\cancel{(x-b)}(x+b)}\cdot\frac{\cancel{x-b}}{2b}=\frac{\cancel{x^2}+2bx+\cancel{b^2}-\cancel{x^2}+2bx-\cancel{b^2}+4b^2}{4b(x+b)}=$$

$$=\frac{4bx+4b^2}{4b(x+b)}=\frac{\cancel{4b(x+b)}}{\cancel{4b(x+b)}}=1.$$

I calcoli hanno senso se $x\neq\pm b\wedge b\neq 0$.

Esempio 3.30.

$$\frac{\dfrac{1}{1+a}+\dfrac{a}{1-a}}{\dfrac{1}{1-a}-\dfrac{a}{1+a}}:\left(1-\frac{1-\dfrac{1}{a}}{1+\dfrac{1}{a}}\right)=\frac{\dfrac{1-a+a+a^2}{\cancel{1-a^2}}}{\dfrac{1+a-a+a^2}{\cancel{1-a^2}}}:\left(1-\frac{\dfrac{a-1}{\cancel{a}}}{\dfrac{a+1}{\cancel{a}}}\right)=$$

$$=\frac{\cancel{1+a^2}}{\cancel{1+a^2}}:\left(1-\frac{a-1}{a+1}\right)=1:\frac{a+1-a+1}{a+1}=1:\frac{2}{a+1}=\frac{a+1}{2}\,.$$

I calcoli hanno senso se $a\neq 1\wedge a\neq -1\wedge a\neq 0$.

3.9. Progressioni

3.9.1. Progressioni aritmetiche

Definizione 3.15 (Progressione aritmetica)**.** *Si dice che $n+1$ numeri reali a_0, a_1, a_2, …, a_n sono in* progressione aritmetica *se è costante la differenza tra a_k e a_{k-1}, per $1\leq k\leq n$. La differenza a_k-a_{k-1} si indica con d e si chiama* ragione *della progressione. I numeri dati si dicono anche* termini *della progressione.*

Si ha, facilmente,

$$a_1 = a_0 + d,\ a_2 = a_1 + d = a_0 + 2d,\ a_3 = a_2 + d = a_0 + 3d,\ \dots,\ a_k = a_0 + kd,\ \dots,\ a_n = a_0 + nd\,.$$

Osserviamo che a_k e a_{n-k} sono, per $1 \le k \le n$, termini equidistanti dagli estremi e che si ha:

$$a_k + a_{n-k} = (a_0 + kd) + \big(a_0 + (n-k)d\big) = a_0 + kd + a_0 + nd - kd = a_0 + a_0 + nd = a_0 + a_n\,,$$

ovvero che la somma di due termini equidistanti dagli estremi è costante.

Posto $S_n = a_0 + a_1 + \dots + a_n$, si ha

$$\begin{aligned}2S_n &= (a_0 + a_1 + \dots + a_n) + (a_0 + a_1 + \dots + a_n) = (a_0 + a_1 + \dots + a_n) + (a_n + \dots + a_1 + a_0) = \\ &= \underbrace{(a_0 + a_n) + (a_1 + a_{n-1}) + \dots + (a_n + a_0)}_{n+1 \text{ addendi uguali}} = (n+1)(a_0 + a_n)\,.\end{aligned}$$

Da qui

$$S_n = (n+1)\frac{a_0 + a_n}{2}\,, \tag{3.15}$$

ovvero: *la somma dei termini di una progressione aritmetica è data dal numero dei termini moltiplicato per la semisomma tra il primo e l'ultimo termine*.

3.9.2. Progressioni geometriche

Definizione 3.16 (Progressione geometrica)**.** *Si dice che $n+1$ numeri reali non nulli a_0, a_1, a_2, …, a_n sono in* progressione geometrica *se è costante il rapporto tra a_k e a_{k-1}, per $1 \le k \le n$. Il rapporto a_k/a_{k-1} si indica con q e si chiama* ragione *della progressione. I numeri dati si dicono anche* termini *della progressione.*

Si ha, facilmente,

$$a_1 = a_0 \cdot q,\ a_2 = a_1 \cdot q = a_0 \cdot q^2,\ a_3 = a_2 \cdot q = a_0 \cdot q^3,\ \dots,\ a_k = a_0 \cdot q^k,\ \dots,\ a_n = a_0 \cdot q^n\,.$$

Osserviamo che a_k e a_{n-k} sono, per $1 \le k \le n$, termini equidistanti dagli estremi e che si ha:

$$a_k \cdot a_{n-k} = (a_0 \cdot q^k) \cdot \big(a_0 \cdot q^{n-k}\big) = a_0 \cdot a_0 \cdot q^k \cdot q^{n-k} = a_0 \cdot a_0 \cdot q^n = a_0 \cdot a_n\,,$$

ovvero che il prodotto di due termini equidistanti dagli estremi è costante.

Posto $P_n = a_0 \cdot a_1 \dots a_n$, si ha

$$\begin{aligned}P_n^2 &= (a_0 \cdot a_1 \dots a_n) + (a_0 \cdot a_1 \dots a_n) = (a_0 \cdot a_1 \dots a_n) + (a_n \dots a_1 \cdot a_0) = \\ &= \underbrace{(a_0 \cdot a_n) + (a_1 \cdot a_{n-1}) \dots (a_n \cdot a_0)}_{n+1 \text{ fattori uguali}} = (a_0 \cdot a_n)^{n+1}\,.\end{aligned}$$

Da qui

$$P_n = \sqrt{(a_0 \cdot a_n)^{n+1}} = (a_0 \cdot a_n)^{(n+1)/2}\,, \tag{3.16}$$

ovvero: *il prodotto dei termini di una progressione geometrica è dato dal prodotto tra il primo e l'ultimo termine elevato alla semisomma del numero dei termini.*

Posto poi $S_n = a_0 + a_1 + \cdots + a_n$, si ha

$$S_n = a_0 + a_1 + \cdots + a_n = a_0 + a_0 q + a_1 q^2 + \cdots + a_0 q^n = a_0(1 + q + q^2 + \cdots + q^n).$$

Osserviamo ora che, applicando la formula (3.6), si ottiene

$$q^{n+1} - 1 = (q-1)(q^n + q^{n-1} + \cdots + 1),$$

da cui, se $q \neq 1$,

$$q^n + q^{n-1} + \cdots + 1 = \frac{q^{n+1}-1}{q-1}.$$

Ne viene

$$S_n = a_0 \frac{q^{n+1}-1}{q-1}, \quad q \neq 1. \tag{3.17}$$

Se invece $q = 1$, tutti i termini della progressione geometrica coincidono con a_0 e la loro somma è semplicemente $(n+1)a_0$.

3.10. Esercizi

Anche nella risoluzione degli esercizi di algebra elementare occorre avere una buona manualità. Qui proponiamo solo alcuni spunti, che evidenzino le principali conoscenze richieste, invitando nuovamente i lettori a riprendere in mano i testi della scuola secondaria per ulteriori problemi, tralasciando comunque gli esercizi troppo lunghi o complessi che riteniamo più dannosi che altro.

Esercizio 3.1. *Semplificare la frazione algebrica*

$$\frac{x^3 - 5x^2 + 8x - 4}{x^3 - 4x^2 + 5x - 2}.$$

Risoluzione. Applicando il teorema degli zeri razionali si trova che 1 è uno zero sia per il numeratore che per il denominatore. eseguendo la divisione (per esempio con la regola di Ruffini) sia del numeratore che del denominatore per $x - 1$ si ottiene

$$\frac{x^3 - 5x^2 + 8x - 4}{x^3 - 4x^2 + 5x - 2} = \frac{(x^2 - 4x + 4)\cancel{(x-1)}}{x^2 - 3x + 2)\cancel{(x-1)}} = \frac{x^2 - 4x + 4}{x^2 - 3x + 2}.$$

A questo punto è possibile trovare gli zeri di numeratore e denominatore applicando tecniche standard (quadrato di un binomio e zeri di un trinomio di secondo grado). Si ottiene

$$\frac{x^2 - 4x + 4}{x^2 - 3x + 2} = \frac{(x-2)^{\cancel{2}}}{(x-1)\cancel{(x-2)}} = \frac{x-2}{x-1}.$$

Condizioni iniziali $x \neq 1$ (compresa anche nella scrittura finale) e $x \neq 2$. □

Esercizio 3.2. *Scomporre in fattori il polinomio seguente*

$$a^3 - b^3 - 9a + 9b + a^2b - ab^2 .$$

Risoluzione. Si ha

$$\begin{aligned} a^3 - b^3 - 9a + 9b + a^2b - ab^2 &= (a-b)(a^2+ab+b^2) - 9(a-b) + ab(a-b) = \\ &= (a-b)(a^2+ab+b^2-9+ab) = (a-b)(a^2+2ab+b^2-9) = \\ &= (a-b)\big((a+b)^2-9\big) = (a-b)(a+b-3)(a+b+3). \quad \square \end{aligned}$$

Esercizio 3.3. *Scomporre in fattori il polinomio seguente*

$$x^2 - 3xy + 2y^2 - x + 2y .$$

Risoluzione. Si ha

$$\begin{aligned} x^2 - 3xy + 2y^2 - x + 2y &= x^2 - 4xy + 4y^2 + xy - 2y^2 - x + 2y = \\ &= (x-2y)^2 + y(x-2y) - (x-2y) = (x-2y)(x-2y+y-1) = (x-2y)(x-y-1). \quad \square \end{aligned}$$

Esercizio 3.4. *Scomporre il polinomio* $x^{100}+1$ *nel prodotto di due polinomi dello stesso grado.*

Risoluzione. Si ha

$$\begin{aligned} x^{100}+1 &= x^{100}+2x^{50}+1-2x^{50} = \left(x^{50}+1\right)^2 - \left(\sqrt{2}x^{25}\right)^2 = \\ &= \left(x^{50}+1-\sqrt{2}x^{25}\right)\left(x^{50}+1+\sqrt{2}x^{25}\right). \quad \square \end{aligned}$$

Esercizio 3.5. *Trovare il massimo comun divisore dei seguenti due polinomi, usando il metodo delle divisioni successive, come nel caso degli interi.*

$$A(x) = x^5 + x^4 - 2x^3 + x^2 + 3x - 2, \quad B(x) = x^4 + 3x^3 - 2x^2 - 5x + 3 .$$

Risoluzione. Si ha, successivamente,

- $x^5 + x^4 - 2x^3 + x^2 + 3x - 2 = (x-2)\left(x^4+3x^3-2x^2-5x+3\right) + \left(6x^3+2x^2-10x+4\right)$;
- $x^4 + 3x^3 - 2x^2 - 5x + 3 = \left(\dfrac{x}{6}+\dfrac{4}{9}\right)\left(6x^3+2x^2-10x+4\right) + \left(-\dfrac{11x^2}{9} - \dfrac{11x}{9} + \dfrac{11}{9}\right)$;
- $\left(6x^3+2x^2-10x+4\right) = \left(-\dfrac{54x}{11}+\dfrac{36}{11}\right)\left(-\dfrac{11x^2}{9} - \dfrac{11x}{9} + \dfrac{11}{9}\right)$.

Poiché l'ultimo resto è zero, si considera il penultimo resto:

$$-\frac{11x^2}{9} - \frac{11x}{9} + \frac{11}{9} = -\frac{11}{9}\left(x^2+x-1\right).$$

Tenendo conto che abbiamo convenuto di prendere sempre il massimo comun divisore con primo coefficiente uguale a 1, otteniamo

$$\mathrm{MCD}\big(x^5+x^4-2x^3+x^2+3x-2,\,,x^4+3x^3-2x^2-5x+3\big)=x^2+x-1\,.$$

Per controllo eseguiamo le divisioni di $A(x)$ e $B(x)$ con x^2+x-1. Si trova

$$\begin{aligned} x^5+x^4-2x^3+x^2+3x-2&=(x^3-x+2)(x^2+x-1)\,,\\ x^4+3x^3-2x^2-5x+3&=(x^2+2x-3)(x^2+x-1)\,. \end{aligned}$$

Si noti che le scomposizioni trovate consentono di trovare due zeri del polinomio $A(x)$, e tutti i quattro zeri del polinomio $B(x)$: nessuno di questi zeri è razionale. Il fattore x^3-x+2 del polinomio $A(x)$ deve avere almeno uno zero, in quanto di grado dispari, ma nemmeno questo zero è razionale. Dunque la ricerca del massimo comun divisore di questi due polinomi non poteva essere fatta per scomposizione con le regole da noi studiate. □

Esercizio 3.6. *Tenendo conto della scomposizione di x^5+1, riconoscere che il polinomio*

$$x^4-x^3+x^2-x+1$$

non può avere zeri reali.

Risoluzione. Si ha

$$x^5+1=(x+1)(x^4-x^3+x^2-x+1)\,.$$

Il polinomio x^5+1 può avere come zero reale solo il numero -1, in quanto nessun altro reale elevato alla quinta potenza può dare come risultato -1. Dunque anche il prodotto $(x+1)(x^4-x^3+x^2-x+1)$ deve avere come unico zero reale -1. Poiché $x-1$ ha proprio -1 come zero, mentre -1 non è zero di $x^4-x^3+x^2-x+1$, si conclude che $x^4-x^3+x^2-x+1$ non può avere zeri reali. □

Esercizio 3.7. *Scomporre in fattori il seguente polinomio*

$$x^4-y^4+2x^3y-2xy^3\,.$$

Risoluzione. Si ha

$$\begin{aligned} x^4-y^4+2x^3y-2xy^3&=(x-y)(x+y)(x^2+y^2)+2xy(x-y)(x+y)=\\ &=(x-y)(x+y)(x^2+y^2+2xy)=(x-y)(x+y)^3\,. \end{aligned}$$ □

Esercizio 3.8. *Scomporre in fattori il polinomio*

$$x^2+(a-3)x-3a\,.$$

Risoluzione. Se si considera il polinomio dato come polinomio nella sola indeterminata x, esso è di secondo grado e se ne possono trovare gli zeri con la nota formula

$$x_{1,2}=\frac{-(a-3)\pm\sqrt{a^2-6a+9+12a}}{2}=\frac{3-a\pm(a+3)}{2}=\begin{cases}3\\-a\end{cases}\,.$$

Il polinomio dato si scompone dunque in $(x-3)(x+a)$.

Un polinomio monico di secondo grado, x^2+bx+c si può a volte anche scomporre senza determinare direttamente gli zeri con la formula (3.12). Se infatti si riescono a trovare due numeri α e β tali che $\alpha+\beta=b$ e $\alpha\beta=c$, si ha

$$x^2+bx+c=x^2+\alpha x+\beta x+\alpha\beta=x(x+\alpha)+\beta(x+\alpha)=(x+\alpha)(x+\beta).$$

Nel caso in esame i due numeri richiesti sono -3 e a, e con questa tecnica si ottiene il medesimo risultato (forse un po' più velocemente). Tuttavia riteniamo che l'applicazione della formula (3.12) per la determinazione degli eventuali zeri di un trinomio di secondo grado, e la scomposizione che ne deriva, sia da preferire, in quanto di carattere più generale e, a nostro avviso, meno soggetta ad errori. Si noti, in particolare, che i due numeri α e β eventualmente trovati non sono gli zeri del trinomio, ma i loro opposti. □

Esercizio 3.9. *Semplificare l'espressione*

$$\frac{2+x}{x+3}-\frac{3x-1}{x^2+x-6}-\frac{x}{x+3}.$$

Risoluzione. Il trinomio x^2+x-6 ha due zeri reali, -3 e 2, dunque si scompone in $(x+3)(x-2)$. Si ha allora

$$\frac{2+x}{x+3}-\frac{3x-1}{x^2+x-6}-\frac{x}{x+3}=\frac{(2+x)(x-2)-(3x-1)-x(x-2)}{(x+3)(x-2)}=\cdots=\frac{-1}{x-2}=\frac{1}{2-x}.$$

Condizioni iniziali: $x\neq-3\wedge x\neq 2$. La condizione $x\neq 2$ è comunque implicita anche nel risultato finale, ovvero "non è andata persa nella risoluzione". □

Esercizio 3.10. *Semplificare l'espressione*

$$\frac{1}{(a-b)(a-x)}+\frac{1}{(b-a)(b-x)}+\frac{1}{(x-a)(x-b)}.$$

Risoluzione. Cominciamo con l'osservare che alcuni fattori che compaiono ai denominatori differiscono solo per il segno: conviene trasformarli in modo che abbiano lo stesso segno.

$(a-x)=-(x-a)$, $(b-a)=-(a-b)$, $(b-x)=-(x-b)$, e quindi $(b-a)(b-x)=(a-b)(x-b)$.

Dunque

$$\frac{1}{(a-b)(a-x)}+\frac{1}{(b-a)(b-x)}+\frac{1}{(x-a)(x-b)}=$$

$$=\frac{-1}{(a-b)(x-a)}+\frac{1}{(a-b)(x-b)}+\frac{1}{(x-a)(x-b)}=$$

$$=\frac{-(x-b)+(x-a)+(a-b)}{(a-b)(x-a)(x-b)}=\cdots=\frac{0}{(a-b)(x-a)(x-b)}=0.$$

Condizioni iniziali: $a\neq b\wedge a\neq x\wedge b\neq x$. □

Esercizio 3.11. *Semplificare l'espressione*

$$\left(\frac{a}{b}+\frac{4b}{a}+4\right)\cdot\left(\frac{a^3-2a^2b+4ab^2}{a^3-8b^3}:\frac{a^3+8b^3}{a^3+2a^2b+4ab^2}\right)\cdot\frac{b}{a}.$$

Risoluzione. Si ha

$$\begin{aligned}
&\left(\frac{a}{b}+\frac{4b}{a}+4\right)\cdot\left(\frac{a^3-2a^2b+4ab^2}{a^3-8b^3}:\frac{a^3+8b^3}{a^3+2a^2b+4ab^2}\right)\cdot\frac{b}{a}=\\
&\quad=\frac{a^2+4b^2+4ab}{ab}\cdot\left(\frac{a\cancel{(a^2-2ab+4b^2)}}{(a-2b)\cancel{(a^2+2ab+4b^2)}}\cdot\frac{a\cancel{(a^2+2ab+4b^2)}}{(a+2b)\cancel{(a^2-2ab+4b^2)}}\right)\cdot\frac{b}{a}=\\
&\qquad=\frac{(a+2b)^{\cancel{2}}}{\cancel{a}\cancel{b}}\cdot\frac{\cancel{a^2}}{(a-2b)\cancel{(a+2b)}}\cdot\frac{\cancel{b}}{\cancel{a}}=\frac{a+2b}{a-2b}.
\end{aligned}$$

Condizioni iniziali: $a\neq 0\wedge b\neq 0\wedge a\neq\pm 2b$. □

4. Funzioni

4.1. Definizioni

Abbiamo già introdotto, vedi il paragrafo 1.9 del capitolo 1, il concetto di funzione: vogliamo trattare questo concetto un po' più in dettaglio, segnalando che ci occuperemo in particolare delle funzioni *reali di variabile reale*, dette brevemente anche *funzioni reali*, ovvero delle funzioni in cui il dominio e il codominio sono sottoinsiemi dell'insieme dei numeri reali.

Cominciamo con il richiamare, nelle loro parti essenziali, la definizione 1.17.

Definizione 4.1 (Funzione)**.** *Dati due insiemi A e B, si dice* funzione *di A in B una qualunque legge che faccia corrispondere* a ogni *elemento di A* uno ed un solo elemento *di B.*

L'insieme A è detto dominio *della funzione, l'insieme B è detto* codominio*. Se x è un elemento dell'insieme A e y è l'*unico *elemento di B che corrisponde ad A, si dice che y è* funzione *di x e si scrive $y = f(x)$. L'elemento y di B è anche detto* immagine *di x tramite f.*

La notazione più completa per le funzioni è la seguente:

$$f : A \to B,\ x \mapsto f(x), \tag{4.1}$$

ma spesso si scrive solo

$$x \mapsto f(x),$$

se gli insiemi A e B sono già stati precisati o sono chiari dal contesto. Si può anche dire semplicemente *la funzione $y = f(x)$*, anche se questo tipo di scrittura può dar luogo ad equivoci e per questo fatto può far storcere il naso ai puristi: in ogni caso occorre tenere sempre presente che in alcune circostanze il rigore è utile (anzi indispensabile), in altre si limita solo ad appesantire l'esposizione.

È opportuno mettere in evidenza, nella definizione di funzione, le seguenti due proprietà.

1. Per assegnare una funzione occorrono tre elementi: il dominio, il codominio e la legge che permette di associare ad ogni punto del dominio uno e un solo punto del codominio.
2. È *obbligatorio* che ad ogni punto dell'insieme A corrisponda esattamente un punto dell'insieme B, cioè che ogni punto dell'insieme A sia in relazione con esattamente un punto dell'insieme B, mentre non è affatto richiesto che ogni punto dell'insieme B sia in relazione con qualche punto dell'insieme A.

L'insieme dei punti di B che è in relazione con qualche punto di A, cioè che è immagine di qualche punto di A, è molto importante e si dà la seguente definizione[(1)].

[1] Purtroppo anche questa nomenclatura non è universale: alcuni autori chiamano "codominio" l'insieme delle immagini e semplicemente "insieme di arrivo" l'insieme B. Tuttavia la nomenclatura che abbiamo proposto è la più diffusa e quella che ci pare più conveniente.

Definizione 4.2 (Insieme immagine)**.** *Se $f : A \to B$ è una funzione, l'insieme degli $y \in B$ che siano immagine di almeno un $x \in A$ si chiama* insieme immagine *o semplicemente* immagine *di A tramite f e si indica con* $\mathrm{Im}(f)$ *o anche con $f(A)$. In formule:*

$$\mathrm{Im}(f) = f(A) = \{\, y \in B \mid \exists x \in A,\ y = f(x) \,\}. \tag{4.2}$$

4.2. Rappresentazioni grafiche

In sostanza per assegnare una funzione $f : A \to B$ occorre assegnare un sottoinsieme del prodotto cartesiano $A \times B$, costituito da coppie (x, y) con la caratteristica che *tutti* gli elementi di A compaiano *esattamente* una volta come primo elemento della coppia. A questo proposito si dà la seguente definizione.

Definizione 4.3 (Grafico)**.** *Si chiama* grafico *di una funzione $f : A \to B$, l'insieme*

$$\Gamma(f) = \{\, (x, y) \subset A \times B \mid y = f(x), x \in A \,\}. \tag{4.3}$$

Sarebbe addirittura possibile identificare una funzione con il suo grafico, e molti lo fanno: noi preferiamo, seguendo la tradizione, mantenere distinti i due concetti. Ci preme in ogni caso segnalare che il concetto di grafico di una funzione non deve essere confuso con quello di rappresentazione grafica, che consiste di tecniche "grafiche" atte a rendere immediatamente evidenti alcune caratteristiche delle funzioni. Torneremo ancora su questo problema.

La prima tecnica grafica che viene utilizzata per visualizzare le funzioni è quella dei *diagrammi a frecce*, come quello della figura 4.1.

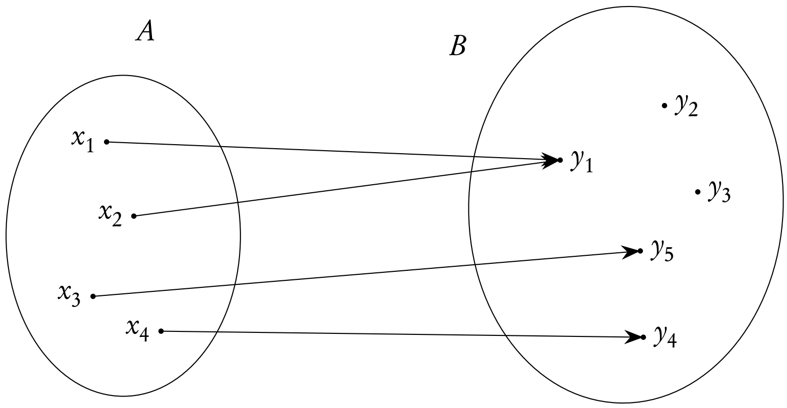


Figura 4.1.: *Diagramma "a frecce" per visualizzare una funzione (tra insiemi finiti)*

Si noti che è *obbligatorio* che da *ogni* punto (elemento) dell'insieme A parta *esattamente* una freccia, mentre sui punti dell'insieme B possono anche arrivare *più* frecce, oppure *nessuna* freccia. Si potrebbe dire, usando un linguaggio figurato, che A è l'insieme degli arcieri, B l'insieme dei bersagli e che ogni arciere ha a disposizione nella propria faretra solo una freccia che è costretto a lanciare, mentre non ci sono limitazioni sui bersagli da colpire: ci possono essere bersagli colpiti da più frecce, e anche bersagli non colpiti da alcuna freccia.

Nella rappresentazione della figura 4.1 si vede chiaramente che l'immagine della funzione è costituita dall'insieme $\{ y_1, y_3, y_5 \}$, che è un sottoinsieme proprio di B.

È chiaro che rappresentazioni grafiche come quella appena vista hanno senso solo se gli insiemi in questione sono finiti: in caso contrario si dovrebbero disegnare infinite frecce, cosa chiaramente impossibile.

Si usano anche altri tipi di rappresentazione per le funzioni. Per esempio se si considera la funzione che a ogni numero naturale compreso tra 1 e 5 fa corrispondere la sua metà (funzione che ha come dominio i numeri naturali citati e come codominio i numeri razionali), si può usare una *tabella* a doppia entrata, in cui nella prima colonna si scrivono i numeri naturali $1, 2, \dots, 5$ e nella seconda colonna le *corrispondenti* metà di questi numeri; vedi la tabella 4.1.

Tabella 4.1.: *Rappresentazione "tabulare" di una funzione*

x	$x/2$
1	$1/2$
2	1
3	$3/2$
4	2
5	$5/2$

Un altro tipo di rappresentazione è quello dei diagrammi a torta, molto significativo in casi speciali. Consideriamo, ad esempio, un corso universitario dove si sono iscritti 120 alunni, provenienti da varie provincie, come nella tabella che segue:

Gorizia	Pordenone	Treviso	Trieste	Udine
5	70	15	10	20

Si comincerà con il calcolare le percentuali relative alle varie provincie:

Gorizia	Pordenone	Treviso	Trieste	Udine
4.17	58.33	12.5	8.33	16.67

Successivamente si calcoleranno le ampiezze delle "fette di torta" da utilizzare per ciascuna provincia, tenendo conto che la torta totale ha un'apertura di 360°:

Gorizia	Pordenone	Treviso	Trieste	Udine
15°	210°	45°	30°	60°

Il grafico è a questo punto immediato ed è mostrato nella figura 4.2.

Una variante ovvia è il diagramma a semitorta, che funziona sullo stesso principio, con un semicerchio anziché con un cerchio: è usato per esempio per esempio nella rappresentazione della distribuzione dei seggi nel parlamento, visto che l'aula di riunione ha proprio la forma di un emiciclo.

Ancora un'altra possibilità è quella di un diagramma a barre, che proponiamo nella figura 4.3: qui si tratta di disegnare dei rettangoli, di base costante e altezza proporzionale alle percentuali relative a ciascuna provincia.

È evidente che, sia in questo diagramma che nel diagramma a torta, l'ordine in cui sono situati i vari elementi del dominio non ha alcun interesse.

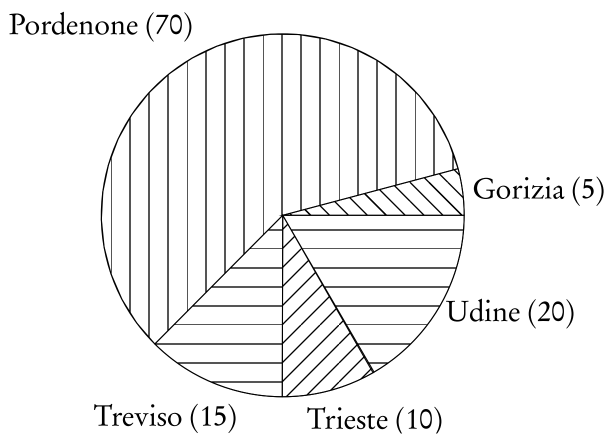


Figura 4.2.: *Provenienza degli studenti del Corso ..., ripartiti per Provincia, diagramma "a torta"*

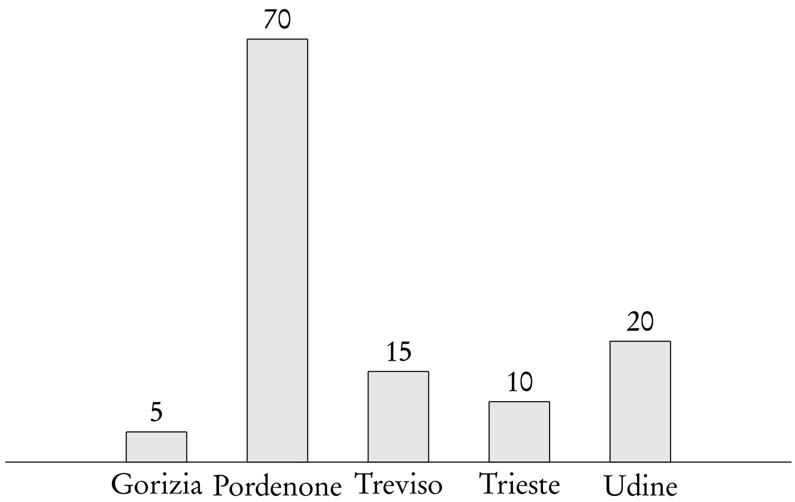


Figura 4.3.: *Provenienza degli studenti del Corso ..., ripartiti per provincia, diagramma "a barre"*

La rappresentazione più conveniente nel caso delle funzioni tra due insiemi di numeri reali è però quella dei diagrammi o grafici cartesiani, in particolare nel caso in cui gli insiemi siano infiniti, quando le rappresentazioni precedenti non sono utilizzabili. L'idea è di considerare un piano in cui si sia fissato un sistema di coordinate cartesiane (ortogonali per semplicità) $\mathrm{O}xy$ e rappresentarvi tutte le coppie (x, y) in cui x è un punto (numero) del dominio della funzione e $y = f(x)$ è il *corrispondente* valore nel codominio della funzione. Riprendendo in esame l'esempio proposto nella tabella 4.1, dobbiamo rappresentare i punti

$$\mathsf{A} = (1, 1/2),\ \mathsf{B} = (2, 1),\ \mathsf{C} = (3, 3/2),\ \mathsf{D} = (4, 2),\ \mathsf{E} = (5, 5/2),$$

ottenendo il grafico della figura 4.4.

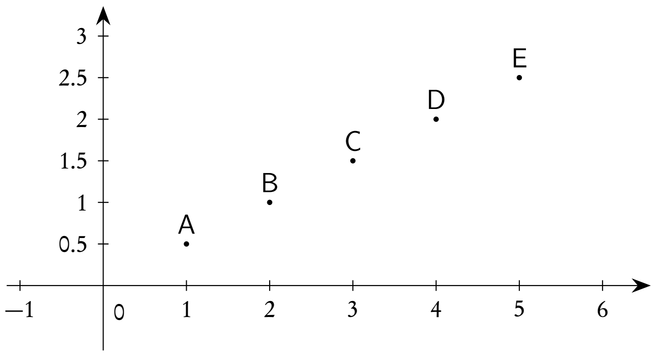


Figura 4.4.: *Esempio di grafico cartesiano*

Il grafico della figura 4.4 è in realtà un grafico a frecce "compattato": siccome i valori del dominio sono punti dell'asse x e quelli del codominio punti dell'asse y, possiamo sempre pensare di tracciare delle frecce che colleghino i punti del dominio con i corrispondenti del codominio, come quelle della figura 4.1, solo che è opportuno che le frecce "passino" per i punti A, B, …, come nella figura 4.5.

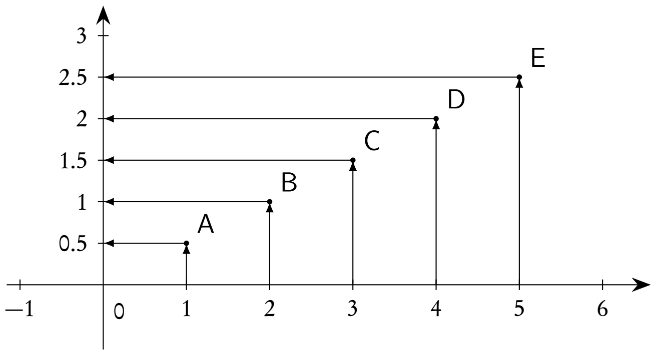


Figura 4.5.: *Esempio di grafico cartesiano, con frecce*

Il grafico 4.4 "compatta" il grafico 4.5 nel senso che ne prende solo gli elementi essenziali, cioè gli "spigoli delle frecce": è evidente che dalla conoscenza degli spigoli si possono facilmente ricostruire le frecce.

Se si confronta la figura 4.4 con la tabella 4.1, ci si rende immediatamente conto dei notevoli vantaggi che il grafico presenta: da esso si può per esempio capire, "a colpo d'occhio", che al crescere di x nel dominio la corrispondente y del codominio cresce, e che tale crescita è *costante*. La cosa diventa ancora più significativa se si vuole considerare la funzione che a ogni numero reale x faccia corrispondere la sua metà: a differenza di quanto succedeva con la funzione rappresentata nella tabella 4.1, questa volta la x non varia più in un insieme finito e quindi una rappresentazione tabulare non ha alcun senso[(2)]. Un diagramma cartesiano è decisamente più significativo.

[2]Si noti comunque che la regola (legge) che collega la x alla y è la stessa del caso precedente: il valore di $f(x)$ è la metà di quello di x. Questo rende evidente che per assegnare una funzione *non* è sufficiente assegnare la regola di calcolo, occorre anche fissare il dominio e il codominio.

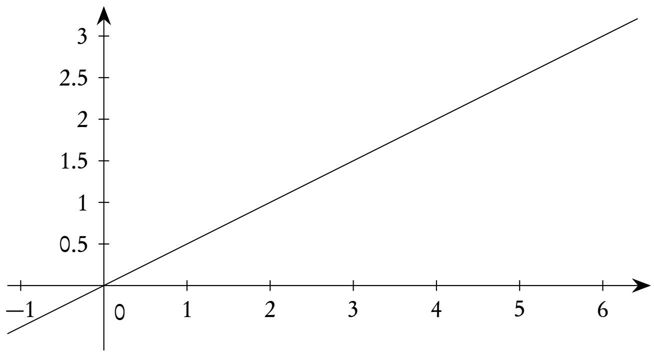


Figura 4.6.: *Grafico cartesiano della funzione* $f(x) = {}^{x}/_{2}$

Naturalmente il diagramma della figura 4.6 contiene anche i punti già rappresentati nel diagramma della figura 4.4, visto che il dominio di quella funzione è un sottoinsieme di quello della funzione ora in esame.

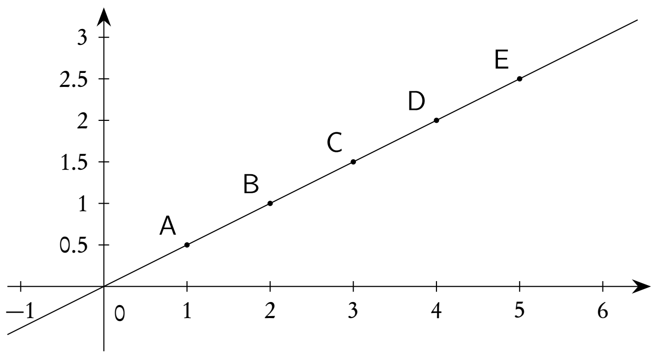


Figura 4.7.: *Grafico cartesiano della funzione* $f(x) = {}^{x}/_{2}$*, con evidenziati alcuni punti*

In tutti i grafici cartesiani che abbiamo fatto abbiamo usato la stessa unità di misura sui due assi: sistemi cartesiani siffatti sono detti *monometrici*. Di solito però nelle applicazioni la cosa non è possibile, e ne vedremo in seguito i motivi. È opportuno tenere presente che se un sistema cartesiano nel piano non è monometrico, le figure possono essere deformate. Per esempio i due grafici della figura 4.8 mostrano la circonferenza di centro l'origine e raggio 1, in due diversi sistemi di coordinate, di cui solo il primo è monometrico.

Discorso simile per la bisettrice del primo e terzo quadrante, rappresentata nella figura 4.9 negli stessi sistemi cartesiani della figura 4.8.

Nelle funzioni che avremo modo di considerare in questo corso, i grafici cartesiani saranno costituiti normalmente da "curve", nel senso intuitivo del termine. Occorre però tenere ben presente l'osservazione che segue.

Osservazione 4.4. La caratteristica delle funzioni di associare ad ogni x del dominio uno ed un solo y del codominio ha un'immediata interpretazione geometrica in termini di grafici cartesiani per le funzioni reali: una retta verticale (cioè parallela all'asse delle ordinate) condotta a partire da un punto x del

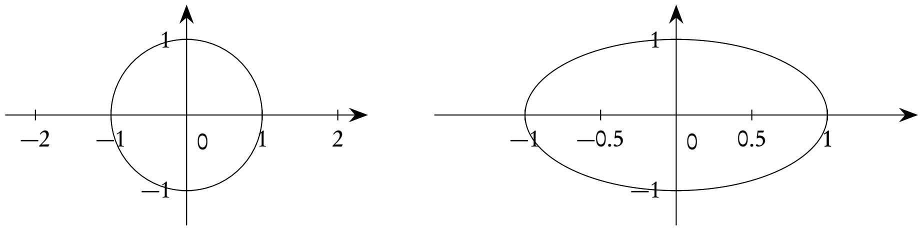


Figura 4.8.: *Circonferenza di centro l'origine e raggio* 1*, in due diversi sistemi di coordinate, il primo monometrico, il secondo no*

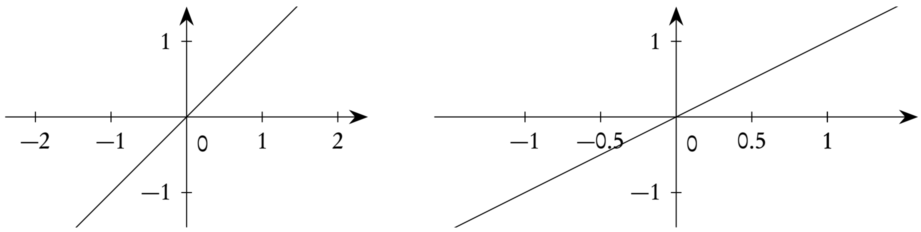


Figura 4.9.: *La bisettrice del primo e terzo quadrante, in due diversi sistemi di coordinate, il primo monometrico, il secondo no*

dominio della funzione, incontra sempre il grafico in un solo punto: se così non fosse ad uno stesso valore di x potrebbero corrispondere più valori di y.

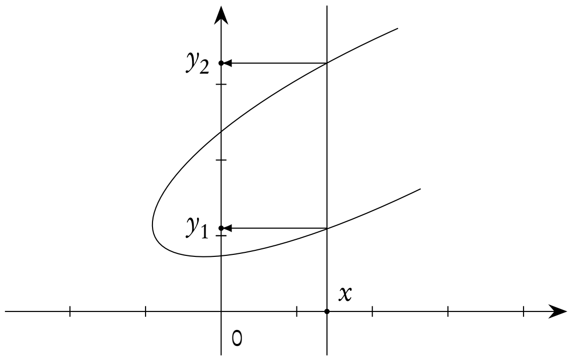


Figura 4.10.: *Curva che non è il grafico cartesiano di una funzione*

I grafici cartesiani, per le funzioni reali, sono in grado di evidenziare anche chiaramente quali sono il dominio e l'insieme immagine di una funzione: il dominio coincide con la proiezione del grafico stesso sull'asse x, l'insieme immagine con la proiezione del grafico sull'asse y. Si veda la figura 4.11.

Un modo suggestivo di pensare alle funzioni è quello di immaginare una "scatola nera", piena di rotelline ed ingranaggi, dotata di una porta di ingresso (input) e di una porta di uscita (output): se inserisco un elemento del dominio, cioè un dato (che nei casi di nostro interesse sarà un numero reale), diciamolo x, attraverso la porta di ingresso, la funzione esegue delle elaborazioni su questo dato e ci restituisce, attraverso la porta di uscita, un risultato, cioè un y, che chiamiamo "immagine" di x tramite la funzione. La figura seguente illustra la situazione relativa alla funzione $f(x)=x^3-x$.

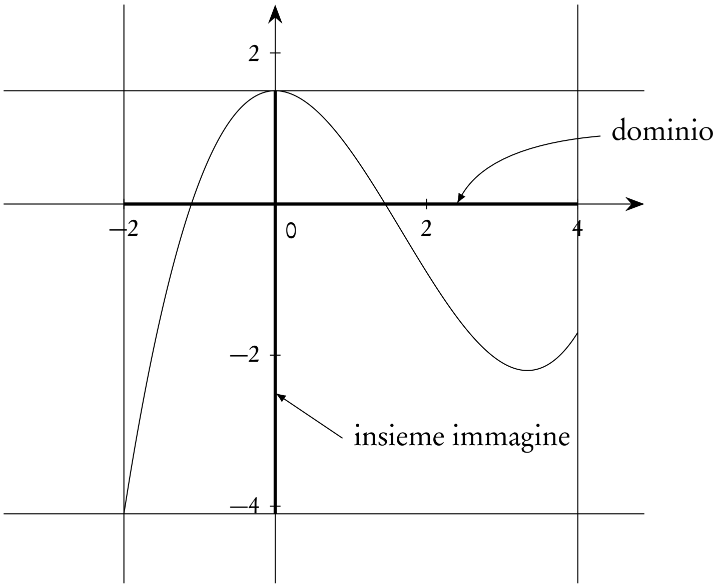


Figura 4.11.: *Dominio e immagine di una funzione reale, a partire dal grafico cartesiano*

$$\textit{Input} \longrightarrow \boxed{x^3 - x} \longrightarrow \textit{Output}$$

Se per esempio il dato in ingresso è il numero 5, gli "ingranaggi interni della scatola nera" calcoleranno il cubo di 5 e successivamente ne sottrarranno il numero 5 stesso, fornendo 120 come output.

$$5 \longrightarrow \boxed{x^3 - x} \longrightarrow 120$$

4.3. Funzioni reali e dominio naturale

Il caso delle funzioni reali (cioè di quelle funzioni in cui sia il dominio che il codominio sono sottoinsiemi dell'inseme dei numeri reali) è di grande importanza e costituisce l'argomento fondamentale di questo testo. Nei casi che ci interessano, inoltre, la legge che fornisce i valori di $f(x)$ è data da un insieme di operazioni matematiche da eseguirsi sui valori di x: addizioni, moltiplicazioni, estrazioni di radici, ecc. In questi casi si dà la seguente definizione.

Definizione 4.5 (Dominio naturale)**.** *Data una funzione reale, esprimibile mediante operazioni matematiche da eseguirsi sulla variabile x, si chiama* dominio naturale *della funzione il massimo sottoinsieme di $\mathbb{R}$ in cui le operazioni da eseguire hanno senso. Inoltre se il codominio della funzione non è precisato si conviene che esso sia sempre tutto $\mathbb{R}$.*

Tenendo conto di questa definizione, si potrà definire una funzione reale assegnando semplicemente l'insieme di operazioni da eseguire sulla variabile x.

Esempio 4.1. Dicendo: "la funzione $f(x)=x^2$", intendiamo la funzione $f:\mathbb{R}\to\mathbb{R}$, $x\mapsto x^2$.

Esempio 4.2. Dicendo: "la funzione $f(x)=\sqrt{x}$", intendiamo la funzione $f:\mathbb{R}^+\to\mathbb{R}$, $x\mapsto\sqrt{x}$.

Esempio 4.3. Dicendo: "la funzione $f(x)={}^1\!/x$", intendiamo la funzione $f:\mathbb{R}\setminus\{0\}\to\mathbb{R}$, $x\mapsto {}^1\!/x$.

A seguito di queste convenzioni avranno senso domande del tipo: "si calcoli il dominio della funzione...", dove per dominio si deve intendere il dominio naturale.

Abbiamo detto che ci occuperemo esclusivamente di funzioni in cui la legge di passaggio da x a $f(x)$ è "di natura matematica". Ciò non significa però che la legge stessa debba essere unica per tutto il dominio: potremo considerare anche funzioni *definite a tratti*, o *in modo composito*, come nell'esempio che segue.

$$f(x)=\begin{cases} x^2, & \text{se} \quad x<-2 \\ 2x+1, & \text{se} \quad -2\le x<5 \\ x+x^2, & \text{se} \quad x\ge 5 \end{cases}.$$

4.4. Funzioni iniettive, suriettive, biiettive

Tra tutte le funzioni hanno particolare interesse le funzioni che soddisfano alle caratteristiche indicate nella definizione che segue.

Definizione 4.6. *Si dicono*

1. iniettive *le funzioni in cui punti diversi del dominio hanno immagini diverse, ovvero le funzioni in cui ogni punto del codominio è immagine di* al più un *punto del dominio;*
2. suriettive *le funzioni in cui ciascun punto del codominio è immagine di* almeno un *punto del dominio*
3. biiettive *o* biunivoche *le funzioni che sono contemporaneamente iniettive e suriettive ovvero le funzioni in cui ciascun punto del codominio è immagine di* esattamente un *punto del dominio.*

Detto in termini più formali: una funzione $f:A\to B$

1. *è iniettiva se*

$$\text{dati } x_1 \text{ e } x_2\in A, \text{ con } x_1\ne x_2, \text{ si ha } f(x_1)\ne f(x_2); \tag{4.4}$$

2. *è suriettiva se*

$$\forall y\in B,\ \exists x\in A \text{ tale che } f(x)=y; \tag{4.5}$$

3. *è biiettiva se*

$$\forall y\in B,\ \exists! x\in A \text{ tale che } f(x)=y. \tag{4.6}$$

In termini di grafici cartesiani, per le funzioni reali, si possono fare le seguenti osservazioni:

1. una funzione è iniettiva se una retta orizzontale (parallela all'asse x) qualunque interseca il grafico della funzione al più una volta;
2. una funzione è suriettiva se ogni retta orizzontale (parallela all'asse x) interseca il grafico della funzione almeno una volta;

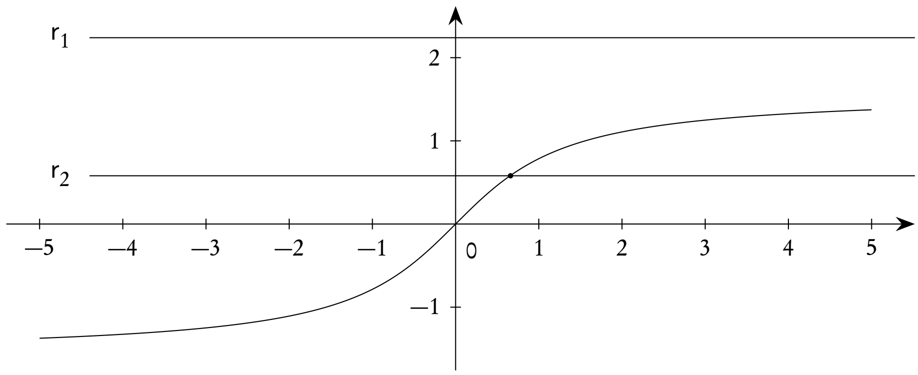


Figura 4.12.: *Esempio di funzione iniettiva*

3. una funzione è biunivoca se ogni retta orizzontale (parallela all'asse x) interseca il grafico della funzione esattamente una volta.

Nel grafico della figura 4.12, che visualizza un esempio di funzione iniettiva, la retta r_1 non interseca mai il grafico della funzione, mentre la retta r_2 lo interseca una sola volta: nessuna retta orizzontale, comunque, può intersecare il grafico più di una volta.

Nel grafico della figura 4.13, che visualizza un esempio di funzione suriettiva, la retta r_1 interseca il grafico della funzione in tre punti, mentre la retta r_2 lo interseca una sola volta: ogni retta orizzontale comunque deve intersecare il grafico almeno una volta.

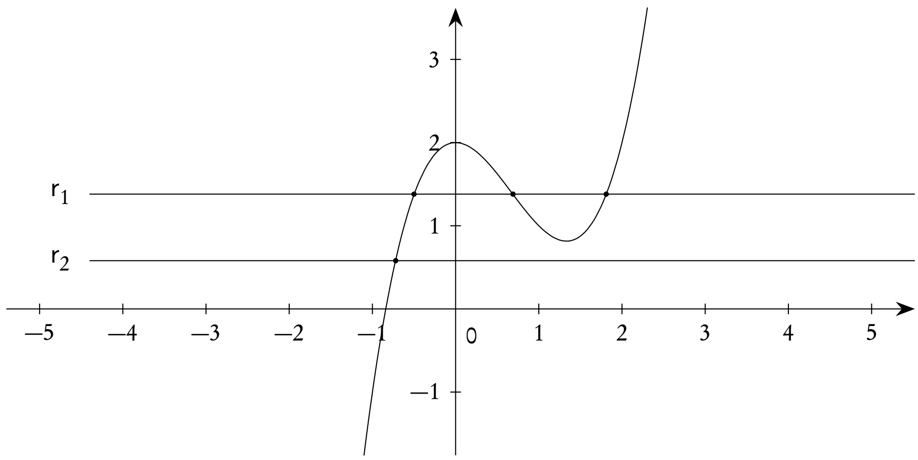


Figura 4.13.: *Esempio di funzione suriettiva*

Nel grafico della figura 4.14, che visualizza un esempio di funzione biiettiva, sia la retta r_1 che la retta r_2 intersecano il grafico una sola volta, e questo succederebbe per ogni retta orizzontale.

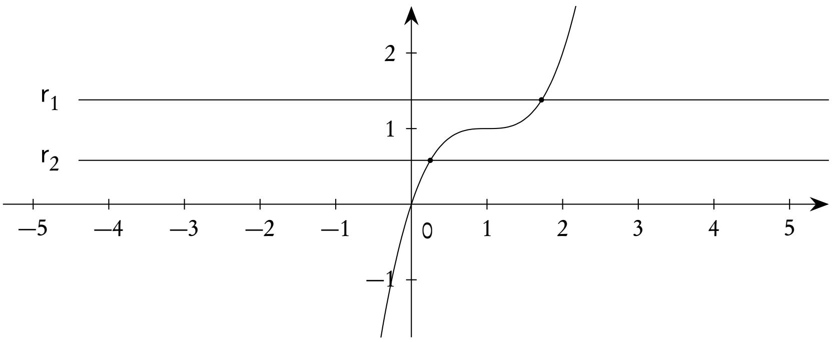


Figura 4.14.: *Esempio di funzione biiettiva*

4.5. Restrizioni di una funzione

Come già osservato, per assegnare una funzione occorre assegnare un dominio, un codominio e una legge: solo nel caso particolare delle funzioni reali, se non ci sono altre indicazioni, si assume quanto contenuto nella definizione 4.5.

In molti casi è opportuno, a partire da certe funzioni, ottenerne altre, con modifiche opportune di uno dei tre elementi che servono per la definizione, in particolare con modifiche sul dominio e sul codominio: parleremo in questi casi di *restrizioni di una funzione*.

Definizione 4.7 (Restrizione sul dominio)**.** *Data una funzione $f: A \to B$, $x \mapsto f(x)$, se $C \subset A$ è un sottoinsieme di A, possiamo considerare una nuova funzione $g: C \to B$, ponendo $g(x) = f(x)\ \forall x \in C$: in altre parole una nuova funzione in cui il codominio e la legge rimangono invariati rispetto a quelli di f, solo con un dominio "ristretto": parleremo appunto di* restrizione sul dominio *e useremo la scrittura*

$$g = f_{|C}\,, \tag{4.7}$$

che si legge: "la funzione g è la funzione f ristretta a C".

Esempio 4.4. Sia data la funzione $f: \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. La funzione $g: \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto x^2$ è una restrizione della funzione f ai reali maggiori o uguali a zero. Il grafico cartesiano di g si ottiene prendendo una parte del grafico cartesiano di f, come mostrato nella figura 4.15. Per la funzione f, seguendo le convenzioni adottate per le funzioni reali, potremo scrivere semplicemente $f(x) = x^2$, per g dovremo invece precisarne esplicitamente il dominio (se sottintendiamo che il codominio sia tutto $\mathbb{R}$).

Consideriamo ora una funzione $f: A \to B$ non suriettiva. Se eliminiamo dal codominio alcuni, o anche tutti, i punti che non sono immagine di alcun x del dominio, le caratteristiche essenziali della funzione non cambiano; non potremo però eliminare alcun punto che sia immagine di qualche x del dominio, a meno di altri cambiamenti. Di solito in questi casi si eliminano dal codominio tutti i punti che sono fuori dall'insieme immagine. Si dà in proposito la seguente definizione.

Una funzione ottenuta da un'altra per restrizione all'immagine è sempre suriettiva: tuttavia questa operazione che, in un certo senso, "semplifica" una funzione non è sempre conveniente e quando la si vuole fare bisogna esplicitamente indicarlo.

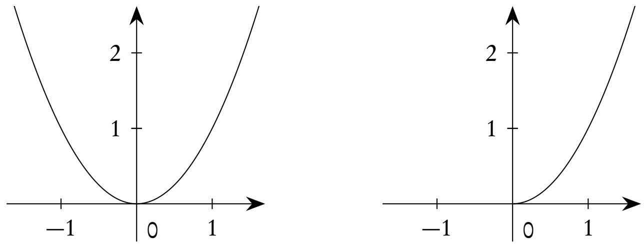


Figura 4.15.: *Grafici di $f(x) = x^2$ e di $f_{|\mathbb{R}^+}$*

Definizione 4.8 (Restrizione all'immagine)**.** *Data una funzione $f : A \to B$, $x \mapsto f(x)$, la funzione da A in $f(A)$, definita dalla stessa legge che definisce f, si chiama* restrizione *della funzione f* all'immagine *e si indica ancora con lo stesso simbolo di f.*

4.6. Operazioni tra funzioni

Nell'insieme delle funzioni reali è possibile definire le operazioni di addizione, sottrazione, moltiplicazione e divisione, che forniscono come risultato ancora funzioni reali.

Siano f e g due funzioni reali e sia A l'intersezione dei rispettivi domini.

1. Diciamo *somma* delle due funzioni f e g la funzione reale $f + g$ definita in A da
$$(f + g)(x) = f(x) + g(x).$$
2. Diciamo *differenza* delle due funzioni f e g la funzione reale $f - g$ definita in A da
$$(f - g)(x) = f(x) - g(x).$$
3. Diciamo *prodotto* delle due funzioni f e g la funzione reale fg definita in A da
$$(fg)(x) = f(x) \cdot g(x).$$
4. Se consideriamo poi l'insieme A_1 degli $x \in A$ dove $g(x) \neq 0$, diciamo *quoziente* di f e g la funzione f/g definita in A_1 da
$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}.$$

In sostanza le operazioni "algebriche" tra funzioni si riconducono a operazioni sulle immagini delle funzioni stesse: poiché queste immagini sono numeri reali, valgono tutte le usuali proprietà di queste operazioni.

Un ulteriore operazione di grande importanza tra funzioni è l'operazione di *composizione*. La definizione formale è la seguente.

Definizione 4.9 (Funzione composta)**.** *Siano $g : A \to B$ e $f : C \to D$, due funzioni reali. Se $g(A) \subseteq C$, si definisce la* funzione composta *di g ed f, nell'ordine, la funzione $f \circ g$ (leggi "f composto g") di A in D, definita da*

$$(f \circ g)(x) = f\big(g(x)\big). \tag{4.8}$$

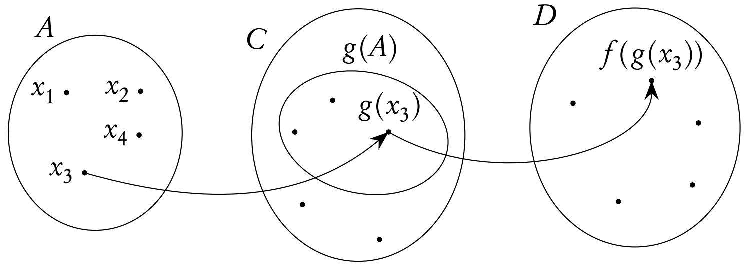
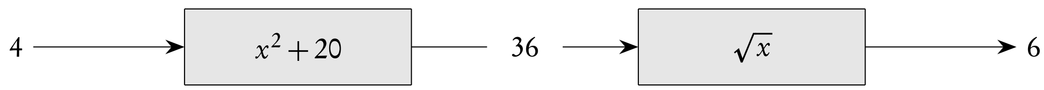


Figura 4.16.: *Composizione di funzioni, diagramma a frecce*

In termini di diagrammi a frecce, la situazione è schematizzata nella figura 4.16.

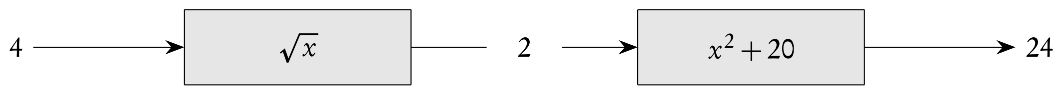
Nella suggestiva rappresentazione mediante "scatole nere" possiamo pensare di avere due scatole in serie: l'output della prima scatola non viene utilizzato subito, ma inserito direttamente come input nella seconda, la quale invece fornisce l'output finale. Se per esempio le due funzioni in questione sono $g(x) = x^2 + 20$ e $f(x) = \sqrt{x}$ e se forniamo come input per g il numero 4, l'output ottenuto, 36, viene inserito nella seconda scatola ottenendo come risultato finale 6.



Potremmo anche pensare, e la cosa sarà molto conveniente, di trascurare il fatto di avere due scatole nere che operano in successione, e immaginare invece di riunire tutti i meccanismi delle due scatole (nella giusta maniera) ottenendo un'unica scatola che, ricevendo il numero 4 in ingresso, fornisce direttamente il numero 6 in uscita: l'unica scatola che sostituisce le due scatole componenti è proprio la funzione composta, che in questo caso diventa

$$f \circ g : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \sqrt{x^2 + 20}\,.$$

È estremamente importante l'ordine in cui le due funzioni agiscono: se scambiassi il ruolo, facendo agire prima la f e poi la g, in corrispondenza del numero 4 in ingresso otterrei un risultato completamente diverso, precisamente 24.



In questo caso la funzione composta è

$$f \circ g : \mathbb{R}^+ \to \mathbb{R}, \quad x \mapsto \left(\sqrt{x}\right)^2 + 20 = x + 20\,,$$

dove, vedi anche la formula (2.29), dobbiamo tenere conto che

$$\left(\sqrt{x}\right)^2 = x \quad \text{solo se} \quad x \geq 0.$$

Dunque la composizione di funzioni *non* gode della proprietà commutativa.

Occorre anche tenere presente che l'output della prima scatola, che diventerà input della seconda, dovrà essere "accettabile" come input, cioè dovrà appartenere al dominio (di solito il dominio naturale) della seconda funzione. Se per esempio la prima funzione fosse $g(x) = x - 5$ e la seconda $f(x) = \sqrt{x}$, è chiaro che il dato 4 in ingresso non sarà accettabile, nel complesso: lo sarebbe se avessi solo la g, ma non lo è più se l'output di g deve poi venir inserito come input di f.

Si noti, nella scrittura $f \circ g$, che g è la prima funzione che agisce, mentre f è la seconda funzione. Questo è in accordo con le nostre notazioni: quando scrivo, per una generica funzione, $h(x)$, intendo che devo considerare il numero x e successivamente inserirlo come dato di ingresso nella funzione f, per ottenere l'opportuno output.

Dal punto di vista pratico, per costruire correttamente la composta di due funzioni converrà denotare con x il nome della variabile della prima funzione che opera (quella più "interna"), con t il suo output, e ancora con t la variabile della seconda funzione (quella più "esterna").

Esempio 4.5. Date le funzioni

$$f(x) = \frac{x}{x+2} \quad \text{e} \quad g(x) = \frac{x^2+1}{x-1},$$

per scrivere esplicitamente la funzione composta $f \circ g$ poniamo

$$t = g(x) = \frac{x^2+1}{x-1} \quad \text{e} \quad f(t) = \frac{t}{t+2}.$$

Basterà ora, nell'espressione di f, sostituire a t il valore $g(x)$:

$$(f \circ g)(x) = f\big(g(x)\big) = \frac{\dfrac{x^2+1}{x-1}}{\dfrac{x^2+1}{x-1} + 2} = \frac{x^2+1}{x^2+2x-1}.$$

Dal dominio della funzione così ottenuta dovremo escludere il punto 1 (perché non fa parte del dominio della prima funzione componente) e i valori di x tali che $g(x) = -2$, ovvero $x = -1 \pm \sqrt{2}$, come si può trovare facilmente. Si noti che il punto 1 potrebbe anche appartenere al dominio naturale della funzione sopra ottenuta, ma deve invece essere escluso perché la funzione si pensa ottenuta come composizione di due funzioni.

Se invece vogliamo scrivere esplicitamente la funzione composta $g \circ f$, poniamo

$$t = f(x) = \frac{x}{x+2} \quad \text{e} \quad g(t) = \frac{t^2+1}{t-1}.$$

Basterà ora, nell'espressione di g, sostituire a t il valore $f(x)$:

$$(g \circ f)(x) = g\big(f(x)\big) = \frac{\left(\dfrac{x}{x+2}\right)^2 + 1}{\dfrac{x}{x+2} - 1} = -\frac{x^2+2x+2}{x+2}.$$

Per quanto riguarda il dominio, basterà solo escludere il punto -2 (che non fa parte del dominio di f); si verifica infatti facilmente che $f(x)$ non assume mai il valore 1 (che è proibito per g).

4.7. Funzione inversa

Se la funzione $f: A \to B$ è biunivoca significa che non solo ad ogni $x \in A$ corrisponde un unico $y \in B$, come per tutte le funzioni, ma che anche, viceversa, ogni $y \in B$ proviene da un solo $x \in A$. In un diagramma a frecce la situazione sarebbe simile a quella della figura 4.17.

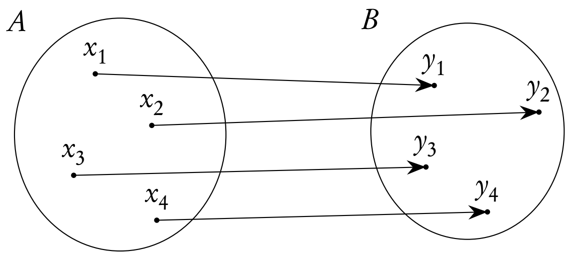


Figura 4.17.: *Funzione biunivoca, diagramma a frecce*

Con riferimento alla figura 4.17, se immaginiamo di invertire il verso di tutte le frecce otteniamo la situazione rappresentata nella figura 4.18 e possiamo pensare che anch'essa si riferisca ad una funzione, precisamente la funzione, questa volta da B in A che "rispedisce le frecce al mittente": anch'essa è una funzione biunivoca che si chiama *funzione inversa* della f. Per questo motivo una funzione biunivoca si dice anche *invertibile*.

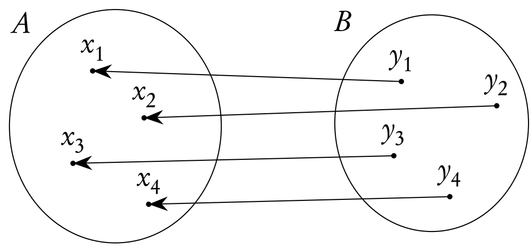


Figura 4.18.: *Funzione inversa, diagramma a frecce*

Precisamente si dà la seguente definizione.

Definizione 4.10 (Funzione inversa). *Se $f: A \to B$ è una funzione biunivoca, si può definire una funzione $g: B \to A$ nel seguente modo: per ogni $y \in B$, $g(y)$ è quell'unico $x \in A$ tale che sia $y = f(x)$. La funzione g si chiama* funzione inversa *della f e si indica con f^{-1}.*

Purtroppo la notazione f^{-1} è, almeno in parte, infelice, in quanto fa pensare al reciproco: si deve invece tenere sempre presente che la funzione inversa di una funzione biunivoca non ha niente a che fare con la funzione reciproco[3]. Alcuni testi preferiscono usare scritture come $f^{\leftarrow}$ o simili. Tuttavia la notazione f^{-1} è quella più diffusa e quella ufficialmente adottata dalle norme ISO, per cui preferiamo attenerci ad essa.

Si noti che, in base alla definizione stessa di funzione inversa, valgono le seguenti due identità.

$$(f \circ f^{-1})(y) = f\big(f^{-1}(y)\big) = y\,, \forall y \in B\,, \quad \text{e} \quad (f^{-1} \circ f)(x) = f^{-1}\big(f(x)\big) = x\,, \forall x \in A. \tag{4.9}$$

Per quanto riguarda il grafico cartesiano delle funzioni reali, possiamo osservare che, dato il grafico di una funzione f invertibile, esso può essere pensato anche come grafico della inversa f^{-1}, purché, per l'inversa, si pensi il dominio sull'asse y e il codominio sull'asse x. Poiché però per tradizione, e per evitare confusioni, si suole sempre rappresentare il dominio sull'asse x, per passare dal grafico di una funzione f a quello della sua inversa, basterà fare il simmetrico del grafico di f rispetto alla bisettrice del primo e terzo quadrante, operazione che consiste nello scambiare, per l'inversa, l'asse y con l'asse x.

In effetti nello scrivere la formula (4.9) avremmo anche potuto usare sempre la variabile x, anche per la funzione f^{-1}, proprio perché, di solito, si preferisce indicare con x la variabile indipendente:

$$(f \circ f^{-1})(x) = f\big(f^{-1}(x)\big) = x\,, \forall x \in B\,, \quad \text{e} \quad (f^{-1} \circ f)(x) = f^{-1}\big(f(x)\big) = x\,, \forall x \in A. \tag{4.10}$$

Osservando la figura 4.19 ci si rende facilmente conto del fatto che, senza didascalie, non sarebbe possibile scoprire qual è il grafico di f e quale quello di f^{-1}: in effetti le due funzioni sono semplicemente una l'inversa dell'altra. Questo si traduce nella formula

$$\big(f^{-1}\big)^{-1} = f\,. \tag{4.11}$$

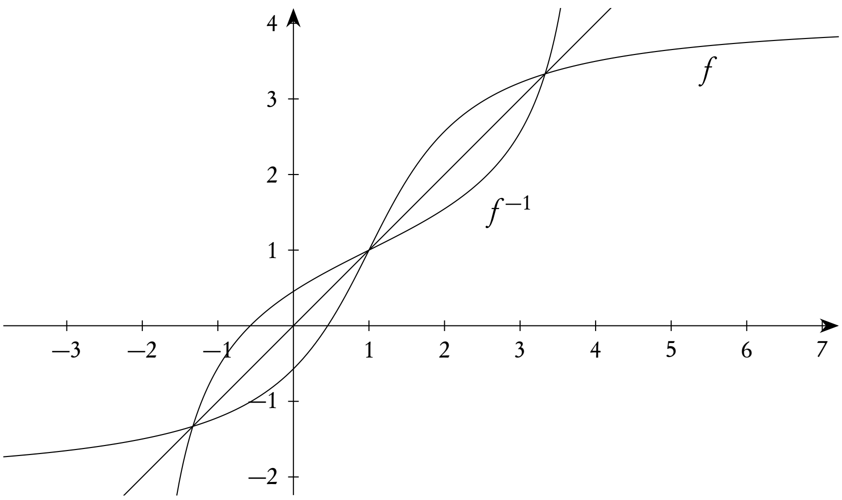


Figura 4.19.: *Grafico di una funzione e della sua inversa*

In molti casi si ha necessità di costruire funzioni inverse anche nel caso di funzioni che non sono biunivoche: per procedere occorre prevedere opportune restrizioni, sia sul dominio che sul codominio. Vediamo un esempio di grande importanza.

[3] Si noti, tra l'altro, che per poter parlare di reciproco di una funzione occorre che sul codominio della funzione sia definito il passaggio al reciproco: da $f(x)$ si deve poter passare ad $1/f(x)$. Ciò avviene per le funzioni a valori reali (naturalmente se non si annullano), ma non per una funzione generica.

Esempio 4.6. La funzione $f:\mathbb{R}\to\mathbb{R}$, $x\mapsto x^2$ (o, brevemente, la funzione $f(x)=x^2$) non è nè iniettiva nè suriettiva. Se però consideriamo la funzione $g:\mathbb{R}^+\to\mathbb{R}^+$, $x\mapsto x^2$, ottenuta operando una restrizione sia sul dominio che sul codominio, otteniamo una funzione biunivoca e quindi invertibile. Ebbene è proprio l'inversa di g che si chiama funzione *radice quadrata*, e *non* l'inversa di f (che non esiste!): $g^{-1}(x)=\sqrt{x}$. Dunque, in accordo con quanto già visto nella formula (2.29), dalla

$$g\big(g^{-1}(x)\big)=x$$

si ricava

$$\left(\sqrt{x}\right)^2=x\,,\quad \text{solo se}\quad x\geq 0\,. \tag{4.12}$$

Si ha poi

$$\sqrt{x^2}\neq x\,,$$

in quanto

$$\sqrt{x^2}$$

significa

$$g^{-1}\big(f(x)\big)\,.$$

Si ha invece, come già visto nella formula (2.28),

$$\sqrt{x^2}=|x|,\quad \forall x\in\mathbb{R}. \tag{4.13}$$

Dal punto di vista operativo, per trovare l'inversa di una funzione reale, dopo aver controllato che si tratta di una funzione biunivoca, si può procedere a scrivere l'equazione $y=f(x)$ e successivamente cercare di "ricavare x in funzione di y": la funzione di y ottenuta è l'inversa cercata e, se si vuole, a questo punto si può anche cambiare la y in x per soddisfare alla convenzione sul nome della variabile indipendente.

Esempio 4.7. Sia

$$f(x)=\frac{x-1}{x+1}\,,$$

di dominio naturale $\mathbb{R}\setminus\{-1\}$. Si può provare, mediante i calcoli che ci accingiamo a fare, che se si restringe il codominio a $\mathbb{R}\setminus\{1\}$ si ottiene una funzione biunivoca e quindi invertibile, che indichiamo con g. Si ha allora

$$y=\frac{x-1}{x+1}\Rightarrow xy+y=x-1\Rightarrow x-xy=y+1\Rightarrow x=\frac{y+1}{1-y}.,$$

da cui

$$g^{-1}(x)=\frac{x+1}{1-x}\,.$$

È facile provare che, per $x\neq 1$, $g\big(g^{-1}(x)\big)=x$ e che, per $x\neq -1$, $g^{-1}\big(g(x)\big)=x$:

$$g\big(g^{-1}(x)\big)=\frac{\dfrac{x+1}{1-x}-1}{\dfrac{x+1}{1-x}+1}=\frac{2x}{2}=x\,,\quad g^{-1}\big(g(x)\big)=\frac{\dfrac{x-1}{x+1}+1}{1-\dfrac{x-1}{x+1}}=\frac{2x}{2}=x\,.$$

Purtroppo i calcoli non sono sempre così semplici e anche le eventuali restrizioni da adottare non sono così immediatamente determinabili.

4.8. Qualche funzione elementare

Alcune funzioni “di base” giocano un ruolo molto importante nelle applicazioni e costituiscono i pilastri su cui è costruita una gran parte delle funzioni di uso comune. La loro importanza risulta evidente se si pensa che per ognuna di esse è previsto un apposito tasto in tutte le calcolatrici scientifiche: funzioni più complesse si ottengono come combinazioni di quei tasti (composizione di funzioni) o mediante l’uso delle operazioni di somma, sottrazione, prodotto e quoziente. Esamineremo qui di seguito le più importanti.

Molti dei risultati che presenteremo dovranno essere accettati “senza discussione”: nei successivi corsi universitari verranno date le giustificazioni necessarie.

4.8.1. La funzione polinomiale di primo grado

La prima funzione che esaminiamo è *la funzione polinomiale di primo grado*, definita da

$$f(x) = mx + q\,, \text{ oppure, brevemente, } y = mx + q\,. \tag{4.14}$$

Essa ha come grafico cartesiano una retta (*non* parallela all’asse y). Il numero m (coefficiente della x) si chiama il *coefficiente angolare* o *pendenza*, il numero q si chiama *ordinata all’origine*. Il nome e il significato geometrico di q sono evidenti: al valore $x = 0$ (origine) corrisponde l’ordinata $y = q$. Per capire il significato di m consideriamo due esempi: la funzione $y = 2x + 1$ e la funzione $y = -2x + 1$.

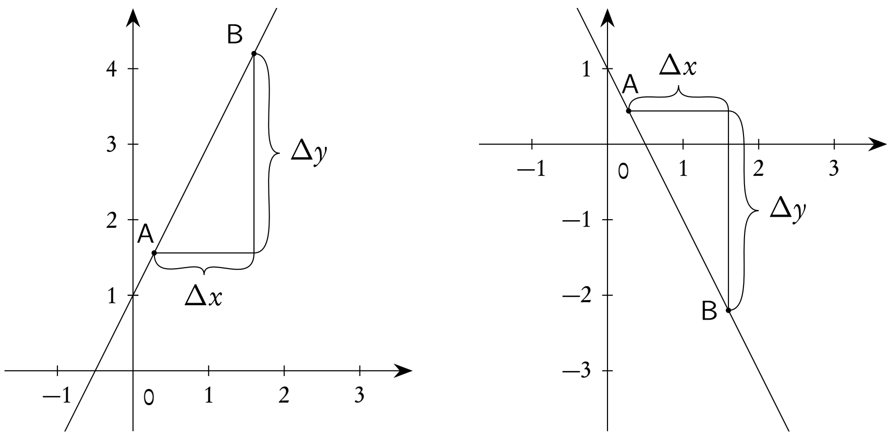


Figura 4.20.: *Significato geometrico del coefficiente angolare*

Se consideriamo due punti A e B qualunque sulle rette, e indichiamo con Δx e Δy, rispettivamente, lo spostamento orizzontale (col segno + se è verso destra, − se è verso sinistra) e quello verticale (col segno + se è verso l’alto, − se è verso il basso) che si ha nel passaggio dal primo al secondo punto (nella figura da A a B), si ha

$$\Delta y = y_{\mathsf{B}} - y_{\mathsf{A}} \quad , \quad \Delta x = x_{\mathsf{B}} - x_{\mathsf{A}}\,,$$

e quindi

$$\Delta y = y_{\mathsf{B}} - y_{\mathsf{A}} = (mx_{\mathsf{B}} + q) - (mx_{\mathsf{A}} + q) = m(x_{\mathsf{B}} - x_{\mathsf{A}}) = m\Delta x\,,$$

da cui

$$m = \frac{\Delta y}{\Delta x}. \tag{4.15}$$

Se si scelgono A e B in modo che $\Delta x = 1$, allora $m = \Delta y$; se si scelgono invece A e B in modo che $\Delta x = 100$, allora $\Delta y = 100\,m$ e il numero $100\,m$ si chiama pendenza percentuale. Le due rette precedenti hanno una pendenza percentuale del 200% e del −200% rispettivamente.

È ovvio che pendenza positiva implica "retta in salita", pendenza negativa implica "retta in discesa" (procedendo da sinistra a destra naturalmente!).

È naturale conseguenza di quanto detto che due rette $y = m_1x + q_1$ e $y = m_2x + q_2$ sono parallele se e solo se hanno lo stesso coefficiente angolare: $m_1 = m_2$.

Si noti che più m è grande, più la retta tende a essere verticale: se la retta potesse essere verticale, m diventerebbe infinitamente grande in senso positivo ($+\infty$) o negativo ($-\infty$). Le rette verticali *non* possono essere grafici di funzioni, ma dal punto di vista geometrico hanno perfettamente senso e si usa dire che il loro coefficiente angolare "è infinito".

Sono particolarmente importanti i casi con $m = \pm 1$ e $q = 0$. Nel caso $m = 1$ si ha $f(x) = x$ e questa funzione è detta anche *funzione identica in* $\mathbb{R}$, in quanto associa ad ogni x di $\mathbb{R}$ lo stesso x: è indicata anche con $I_{\mathbb{R}}$ o con $1_{\mathbb{R}}$. Il suo grafico (in un sistema monometrico!) è la bisettrice del primo e terzo quadrante (a volte chiamata anche la *diagonale* di $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$). Nel caso $m = -1$ si ha $f(x) = -x$ e questa funzione è anche detta *funzione opposto in* $\mathbb{R}$, in quanto ad ogni x di $\mathbb{R}$ associa il suo opposto: il suo grafico (in un sistema monometrico!) è la bisettrice del secondo e quarto quadrante.

Notiamo che se $m = 0$ la funzione in considerazione non è più di primo grado, ma di grado 0: si tratta delle cosiddette *funzioni costanti*, che hanno come grafico rette parallele all'asse delle ascisse, cioè con pendenza nulla: si tratta dunque di un caso particolare di retta, per cui, a volte, anche questo tipo di funzioni sono comprese nella dicitura "funzioni polinomiali di primo grado".

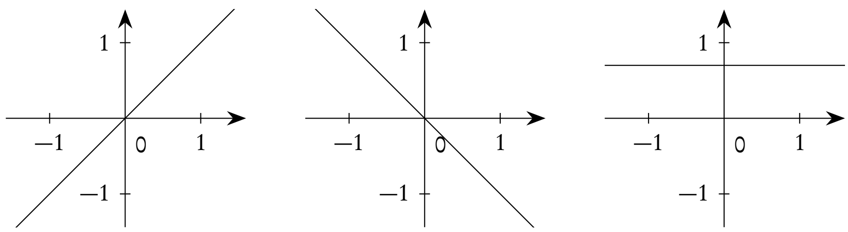


Figura 4.21.: *Funzione identica, funzione opposto, funzione costante, in* $\mathbb{R}$

Il caso in cui $q = 0$ (retta passante per l'origine) è di grande importanza applicativa. In questa situazione si ha

$$\frac{y}{x} = m = \text{costante, se} \quad x \neq 0 :$$

la y e la x sono *direttamente proporzionali*, ovvero se x aumenta o diminuisce, altrettanto succede alla y, e dello stesso fattore.

Le funzioni del tipo $f(x) = mx + q$ in cui $q = 0$ sono anche dette *funzioni lineari*, mentre quelle in cui $q \neq 0$ son anche dette *funzioni affini*[4].

Torneremo su questo argomento e su altre caratteristiche delle funzioni polinomiali di primo grado trattando la retta in geometria analitica.

4.8.2. La funzione polinomiale di secondo grado

La seconda funzione di base di cui ci interessiamo è quella definita da

$$f(x) = ax^2 + bx + c, \text{ oppure, brevemente, } y = ax^2 + bx + c, \quad a \neq 0, \tag{4.16}$$

che chiameremo *funzione polinomiale di secondo grado*.

Essa ha come grafico una parabola con *asse verticale*. Le caratteristiche che più interessano sono le seguenti:

1. la parabola volge la concavità verso l'alto ("sorride") se $a > 0$, volge la concavità verso il basso ("piange")[5] se $a < 0$;
2. il *vertice* della parabola ha ascissa data da
$$x_V = -\frac{b}{2a},$$
mentre l'ordinata dello stesso può essere ricavata semplicemente per sostituzione nell'espressione della funzione, senza necessità di memorizzare formule particolari;
3. la parabola è simmetrica rispetto al suo asse, cioè alla retta verticale passante per il vertice.

Anche su questo argomento torneremo estesamente nel capitolo sulla geometria analitica.

4.8.3. La funzione valore assoluto

Abbiamo già introdotto, nella definizione 2.18, il concetto di valore assoluto di un numero reale. Occupiamoci ora della funzione reale definita mediante il valore assoluto:

$$f(x) = \operatorname{abs} x = |x| = \begin{cases} x, & \text{se } x \geq 0; \\ -x, & \text{se } x < 0. \end{cases}$$

Il suo grafico cartesiano si può ottenere subito, a partire dalla definizione data, ed è mostrato nella figura 4.22.

[4] Anche se non fa strettamente parte delle conoscenze tipiche della matematica di base, segnaliamo che il nome funzioni lineari è legato al fatto che queste funzioni di $\mathbb{R}$ in $\mathbb{R}$ godono delle seguenti due proprietà:

1. $f(kx) = kf(x)$;
2. $f(x_1 + x_2) = f(x_1) + f(x_2)$.

L'importanza di queste proprietà apparirà manifesta nei successivi corsi universitari.

[5] L'origine di questa nomenclatura è chiaramente legata agli smile comunemente usati nel web: ":-)" e ":-(".

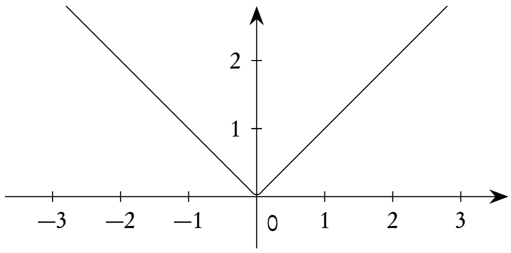


Figura 4.22.: *La funzione valore assoluto*

4.8.4. Le funzioni potenza ad esponente intero

La nozione di potenza ad esponente intero ci consente di definire le cosiddette *funzioni potenza ad esponente intero*: si tratta delle funzioni reali definite dalla legge

$$x \mapsto x^m, \quad m \in \mathbb{Z}.$$

Se $m > 0$ il loro dominio naturale è $\mathbb{R}$, se invece $m \leq 0$ il loro dominio naturale è $\mathbb{R} \setminus \{0\}$. Le indicheremo, a volte, con il simbolo p_m: $p_m(x) = x^m$.

Osserviamo che se m è pari si ha $x^m = (-x)^m$, ovvero $p_m(x) = p_m(-x)$: funzioni che godono di questa proprietà si chiamano *funzioni pari*; se invece m è dispari si ha $x^m = -(-x)^m$: funzioni che godono di questa proprietà si chiamano *funzioni dispari*. Troveremo in seguito altre funzioni con queste proprietà (per esempio le funzioni coseno e seno). Le funzioni pari hanno un grafico cartesiano simmetrico rispetto all'asse delle ordinate, quelle dispari hanno un grafico simmetrico rispetto all'origine (naturalmente perché questo discorso abbia senso il dominio deve essere simmetrico rispetto all'origine, cosa che succede sicuramente per le funzioni potenza che stiamo esaminando. A titolo d'esempio, nella figura 4.23 sono rappresentati i grafici di una funzione pari e di una funzione dispari. La funzione valore assoluto, più sopra introdotta, è un ulteriore esempio di funzione pari.

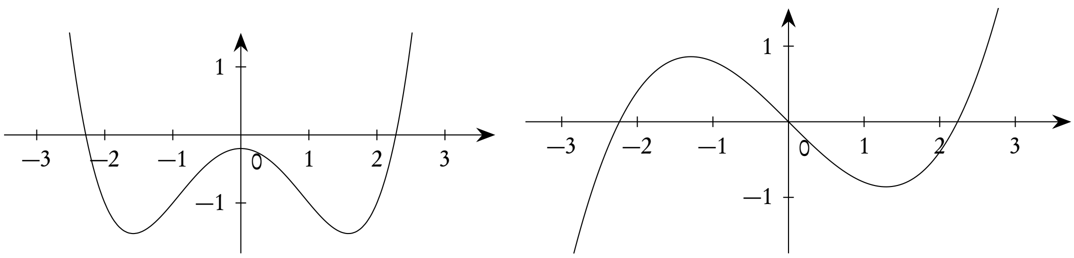


Figura 4.23.: *Esempio di funzione pari, a sinistra, e dispari, a destra*

Dunque per le funzioni potenza sarà sufficiente tracciare il grafico solo per $x \geq 0$ (o per $x > 0$ se l'esponente è minore o uguale a 0).

Alcuni casi sono di particolare importanza nelle applicazioni e li tratteremo per primi.

Potenza con esponente 0

Poiché

$$p_0(x) = x^0 = 1, \quad x \neq 0,$$

il grafico è quello della funzione costantemente uguale ad 1, tranne nell'origine dove non è definita. Tuttavia di solito si "estende" questo grafico fino a comprendere anche l'origine: si dice che si fa un *prolungamento* della funzione all'origine. Attenzione: questo *non* significa definire 0^0 (che sappiamo non avere alcun significato), significa soltanto che si conviene di assegnare come domino a p_0 tutto $\mathbb{R}$, ponendo $p_0(0) = 1$.

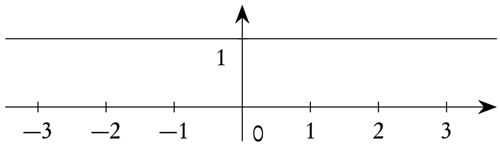


Figura 4.24.: *La funzione p_0, prolungata nell'origine*

Potenza con esponente 1

Il caso di p_m con $m = 1$ rientra tra le funzioni polinomiali di primo grado, precisamente si tratta della funzione identica, già trattato nella pagina 119.

Potenza con esponente 2

Il caso di p_m con $m = 2$ rientra tra le funzioni polinomiali di secondo grado, precisamente con $a = 1$, $b = c = 0$: si tratta dunque di una funzione avente come grafico una parabola con asse verticale e vertice nell'origine.

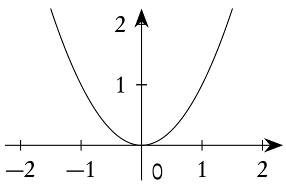


Figura 4.25.: *La funzione $p_2(x) = x^2$*

Potenza con esponente −1

Se $m = -1$ si ha

$$p_{-1}(x) = x^{-1} = \frac{1}{x}, \quad x \neq 0.$$

Questa funzione è detta anche, per ovvi motivi, *funzione reciproco*. Il suo grafico è rappresentato nella figura 4.26 ed è una *iperbole equilatera*.

Si noti che, scrivendo la funzione nella forma $y = 1/x$, si ha, per $x \neq 0$, $xy = 1$, che significa che il prodotto tra x ed y è costantemente uguale ad 1, ovvero x ed y sono inversamente proporzionali: all'aumentare di una, l'altra diminuisce dello stesso fattore. In generale una legge di *proporzionalità inversa* è del tipo $y = k/x$ e la funzione reciproco ne è un caso particolare. Il grafico di $f(x) = k/x$ è rappresentato nella figura 4.27, a sinistra con $k = 2$, a destra con $k = -2$.

Potenze con esponente positivo

Nella figura 4.28 sono riportati i grafici delle funzioni potenza per alcuni esponenti positivi, compresi gli esponenti 1 e 2 già trattati. Sono tracciati i grafici solo per $x \geq 0$, per ottenere il grafico per $x < 0$

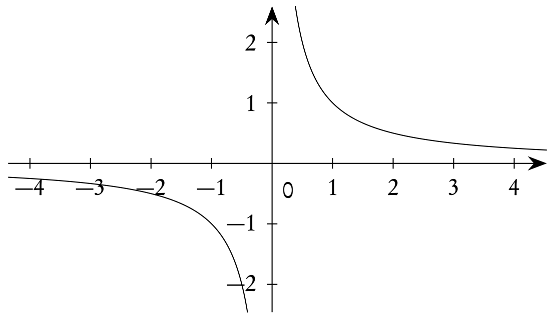


Figura 4.26.: *La funzione* $f(x) = 1/x$

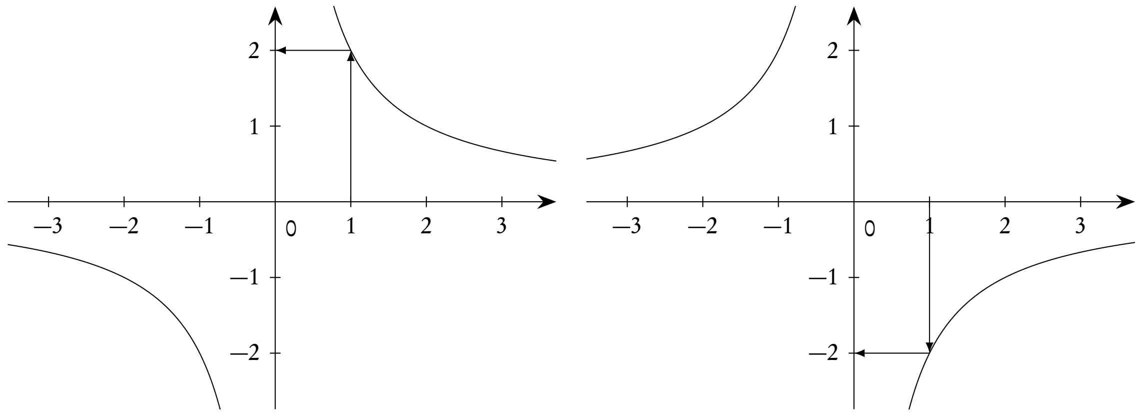


Figura 4.27.: *Leggi di proporzionalità inversa con* $k > 0$ *a sinistra e* $k < 0$ *a destra*

basta operare una simmetria, come già osservato. Si noti che tutte le funzioni valgono 0 per $x = 0$ e 1 per $x = 1$, come è ovvio; inoltre se $0 < x < 1$, quanto più l'esponente m è grande tanto più x^m è piccolo, il contrario se $x > 1$. Nella stessa figura 4.28, sulla destra, è mostrato l'ingrandimento del quadrato $[0,1] \times [0,1]$, per maggiore chiarezza.

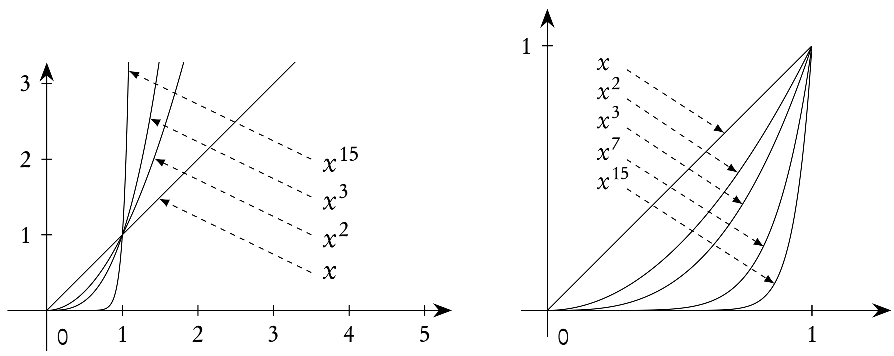


Figura 4.28.: *Potenze con vari esponenti positivi e ingrandimento del quadrato* $[0,1] \times [0,1]$

Potenze con esponente negativo

Nella figura 4.29 sono riportati i grafici delle funzioni potenza per alcuni esponenti negativi, compreso l'esponente -1 già trattato. Sono tracciati i grafici solo per $x > 0$, per ottenere il grafico per $x < 0$ basta operare una simmetria, come già osservato. Si noti che tutte le funzioni risultano non definite per $x = 0$ e valgono 1 per $x = 1$, come è ovvio; inoltre se $0 < x < 1$, quanto più l'esponente m è grande tanto più x^m è grande, il contrario se $x > 1$.

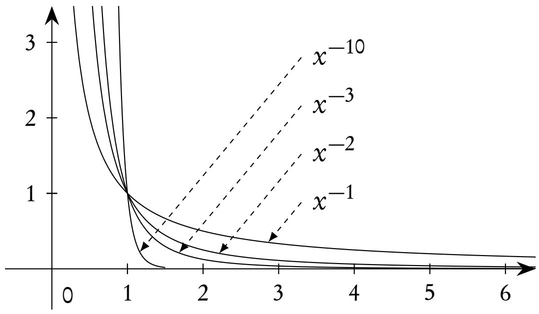


Figura 4.29.: *Potenze con vari esponenti negativi*

4.8.5. Le funzioni radice

Come risulta dai grafici, tutte le funzioni potenza con esponente positivo, se ristrette ai reali non negativi, sono funzioni biunivoche[6] e quindi invertibili (si veda la discussione che abbiamo fatto a proposito della funzione $f(x) = x^2$ nella pagina 117). Le funzioni potenza con esponente positivo dispari maggiore[7] di 1 sono biunivoche, e quindi invertibili, senza bisogno di restrizione. Le inverse di queste funzioni si chiamano *funzioni radice* e sono dunque definite per $x \geq 0$ se $m \geq 2$ pari, per $x \in \mathbb{R}$ se $m > 1$ dispari. I grafici si otterranno dunque per semplice simmetria dei grafici delle funzioni potenza, rispetto alla bisettrice del primo e terzo quadrante: si veda la figura 4.30.

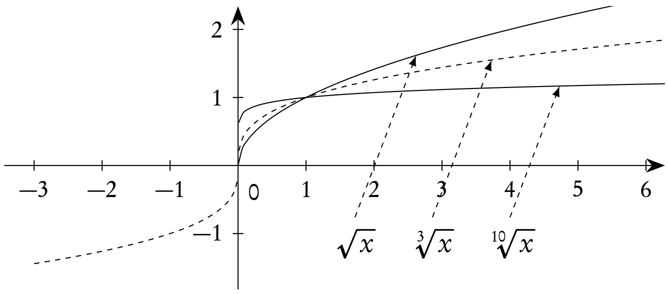


Figura 4.30.: *Alcune funzioni radice*

[6] La cosa andrebbe ovviamente provata in maniera formale, ma qui ci limitiamo semplicemente ad accettarla.
[7] Il caso $m = 1$ è di poco interesse, in quanto l'inversa di $f(x) = x$ coincide con la $f(x)$ stessa.

4.8.6. Le funzioni potenza ad esponente reale

A seguito della estensione della definizione di potenza agli esponenti razionali non interi e agli esponenti irrazionali, potremo definire le funzioni ad esponente reale qualunque α, anche se α è un razionale non intero o un irrazionale:

$$x \mapsto x^{\alpha},$$

aventi dominio naturale $\mathbb{R}^+$ se $\alpha > 0$, $\mathbb{R}^+ \setminus \{0\}$ se $\alpha < 0$.

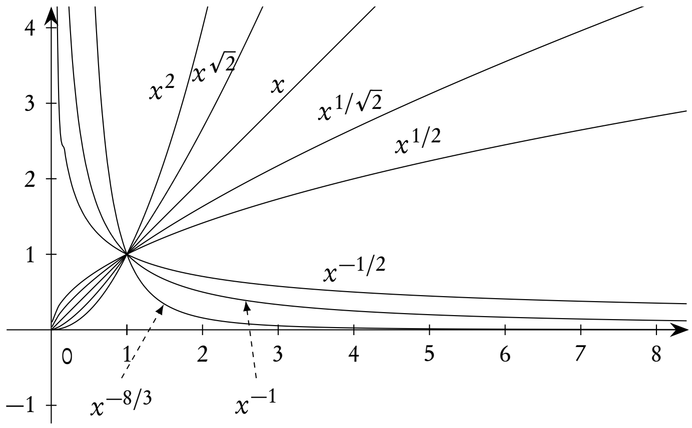


Figura 4.31.: *Alcune funzioni potenza*

Poiché, per definizione,

$$x^{1/m} = \sqrt[m]{x}, \quad m \in \mathbb{N}, \quad m > 1,$$

le funzioni radice potranno essere considerate come caso particolare di potenze ad esponente razionale. Tuttavia segnaliamo che di solito mentre $\sqrt[3]{x}$ si ritiene definita su tutto $\mathbb{R}$, $x^{1/3}$ si ritiene[8] definita solo per $x \geq 0$, e questo per evitare difficoltà con le basi negative, difficoltà già evidenziate nella definizione 2.23. Non entriamo in ulteriori dettagli se non per osservare che queste funzioni estendono le proprietà già viste per le funzioni potenza ad esponente intero e per le funzioni radice. Nella figura 4.31 sono riportati, a titolo d'esempio alcuni grafici per un utile confronto.

4.8.7. Funzioni potenza e invertibilità

Il problema dell'invertibilità delle funzioni potenza è di grande importanza nelle applicazioni e, anche se già sostanzialmente trattato nelle pagine precedenti, presentiamo qui un riepilogo generale, distinguendo i vari casi. In quanto segue il dominio è sempre il dominio naturale e, se necessario, sottindendiamo una restrizione del codominio all'immagine, in modo da non avere problemi con la suriettività.

1. La funzione $p_0(x)$, potenza con esponente zero, che abbiamo esteso a tutto $\mathbb{R}$ ponendo $p_0(0) = 0$, è costante e quindi non invertibile.
2. Le funzioni $f(x) = x^{\pm 1}$ son invertibili e coincidono con le loro inverse.

[8] Questo avviene anche nella maggior parte dei software di calcolo simbolico.

3. Le funzioni $f(x)=x^{\pm 3}$, $f(x)=x^{\pm 5}$, ... (cioè le funzioni potenza con esponente intero dispari diverso da ± 1) sono invertibili e le loro inverse sono $f(x)=\sqrt[3]{x}$, $f(x)=\sqrt[5]{x}$, ... (per esponenti positivi) e $1/\sqrt[3]{x}$, $1/\sqrt[5]{x}$, ... (per esponenti negativi).
4. Le funzioni $f(x)=x^{\pm 2}$, $f(x)=x^{\pm 2}$, ... (cioè le funzioni potenza con esponente intero pari diverso da 0) *non* sono iniettive e quindi *non* sono invertibili. Se però operiamo una restrizione ai reali ≥ 0 (per esponenti positivi) o >0 (per esponenti negativi), otteniamo delle funzioni iniettive e quindi invertibili (ricordiamo che operiamo una restrizione del codominio all'immagine). Le inverse di queste restrizioni sono le funzioni $f(x)=x^{\pm 1/2}$, $f(x)=x^{\pm 1/4}$, ... che coincidono con $f(x)=\sqrt{x}$, $f(x)=\sqrt[4]{x}$, ... (per esponenti positivi) e $f(x)=1/\sqrt{x}$, $f(x)=1/\sqrt[4]{x}$, ... (per esponenti negativi).
5. Le funzioni $f(x)=x^a$, con a non intero, sono invertibili e le loro inverse sono le funzioni $f(x)=x^{1/a}$, con l'avvertenza che, se $a=\pm 1/2$, $a=\pm 1/3$, ..., allora le inverse sono le restrizioni di $f(x)=x^{\pm 2}$, $f(x)=x^{\pm 3}$, ... ai reali ≥ 0 (per a positivo) o >0 (per a negativo): si tenga infatti presente che le potenze con esponente razionale, secondo la nostra convenzione, non sono definite sui reali negativi.

4.8.8. La funzione segno

La *funzione segno* è definita come segue:

$$\operatorname{sgn}(x)=\begin{cases} -1, & \text{se } x<0; \\ 0, & \text{se } x=0; \\ 1, & \text{se } x>0. \end{cases}, \tag{4.17}$$

ed ha il grafico rappresentato nella figura 4.32.

Non tutti concordano nel ritenere la funzione segno definita anche per $x=0$: quella da noi proposta è comunque la definizione ufficialmente adottata dalle norme ISO.

Vale la seguente importante formula che collega le funzioni segno e valore assoluto.

$$x=\operatorname{sgn}(x)\cdot \operatorname{abs}(x)=\operatorname{sgn}(x)\cdot |x|\,. \tag{4.18}$$

Si noti che, dalla (4.18) si deduce

$$\operatorname{sgn}(x)=\frac{x}{|x|}\ \left(=\frac{|x|}{x}\right),\quad x\neq 0. \tag{4.19}$$

Molti autori adottano proprio la (4.19) come definizione della funzione segno (per $x\neq 0$).

4.8.9. Le funzioni parte intera e simili

Le funzioni introdotte in questo paragrafo hanno importanza in molte applicazioni, in particolare sono di uso comune in informatica. Purtroppo la nomenclatura adottata non è universalmente accettata: come al solito qui abbiamo seguito quella prevista dalle norme ISO.

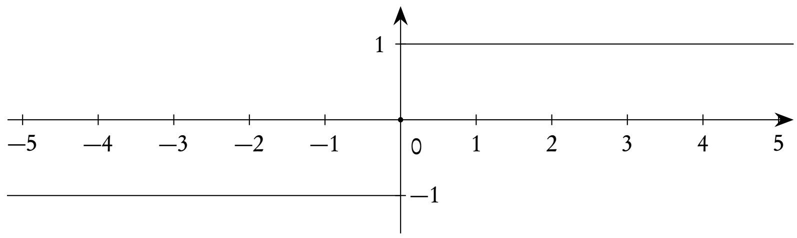


Figura 4.32.: *La funzione segno*

Le funzioni floor e ceil

Fissato un $x \in \mathbb{R}$ esiste[9] un unico intero n tale che $n \leq x < n+1$: si tratta del più grande intero minore o uguale a x. Questo n si indica con $\lfloor x \rfloor$, scrittura che si legge[10] *floor* x. La maggior parte dei matematici chiamano questo n parte intera di x e lo indicano con $[x]$. Noi useremo la dicitura parte intera con un altro significato e sconsigliamo in ogni caso la scrittura con le (troppo abusate!) parentesi quadre. La funzione $x \mapsto \lfloor x \rfloor$, definita su tutto $\mathbb{R}$ ha il grafico rappresentato nella figura 4.33, a sinistra.

Esempio 4.8. $\lfloor -2.5 \rfloor = -3$; $\lfloor 4.6 \rfloor = 4$.

Fissato sempre un $x \in \mathbb{R}$ esiste un unico intero m tale che $m-1 < x \leq m$: si tratta del più piccolo intero maggiore o uguale a x. Questo m si indica con $\lceil x \rceil$, scrittura che si legge[11] *ceil* x. La funzione $x \mapsto \lceil x \rceil$, definita su tutto $\mathbb{R}$ ha il grafico rappresentato nella figura 4.33, a destra.

Esempio 4.9. $\lceil -2.5 \rceil = -2$; $\lceil 4.6 \rceil = 5$.

Si noti che, in base alle definizioni date, si ha

$$\forall x \in \mathbb{R}, \quad \lfloor x \rfloor \leq x \leq \lceil x \rceil, \tag{4.20}$$

ovvero ogni numero reale "sta tra il pavimento e il soffitto".

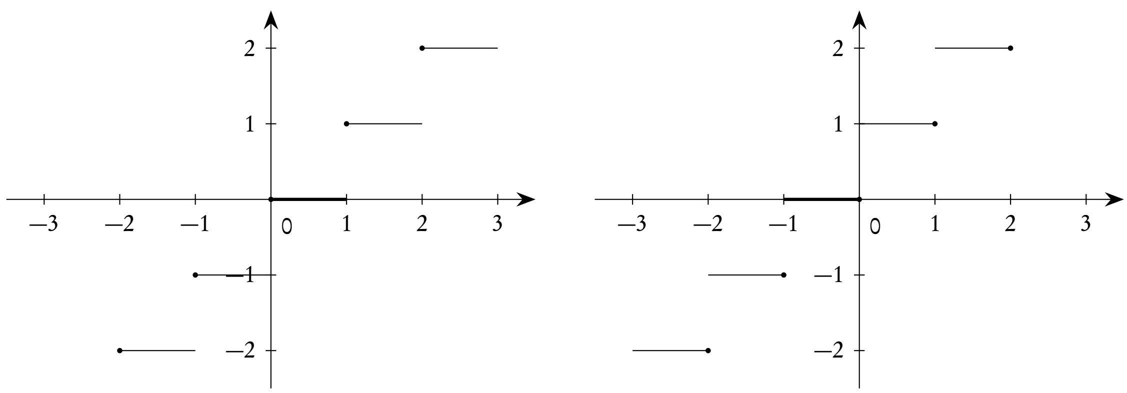


Figura 4.33.: *Le funzioni* floor, *a sinistra, e* ceil, *a destra*

[9] Si tratta di una conseguenza della proprietà di $\mathbb{R}$ di essere archimedeo.

[10] *Floor* ha principalmente il significato di *pavimento* e la nomenclatura è molto appropriata.

[11] *Ceil*, o meglio *ceiling*, ha principalmente il significato di *soffitto* e anche questa nomenclatura è decisamente appropriata. Alcuni software di calcolo simbolico usano proprio il termine *ceiling*, al posto di *ceil*.

La funzione $x - \lfloor x \rfloor$, il cui grafico è rappresentato nella figura 4.34, è chiamata da molti matematici funzione parte frazionaria, mentre secondo la normativa ISO questo nome è usato con un altro significato, di cui parleremo tra poco.

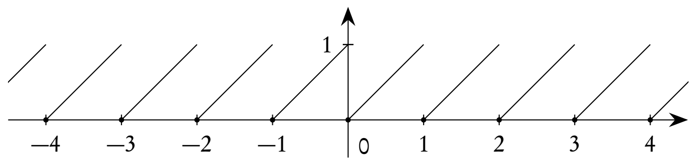


Figura 4.34.: *La funzione* $x - \operatorname{floor} x$

Questa funzione ha comunque il vantaggio di essere sempre positiva e di essere periodica di *periodo* 1: ritorneremo in dettaglio sul concetto di funzione periodica parlando delle funzioni trigonometriche (vedi il capitolo 11).

Le funzioni parte intera e parte frazionaria

Dato un numero reale x, si chiama *parte intera* di x e si indica con $\operatorname{int} x$ il numero definito formalmente da

$$\operatorname{int} x = \operatorname{sgn} x \cdot \lfloor \operatorname{abs} x \rfloor, \tag{4.21}$$

ovvero, brevemente, l'intero con segno che precede la virgola nella rappresentazione decimale di x.

Esempio 4.10. $\operatorname{int}(-0.2) = 0$; $\operatorname{int}(-3.5) = -3$; $\operatorname{int}(4.1) = 4$.

Il grafico della funzione parte intera è rappresentato nella figura 4.35, a sinistra.

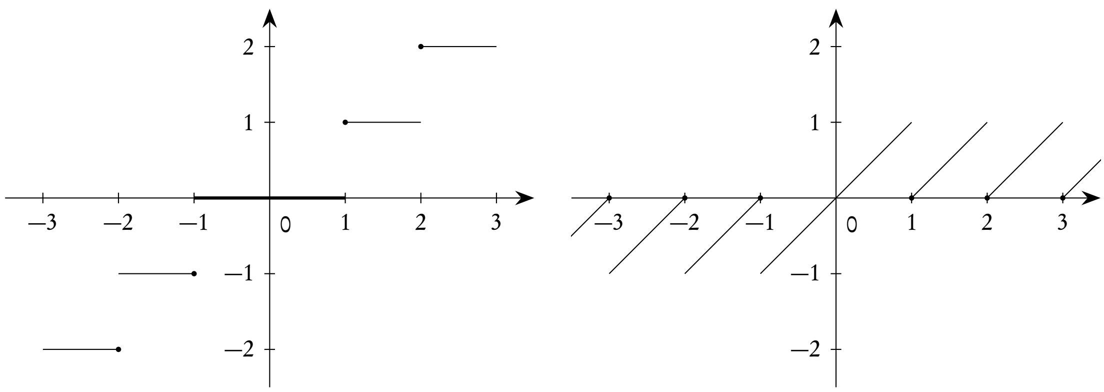


Figura 4.35.: *Funzione parte intera (a sinistra) e parte frazionaria (a destra)*

Dato un numero reale x, si chiama *parte frazionaria* di x e si indica con $\operatorname{frac} x$ il numero definito formalmente da

$$\operatorname{frac} x = x - \operatorname{int} x, \tag{4.22}$$

ovvero, brevemente, la parte che segue la virgola, con il suo segno, nella rappresentazione decimale di x.

Esempio 4.11. $\operatorname{frac}(-0.2) = -0.2 - \operatorname{int}(-0.2) = -0.2 - (0) = -0.2$; $\operatorname{frac}(-3.5) = -3.5 - \operatorname{int}(-3.5) = -3.5 - (-3) = -0.5$; $\operatorname{frac}(4.1) = 4.1 - \operatorname{int}(4.1) = 4.1 - 4 = 0.1$.

Il grafico della funzione parte frazionaria è rappresentato nella figura 4.35, a destra.

Ribadiamo di nuovo che la nomenclatura relativa a questo gruppo di funzioni è ben lungi dall'essere universale: leggendo un testo è sempre bene consultare l'elenco delle notazioni e le relative definizioni.

4.9. Esercizi

Sono qui proposti solo alcuni pochi esercizi di base. Esercizi più significativi sulle funzioni reali saranno proposti nei capitoli successivi.

Esercizio 4.1. *Tracciare il grafico della funzione*

$$f(x) = \operatorname{sgn}(x^2 - 1).$$

Risoluzione. Basterà solo trovare dove $x^2 - 1$ è > 0, $= 0$, < 0: si trova facilmente che si verifica il primo caso per $x > 1 \vee x < -1$, il secondo per $x = \pm 1$, il terzo per $-1 < x < 1$. Il grafico è allora quello della figura 4.36. □

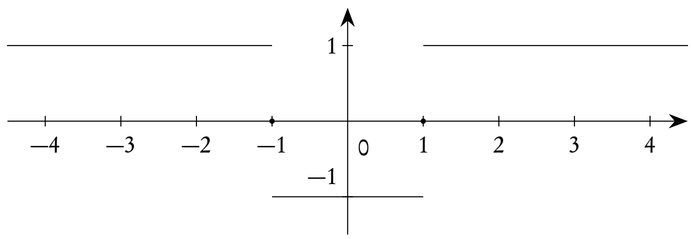


Figura 4.36.: *Grafico della funzione* $f(x) = \operatorname{sgn}(x^2 - 1)$

Esercizio 4.2. *Tracciare il grafico della funzione* $x\lfloor x \rfloor$.

Risoluzione. Conviene procedere esaminando separatamente gli intervalli del tipo $[n, n+1[$, $n \in \mathbb{Z}$.

- In $[-2, -1[$ si ha $x\lfloor x \rfloor = -2x$;
- in $[-1, 0[$ si ha $x\lfloor x \rfloor = -x$;
- in $[0, 1[$ si ha $x\lfloor x \rfloor = 0$;
- in $[1, 2[$ si ha $x\lfloor x \rfloor = x$;
- in $[2, 3[$ si ha $x\lfloor x \rfloor = 2x$;
- $\ldots$;
- in $[n, n+1[$ si ha $x\lfloor x \rfloor = nx$.

Il grafico ha l'aspetto della figura 4.37, dove abbiamo usato diverse unità per i due assi cartesiani. □

Esercizio 4.3. *Tracciare il grafico della funzione* $f(x) = \lfloor x^2 \rfloor$.

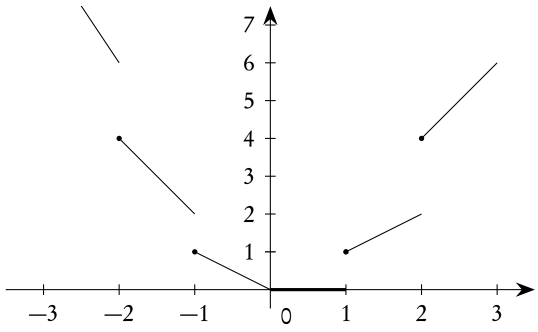


Figura 4.37.: *La funzione* $x\,\mathrm{floor}(x)$

Risoluzione. Si comincia con l'osservare che la funzione è pari (scambiando x con $-x$ non si modifica nulla); successivamente converrà procedere esaminando separatamente gli intervalli del tipo $[n, n+1[$, $n \in \mathbb{N}$.

- Se $0 \leq x^2 < 1$, ovvero $-1 < x < 1$, si ha $f(x) = 0$;
- se $1 \leq x^2 < 2$, ovvero $-\sqrt{2} < x \leq -1 \vee 1 \leq x < \sqrt{2}$, si ha $f(x) = 1$;
- se $2 \leq x^2 < 3$, ovvero $-\sqrt{3} < x \leq -\sqrt{2} \vee \sqrt{2} \leq x < \sqrt{3}$, si ha $f(x) = 2$;
- ecc.

Si ottiene il grafico della figura 4.38, dove abbiamo usato unità diverse sui due assi. □

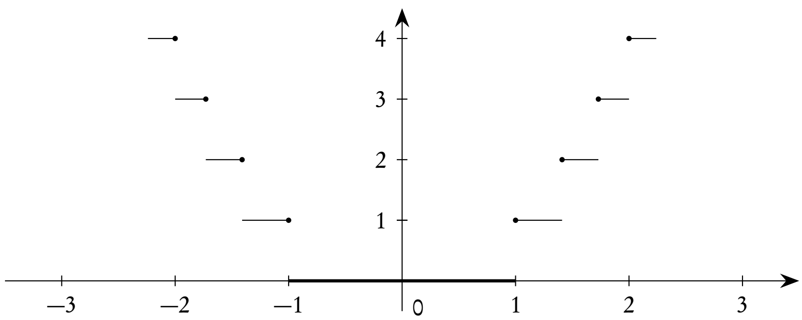


Figura 4.38.: *La funzione* $f(x) = \mathrm{floor}(x^2)$

Esercizio 4.4. *È data la funzione*

$$f(x) = \begin{cases} x, & se\ x \geq 0; \\ x+1, & se\ x < 0. \end{cases} .$$

Si scriva esplicitamente la funzione $h(x) = (f \circ f)(x) = f\big(f(x)\big)$.

Risoluzione. Per evitare ambiguità si possono usare nomi diversi per la funzione "più interna" (la cui variabile chiamiamo x) e per quella più esterna (la cui variabile chiamiamo t), anche se in questo caso coincidono. Dunque posto:

$$t = g(x) = \begin{cases} x, & \text{se } x \geq 0; \\ x+1, & \text{se } x < 0. \end{cases}, \qquad y = f(t) = \begin{cases} t, & \text{se } t \geq 0; \\ t+1, & \text{se } t < 0. \end{cases},$$

dobbiamo calcolare $f\big(g(x)\big)$. Per fare correttamente la composizione dobbiamo determinare, al variare di x, quando $t = g(x)$ è ≥ 0 e quando invece è < 0. Si vede facilmente che $t \geq 0$ se $x \geq -1$, mentre $t < 0$ se $x < -1$. Dunque:

- se $x < -1$, $t = x+1 < 0$ e poiché con questi valori di t si ha $f(t) = t+1$, avremo $f\big(g(x)\big) = (x+1)+1 = x+2$;
- se $-1 \leq x < 0$, $t = x+1 \geq 0$ e poiché con questi valori di t si ha $f(t) = t$, avremo $f\big(g(x)\big) = x+1$;
- se $x \geq 0$, $t = x \geq 0$ e poiché con questi valori di t si ha $f(t) = t$, avremo $f\big(g(x)\big) = x$.

Il grafico di $f\big(f(x)\big)$ è rappresentato nella figura 4.39.

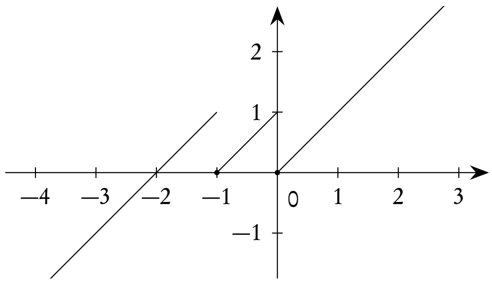


Figura 4.39.: *Grafico di $f\big(f(x)\big)$, ove $f(x) = x+1$ se $x < 0$, $f(x) = x$, se $x \geq 0$*

Si provi per esercizio a scrivere esplicitamente la funzione $f\Big(f\big(f(x)\big)\Big)$. □

Esercizio 4.5. *Tracciare il grafico della funzione*

$$f(x) = x^2 - |x|.$$

Risoluzione. Si può scrivere la funzione data come

$$f(x) = \begin{cases} x^2 - x, & \text{se } x \geq 0; \\ x^2 + x, & \text{se } x < 0. \end{cases}.$$

Il grafico richiesto sarà dunque dato dall'unione di due archi di parabola ed è proposto nella figura 4.40. Si noti che cambiando x in $-x$ la funzione non cambia: questo si riflette nel fatto che il grafico di $f(x)$ a sinistra e a destra dell'asse delle ordinate è identico, ovvero che la funzione data è pari, come mostra chiaramente la figura 4.40. □

Esercizio 4.6. *Tracciare il grafico della funzione*

$$f(x) = |x^2 - 1| + |x - 1|.$$

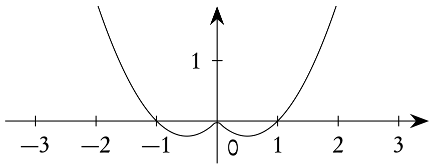


Figura 4.40.: *Grafico di* $f(x) = x^2 - |x|$

Risoluzione. Si può riscrivere la funzione data come

$$f(x) = \begin{cases} x^2 - x, & \text{se } x < -1 \\ -x^2 - x + 2, & \text{se } -1 \leq x < 1 \\ x^2 + x - 2, & \text{se } x \geq 1 \end{cases} .$$

Il grafico è proposto nella figura 4.41. Esso è ottenuto in maniera elementare dall'unione di tre archi di parabola. Si osservi che la funzione data è sempre positiva e questo si riflette nel fatto che il grafico sta sempre sopra l'asse delle ascisse. □

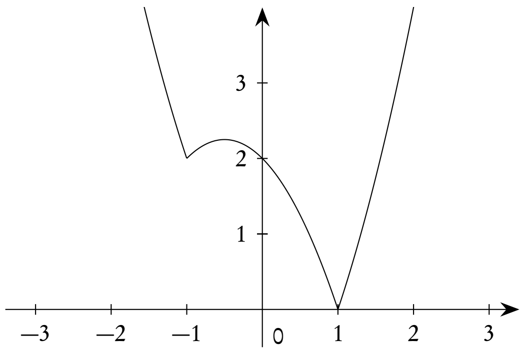


Figura 4.41.: *Grafico della funzione* $f(x) = |x^2 - 1| + |x - 1|$

5. Equazioni e sistemi

5.1. Equazioni in un 'incognita

Risolvere una equazione, nell'incognita reale x, significa risolvere il seguente problema: date due funzioni reali di comune dominio D, $f, g : D \to \mathbb{R}$, trovare il sottoinsieme S di D in cui si ha $f(x) = g(x)$. In termini grafici questo problema equivale a determinare i punti sull'asse x in corrispondenza dei quali i grafici delle due funzioni si intersecano. Se $g(x) = 0$ l'equazione si dice *ridotta in forma normale* e, dal punto di vista dei grafici cartesiani, si riduce a determinare le intersezioni del grafico della funzione f con l'asse delle x.

La variabile x si chiama *incognita* dell'equazione e l'insieme S *insieme delle soluzioni*; ciascuno degli elementi di S si chiama una *soluzione* o anche una *radice* dell'equazione. Nella scrittura $f(x) = g(x)$, $f(x)$ si chiama il *primo membro*, $g(x)$ il *secondo membro* dell'equazione.

Nei casi pratici, di norma, l'insieme D non è assegnato e si conviene che esso sia il dominio naturale comune delle due funzioni f e g: si parla anche di *dominio della equazione*. La ricerca del dominio dell'equazione deve essere sempre fatta prima di eseguire qualsiasi operazione semplificativa sull'equazione data. Chiariamo questo fatto con un semplice esempio.

Esempio 5.1. Per risolvere l'equazione

$$\frac{1}{1/x-1} = 0$$

si può osservare che essa può essere riscritta semplicemente come $x - 1 = 0$: mentre quest'equazione ha solo il numero 1 come soluzione, 1 non può essere soluzione dell'equazione data perché annulla il denominatore.

Le equazioni più importanti sono quelle che, ridotte in forma normale, hanno a primo membro un polinomio o un quoziente di polinomi (funzione razionale). In proposito si dà la seguente definizione.

Definizione 5.1. *Un'equazione che, ridotta in forma normale, ha a primo membro una funzione razionale si dice un'*equazione razionale*; se il primo membro è, in particolare, un polinomio, si dice* razionale intera *e il grado del polinomio si chiama anche grado dell'equazione.*

Definizione 5.2. *Un'equazione nell'incognita reale x si dice un'*identità *se è il suo insieme di soluzioni è* $\mathbb{R}$, *ovvero se è verificata per ogni valore reale di x.*

Spesso si definisce *impossibile* un'equazione che non ha soluzioni, per esempio l'equazione $x^2 + 1 = 0$. In casi come questo ci pare più appropriato dire semplicemente che l'insieme delle soluzioni è vuoto, ma naturalmente si tratta solo di una questione di nomenclatura.

5.1.1. Principi di equivalenza

Definizione 5.3. *Due equazioni si dicono* equivalenti *se hanno lo stesso insieme di soluzioni.*

La risoluzione di un'equazione richiede in genere una serie di passaggi preliminari atti a ridurre l'equazione stessa a una equazione equivalente e appartenente ad una delle forme canoniche che saranno esaminate in seguito. Questi passaggi si basano sulle seguenti note proprietà delle uguaglianze tra numeri, già considerate trattando le proprietà delle operazioni, vedi la pagina 39.

1. Aggiungendo o sottraendo uno stesso numero ai due membri di un'uguaglianza, si ottiene ancora un'uguaglianza.
2. Moltiplicando o dividendo ambo i membri di un'uguaglianza per uno stesso numero *diverso da zero* si ottiene ancora un'uguaglianza.

Da qui si ricavano i seguenti due teoremi detti anche *principi di equivalenza*.

Teorema 5.4 (Primo principio di equivalenza delle equazioni)**.** *Aggiungendo o sottraendo ad ambo i membri di un'equazione, di dominio D, uno stesso numero o una stessa funzione nell'incognita x, con dominio D, si ottiene un'equazione equivalente alla data, naturalmente sempre nel dominio D.*

Esempio 5.2. Le equazioni:

$$x+3=2x-5 \quad \text{e} \quad x+3-2x+5=0, \quad \text{ovvero} \quad -2x+8=0$$

sono equivalenti: la seconda è stata ottenuta dalla prima aggiungendo ad ambo i membri la funzione $-2x+5$.

Esempio 5.3. Le equazioni:

$$x+\frac{1}{x-1}=1+\frac{1}{x-1}, \quad \text{e} \quad x+\frac{1}{x-1}-\frac{1}{x-1}=1+\frac{1}{x-1}-\frac{1}{x-1}, \quad \text{ovvero} \quad x=1$$

non sono equivalenti: la prima non ha soluzioni, la seconda, ottenuta aggiungendo ad ambo i membri la funzione $-1/x-1$ e poi semplificando, ha la soluzione $x=1$. Esse sono però equivalenti se si tiene conto del dominio della disequazione, dato da $\mathbb{R}\setminus\{1\}$: la funzione aggiunta ha in effetti proprio questo dominio.

Questo principio si suole di solito enunciare nella seguente forma semplificata: è possibile portare una quantità da un membro all'altro di un'equazione, *cambiandola di segno*.

Teorema 5.5 (Secondo principio di equivalenza delle equazioni)**.** *Moltiplicando o dividendo ambo i membri di un'equazione di dominio D per uno stesso numero diverso da zero, o per una stessa funzione che abbia sempre per dominio D e che* non si annulli mai *si ottiene un'equazione equivalente alla data.*

Esempio 5.4. Le equazioni

$$x^3+x=x^2+1 \quad \text{e} \quad x=1$$

sono equivalenti: la seconda è stata ottenuta dalla prima dividendo ambo i membri per la funzione x^2+1 che ha dominio $\mathbb{R}$ e non si annulla mai.

Esempio 5.5. Le equazioni

$$x^3+x=x^2+2x \quad \text{e} \quad x^2+1=x+2$$

non sono equivalenti: la prima ha la soluzione $x=0$, la seconda no.

5.1.2. Equazioni di primo grado

Un'equazione di primo grado in un'incognita si può sempre mettere nella forma

$$ax+b=0, \quad a,b\in\mathbb{R},\ a\neq 0. \tag{5.1}$$

Essa ha *sempre* una ed una sola soluzione (si presti attenzione al fatto che si è supposto $a\neq 0$) che si trova semplicemente "portando" b a secondo membro e dividendo per a:

$$x=-\frac{b}{a}. \tag{5.2}$$

Nel caso in cui a e b sono numeri reali non c'è altro da dire; nel caso in cui a o b sono espressioni letterali occorrerà, in genere, una discussione preliminare, come si evince dagli esempi che seguono.

Esempio 5.6. Risolvere l'equazione

$$(a^2-1)x=1.$$

Se $a^2-1\neq 0$ l'equazione è di primo grado e ha l'unica soluzione $x={}^1/_{a^2-1}$. Se invece $a^2-1=0$, ovvero $a=\pm 1$, l'equazione non è di primo grado è non ha nessuna soluzione.

Esempio 5.7. Risolvere l'equazione

$$ax-a+b=0.$$

Se $a\neq 0$ l'equazione è di primo grado e ha l'unica soluzione $x={}^{a-b}/_{a}$. Se $a=0\wedge b\neq 0$ l'equazione non è di primo grado e non ha nessuna soluzione. Se $a=b=0$ si ha un'identità.

Esempio 5.8. Risolvere l'equazione

$$ax=b^2-1.$$

Se $a\neq 0$ l'equazione è di primo grado e ha l'unica soluzione $x={}^{b^2-1}/_{a}$. Se $a=0\wedge b\neq\pm 1$ l'equazione non è di primo grado e non ha nessun soluzione. Se $a=0\wedge b=\pm 1$ si ha un'identità.

È chiaro che risolvere l'equazione $ax+b=0$ equivale a trovare l'unica radice del polinomio di primo grado $P(x)=ax+b$, che, come già noto, è proprio $-{}^b/_a$.

Dal punto di vista grafico la risoluzione dell'equazione $ax+b=0$ equivale a determinare l'intersezione della retta $y=ax+b$ con l'asse delle ascisse: poiché si è supposto $a\neq 0$, esiste una sola intersezione di ascissa, appunto, $-{}^b/_a$.

5.1.3. Equazioni di secondo grado

Un'equazione di secondo grado in un'incognita si può sempre mettere nella forma

$$ax^2+bx+c=0, \quad a,b,c\in\mathbb{R}, \quad a\neq 0. \tag{5.3}$$

Risolvere quest'equazione equivale a determinare gli eventuali zeri del polinomio di secondo grado $P(x)=ax^2+bx+c$ e possiamo dunque riportare semplicemente quanto già detto nella pagina 90:

- se $\Delta=b^2-4ac<0$ l'equazione (5.3) non ha, nei reali, alcuna soluzione;
- se $\Delta=b^2-4ac=0$ l'equazione ha, nei reali, una sola soluzione $x=-{}^b/_{2a}$, o come si usa dire, una soluzione doppia;

— se $\Delta = b^2 - 4ac > 0$ l'equazione ha, nei reali, due soluzioni distinte

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Dal punto di vista grafico risolvere l'equazione (5.3) equivale a determinare le eventuali intersezioni della parabola $y = ax^2 + bx + c$ con l'asse delle ascisse: se $\Delta < 0$ la parabola sta tutta sopra ($a > 0$) o sotto ($a < 0$) l'asse delle ascisse e dunque l'equazione non ha soluzioni; se $\Delta = 0$ la parabola ha il vertice sull'asse delle ascisse in corrispondenza all'unica soluzione dell'equazione; se $\Delta > 0$ la parabola interseca l'asse delle ascisse in due punti con ascissa uguale alle soluzioni dell'equazione. Ritorneremo su questo argomento estesamente nel capitolo sulla geometria analitica.

Se a, b, c sono numeri non c'è altro da dire; se invece a, b, c contengono lettere occorrerà un'analisi dettagliata delle situazioni che si possono presentare, come si evince dagli esempi che seguono.

Esempio 5.9. Risolvere l'equazione

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

Se $a = 0$ l'equazione è di primo grado e ha l'unica soluzione $x = 1/2$. Se $a \neq 0$, si ha $\Delta = 4 - 4a$: dunque se $a = 1$ si ha $\Delta = 0$ e l'equazione ha l'unica soluzione $x = -2/2 \cdot 1 = 1$; se $a < 1$ si ha $\Delta > 0$ e l'equazione ha due soluzioni reali e distinte (date dalla già citata formula); se $a > 1$ si ha $\Delta < 0$ e l'equazione non ha soluzioni.

Esempio 5.10. Risolvere l'equazione

$$(a-1)x^2 - ax + 1 = 0.$$

Se $a = 1$ l'equazione è di primo grado e ha l'unica soluzione $x = 1$; se $a \neq 1$ si ha $\Delta = a^2 - 4a + 4 = (a-2)^2$: dunque se $a = 2$ l'equazione ha la sola soluzione $x = 1$, se $a \neq 2$ l'equazione ha le due soluzioni

$$x_{1,2} = \frac{a \pm (a-2)}{2(a-1)} = \left\langle \begin{array}{c} 1 \\ \\ \dfrac{1}{a-1} \end{array} \right. .$$

Nel caso particolare in cui b è pari è tradizione imparare una "formula ridotta" per risolvere l'equazione di secondo grado (5.3). Riteniamo inutile la memorizzazione di questa ulteriore formula anche perché spesso fonte di errori. Se proprio si vogliono semplificare i calcoli, quando b è pari si possono dividere ambo i membri dell'equazione (5.3) per 2 e applicare poi la solita formula. Si veda l'esempio che segue.

Esempio 5.11. Risolvere l'equazione

$$3x^2 - 4x - 5 = 0, \quad \text{ovvero, dividendo per 2 ambo i membri,} \quad \frac{3}{2}x^2 - 2x - \frac{5}{2} = 0.$$

Applicando la formula nota si ha

$$\Delta = 4 + 15 = 19, \quad \text{quindi} \quad x_{1,2} = \frac{2 \pm \sqrt{19}}{3}.$$

È anche tradizione usare dei nomi particolari (che qui non riportiamo neppure!) per le equazioni di secondo grado nei casi in cui b oppure c o entrambi siano zero: anche in questo caso ci pare assolutamente inutile fare un ulteriore sforzo mnemonico e riteniamo preferibile o utilizzare comunque la formula risolutiva (che funziona sempre!) o servirsi dei ragionamenti elementari che seguono.

— Se $b=0 \land c \neq 0$, si ha $ax^2+c=0$ ovvero $ax^2=-c$ o ancora $x^2=-c/a$: se dunque $-c/a<0$ non si hanno soluzioni, se $-c/a>0$ si hanno le due soluzioni distinte $x_{1,2}=\pm\sqrt{-c/a}$.
— Se $b \neq 0 \land c=0$, si ha $ax^2+bx=0$ ovvero $x(ax+b)=0$ da cui $x_1=0$ e $x_2=-b/a$ (legge dell'annullamento del prodotto).
— Se $b=0 \land c=0$, si ha $ax^2=$ da cui si trova l'unica soluzione $x=0$.

Relazioni tra coefficienti e soluzioni in un'equazione di secondo grado

Se l'equazione (5.3) ha $\Delta \geq 0$ si ha, facilmente,

$$x_1+x_2=\frac{-b-\sqrt{\Delta}}{2a}+\frac{-b+\sqrt{\Delta}}{2a}=-\frac{b}{a}, \tag{5.4}$$

e

$$x_1\cdot x_2=\frac{-b-\sqrt{\Delta}}{2a}\cdot\frac{-b+\sqrt{\Delta}}{2a}=\frac{c}{a}. \tag{5.5}$$

Le formule (5.4) e (5.5) valgono anche se $\Delta<0$, ma in questo caso le radici sono complesse e questo argomento esula dagli scopi di questo testo.

Tenendo conto delle formule (5.4) e (5.5) si ha

$$ax^2+bx+c=0 \quad\Rightarrow\quad x^2+\frac{b}{a}x+\frac{c}{a}=0 \quad\Rightarrow\quad x^2-sx+p=0, \tag{5.6}$$

avendo posto $s=x_1+x_2$ e $p=x_1\cdot x_2$. Si tratta di una forma di scrittura di un 'equazione di secondo grado utile in molte situazioni. Naturalmente, lavorando sui reali, si deve tenere conto che la formula vale solo se $\Delta \geq 0$.

Esempio 5.12. Si voglia scrivere un'equazione di secondo grado avente per soluzioni $\{-1/2, 2/3\}$. Essendo $s=1/6$ e $p=-1/3$, l'equazione cercata sarà

$$x^2-\frac{1}{6}x-\frac{1}{3}=0 \quad \text{ovvero} \quad 6x^2-x-2=0.$$

Si poteva anche procedere osservando che l'equazione deve essere, in base al teorema fondamentale dell'algebra, della forma

$$\left(x+\frac{1}{2}\right)\left(x-\frac{2}{3}\right)=0,$$

da cui si ottiene lo stesso risultato.

Utilizzando le formule (5.4) e (5.5) si possono ricavare anche altre relazioni tra i coefficienti dell'equazione (5.3) e le soluzioni (sempre nell'ipotesi $\Delta \geq 0$ se si vuole lavorare sui reali). Riportiamo alcune delle formule di uso più comune.

— $x_1^2 + x_2^2 = (x_1 + x_2)^2 - 2x_1x_2$.
— $x_1^3 + x_2^3 = (x_1 + x_2)^3 - 3x_1x_2(x_1 + x_2)$.
— $\dfrac{1}{x_1} + \dfrac{1}{x_2} = \dfrac{x_1 + x_2}{x_1x_2}$.

Regola dei segni di Cartesio

In un'equazione di secondo grado $ax^2 + bx + c = 0$, si dice che si ha una *permanenza* se due coefficienti consecutivi (a e b oppure b e c) hanno lo stesso segno, una *variazione* se due coefficienti consecutivi hanno segno opposto.

Vale la seguente *Regola dei segni di Cartesio*: in un'equazione di secondo grado scritta in forma normale e con discriminante positivo, ad ogni variazione corrisponde una soluzione positiva, ad ogni permanenza una soluzione negativa. Se l'equazione ha una permanenza e una variazione le due soluzioni di segno opposto hanno valore assoluto diverso: se la permanenza precede la variazione è maggiore il valore assoluto della soluzione negativa, se la variazione precede la permanenza è maggiore il valore assoluto della soluzione positiva.

5.1.4. Equazioni di grado superiore al secondo

Un'equazione di grado n, scritta in forma normale, è un'equazione del tipo

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0. \tag{5.7}$$

Risolvere un'equazione di questo tipo sarà dunque equivalente a determinare gli zeri del polinomio a primo membro. Come già accennato nella pagina 3.4, esistono formule per trovare le radici di un polinomio di terzo e quarto grado, formule che naturalmente forniranno le radici di equazioni di terzo e quarto grado scritte in forma normale, ma queste formule utilizzano i complessi e quindi esulano dagli scopi di questo testo. Per equazioni di grado superiore al quarto non esiste invece alcuna formula risolutiva e questo tipo di equazioni può essere risolto solo in casi particolari, di alcuni dei quali faremo un breve cenno, senza pretendere di essere esaustivi, per la varietà dei casi che si possono presentare.

Scomposizione in fattori

La prima e più importante tecnica applicabile alla risoluzione di equazioni di grado superiore al secondo è quella della scomposizione del polinomio a primo membro in fattori, con successiva applicazione della legge di annullamento del prodotto. Si possono applicare le tecniche viste nel paragrafo 3.3 e il teorema 3.12 sugli zeri razionali di un polinomio. Non insistiamo oltre su questo argomento, già ampiamente trattato.

Equazioni binomie

Sono così chiamate le equazioni del tipo

$$ax^n + b = 0, \quad a \neq 0, \quad n \in \mathbb{N}, \quad n > 2.$$

La loro risoluzione è immediata:

$$ax^n + b = 0 \quad \Rightarrow \quad x^n = -\frac{b}{a}.$$

A questo punto è sufficiente applicare la definizione di radice n-esima: se n è dispari si ha un'unica soluzione

$$x = \sqrt[n]{-\frac{b}{a}};$$

se invece n è pari, se $-b/a < 0$ non si hanno soluzioni, se $-b/a = 0$ si ha la sola soluzione $x = 0$, se $-b/a > 0$ si hanno le due soluzioni

$$x_{1,2} = \pm\sqrt[n]{-\frac{b}{a}}.$$

Equazioni trinomie

Sono così chiamate le equazioni del tipo

$$ax^{2n} + bx^n + c = 0, \quad a \neq 0, \quad n \in \mathbb{N}, \quad n > 2.$$

Per la loro risoluzione basta porre $x^n = t$ e risolvere l'equazione $at^2 + bt + c$: se questa ha le soluzioni t_1 e t_2 basterà risolvere poi le equazioni $x^n = t_1$ e $x^n = t_2$, con le osservazioni fatte a proposito delle equazioni binomie.

Il caso $n = 2$ è particolarmente frequente e in questo caso l'equazione di chiama anche *biquadratica*.

Equazioni reciproche

Sono così chiamate le equazioni, ridotte a forma normale, in cui il polinomio a primo membro ha i coefficienti equidistanti dagli estremi uguali (equazioni *reciproche di prima specie*) o opposti (equazioni *reciproche di seconda specie*). Notare che le equazioni reciproche di seconda specie di grado pari devono avere il coefficiente del termine centrale nullo, in quanto deve essere l'opposto di se stesso. Le equazioni di prima specie di grado dispari hanno la radice -1, tutte le equazioni di seconda specie hanno come radice il numero 1: in questi casi è possibile applicare la scomposizione in fattori. Le equazioni reciproche di prima specie e di grado pari si riducono ad un'equazione di grado $n/2$ mediante la sostituzione

$$x + \frac{1}{x} = t.$$

Se l'equazione ottenuta è risolubile la tecnica permette la risoluzione anche dell'equazione iniziale. Il metodo funziona ovviamente nel caso $n = 4$, perché l'equazione ottenuta in t è di secondo grado. Si veda l'esempio che segue per capire come procedere.

Esempio 5.13. Si debba risolvere l'equazione $12x^4 + 4x^3 - 41x^2 + 4x + 12 = 0$. Dividiamo ambo i membri per x^2, ottenendo

$$12x^2 + 4x - 41 + 4\frac{1}{x} + 12\frac{1}{x^2} = 0 \quad \Rightarrow \quad 12\left(x^2 + \frac{1}{x^2}\right) + 4\left(x + \frac{1}{x}\right) - 41 = 0.$$

Posto ora $x + {}^1/_x = t$, si ha

$$x+\frac{1}{x}=t \quad \Rightarrow \quad \left(x+\frac{1}{x}\right)=t^2 \quad \Rightarrow \quad x^2+\frac{1}{x^2}=t^2-2.$$

Dunque

$$12(t^2-2)+4t-41=0 \quad \Rightarrow \quad t_{1,2}=\left\langle \begin{array}{l} -\dfrac{5}{2} \\ \dfrac{13}{6} \end{array}\right.$$

Risolvendo ora

$$x+\frac{1}{x}=-\frac{5}{2} \quad \text{e} \quad x+\frac{1}{x}=\frac{13}{6}$$

si trova che l'insieme di soluzioni dell'equazione data è

$$\left\{-2,-\frac{1}{2},\frac{2}{3},\frac{3}{2}\right\}.$$

Tenendo conto della eventuale presenza delle radici ± 1 e della sostituzione indicata per le equazioni di prima specie e di grado pari, si riescono sempre a risolvere le equazioni fino al quinto grado.

Le equazioni in questione si chiamano reciproche perché, come è facile provare, se hanno una radice α, hanno anche la radice reciproca ${}^1/_\alpha$.

5.1.5. Equazioni razionali fratte

Per risolvere un'equazione razionale fratta, ridotta in forma normale, basterà determinare le radici del numeratore e controllare che esse non siano, contemporaneamente, radici del denominatore.

Esempio 5.14. Si debba risolvere l'equazione

$$\frac{x^4-3x^2+2}{x^2-3x+2}=0.$$

Le radici del numeratore sono: $-\sqrt{2},-1,1,\sqrt{2}$. Poiché 1 è anche radice del denominatore, se ne deduce che l'insieme delle soluzioni è $\left\{-\sqrt{2},-1,\sqrt{2}\right\}$.

5.1.6. Equazioni irrazionali

Si chiamano *irrazionali* le equazioni contenenti uno o più radicali e nelle quali l'incognita compare sotto il segno di radice. Non esistono tecniche standard per la risoluzione di questo tipo di equazioni: il procedimento risolutivo consiste comunque nel trasformare l'equazione stessa in una equazione razionale, mediante opportuni elevamenti a potenza, ed è chiaro che l'equazione irrazionale è risolubile se tale è l'equazione razionale ottenuta.

L'applicazione di questo procedimento tuttavia richiede la massima attenzione. Bisogna anzitutto tenere conto che, elevando a potenza ambo i membri di un'equazione, non è detto che si ottenga un'equazione equivalente: occorrerà dunque controllare se le soluzioni ottenute dopo l'elevazione erano anche radici dell'equazione originale. Qualche esempio chiarirà il problema.

Esempio 5.15. L'equazione $x = 1$ è di primo grado e ha la sola radice 1; elevando al quadrato ambo i membri si ottiene l'equazione $x^2 = 1$ che ha le due radici ± 1. In questo caso l'elevazione al quadrato ha introdotto una soluzione estranea.

Esempio 5.16. L'equazione $\sqrt{x^2+1} = -2$ non ha radici (il primo membro è positivo, il secondo negativo); elevando al quadrato e semplificando si ottiene l'equazione $x^2 = 3$ che ha le due radici $\pm\sqrt{3}$. In questo caso l'elevazione al quadrato ha introdotto due soluzioni estranee.

Esempio 5.17. Elevando al quadrato ambo i membri dell'equazione $\sqrt{x^2} = x$ si ottiene l'equazione $x^2 = x^2$ che è un'identità. Solo i reali non negativi sono però soluzioni anche dell'equazione di partenza. In questo caso l'elevazione al quadrato ha introdotto infinite soluzioni estranee

Un'altra difficoltà è data dal fatto che l'elevazione a potenza può complicare il problema anziché semplificarlo, come mostra l'esempio che segue.

Esempio 5.18. Si debba risolvere l'equazione $\sqrt{x}+2 = x$. Se si eleva direttamente al quadrato si ottiene l'equazione $x+4\sqrt{x}+4 = x^2$, che è ancora irrazionale e più complessa della precedente. Se invece la si riscrive in $\sqrt{x} = x-2$ e poi si eleva al quadrato si ottiene $x = x^2-4x+4$ che ha le radici 1 e 4. La prima non soddisfa l'equazione di partenza, da cui si deduce che l'equazione ha la sola radice $x = 4$.

Bisogna dunque procedere con la massima cautela, valutando accuratamente le varie possibilità. In genere se l'equazione contiene un solo radicale conviene isolarlo in uno dei due membri: l'elevazione di ambo i membri all'indice della radice rende l'equazione stessa razionale. Se l'equazione contiene due radicali conviene riscrivere l'equazione con un radicale per ogni membro ed elevare ad un'opportuna potenza ambo i membri, eventualmente ripetendo l'operazione. In altre situazioni (abbastanza rare!) bisogna valutare caso per caso.

Esempio 5.19. Per risolvere l'equazione $x-1-\sqrt[3]{x^3-1} = 0$ la si riscrive in $x-1 = \sqrt[3]{x-1}$; elevando al cubo si ottiene $3x^2-3x = 0$ che ha come radici 0 e 1, che sono anche[1] soluzioni dell'equazione data.

Esempio 5.20. Per risolvere l'equazione $\sqrt{x+2} = \sqrt[3]{3x+2}$ si elevano ambo i membri alla sesta potenza, ottenendo l'equazione $x^3-3x^2+4 = 0$, che ha come radici -1 e 2. Solo 2 è radice anche dell'equazione di partenza.

Esempio 5.21. Per risolvere l'equazione $\sqrt{7-x}+\sqrt{4-x} = 3$ si eleva al quadrato ottenendo, dopo semplificazione, $\sqrt{(7-4x)(4-x)} = x-1$. Una successiva elevazione al quadrato porta alla soluzione $x = 3$, che è anche soluzione dell'equazione di partenza.

Esempio 5.22. Per risolvere l'equazione $\sqrt{2x-1}+\sqrt{2x+1} = 1+\sqrt{4x-1}$ si elevano al quadrato ambo i membri ottenendo $\sqrt{4x^2-1} = \sqrt{4x-1}$. Da un successivo elevamento al quadrato si ottengono le soluzioni 0 e 1, di cui solo la seconda è anche soluzione dell'equazione data.

5.1.7. Equazioni con valori assoluti

La soluzione di equazioni con valori assoluti *non* richiede *nessuna* nuova strategia rispetto a quelle finora considerate: basterà solo ricordare la definizione di valore assoluto data nel paragrafo 4.8.3, nella pagina 120, e distinguere tutti i casi che si possono presentare. Per fare questo occorrerà valutare il segno dell'argomento del valore assoluto: poiché tratteremo in dettaglio questo argomento nel capitolo 6, torneremo su questo argomento nel paragrafo 6.7, nella pagina 170. Per ora ci limitiamo a qualche semplice esempio per chiarire il metodo.

[1] In genere l'elevazione a una potenza dispari non introduce soluzioni estranee.

Esempio 5.23. Risolvere l'equazione $x^2-3|x|+2=0$. Poiché

$$|x|=\begin{cases} -x, & \text{se } x<0 \\ x, & \text{se } x\geq 0 \end{cases},$$

dovremo considerare due casi

1. $x<0$. Allora l'equazione diventa $x^2+3x+2=0$ che ha le soluzioni $x_1=-2$ e $x_2=-1$, entrambe accettabili.
2. $x\geq 0$. Allora l'equazione diventa $x^2-3x+2=0$ che ha le soluzioni $x_3=1$ e $x_4=2$, entrambe accettabili.

L'equazione data ha dunque 4 soluzioni: $S=\{-2,-1,1,2\}$.

Si noti che l'equazione non muta se sostituisco $-x$ al posto di x: dunque se ha una soluzione deve avere anche la sua opposta: questo osservazione avrebbe permesso di considerare uno solo dei due casi (anche se, questa volta, la trattazione di due casi non è poi stata tanto faticosa!).

Si noti altresì che l'equazione $x^2+3|x|+2=0$ non ha invece alcuna soluzione, e questo può essere deciso senza alcun calcolo, in quanto il primo membro non può assumere valori inferiori a 2.

È buona regola, in qualunque tipo di problema, vedere se è possibile trovare qualche scorciatoia prima di partire con le tecniche standard!

Esempio 5.24. Risolvere l'equazione $|x|-|x-1|+1=0$. Poiché

$$|x|=\begin{cases} -x, & \text{se } x<0 \\ x, & \text{se } x\geq 0 \end{cases} \quad \text{e} \quad |x-1|=\begin{cases} -x+1, & \text{se } x<1 \\ x-1, & \text{se } x\geq 1 \end{cases},$$

dovremo considerare 3 casi.

1. $x<0$. Allora l'equazione diventa $(-x)-(-x+1)+1=0$ che si riduce ad una identità: tutti gli $x<0$ sono soluzioni dell'equazione.
2. $0\leq x<1$. Allora l'equazione diventa $x-(-x+1)+1=0$ che ha come soluzione $x=0$.
3. $x\geq 1$. Allora l'equazione diventa $x-(x-1)+1=0$, che non ha soluzioni.

L'insieme delle soluzioni è dunque $S=]-\infty,0]$.

Un caso semplice

Un'equazione del tipo $|f(x)|=a$ può essere trattata in maniera più semplice, sempre tenendo conto della definizione e delle proprietà del valore assoluto. Precisamente:

1. $|f(x)|=a$, con $a<0$, non ha nessuna soluzione;
2. $|f(x)|=0$, equivale a $f(x)=0$;
3. $|f(x)|=a$, con $a>0$, equivale a $f(x)=a\vee f(x)=-a$.

Esempio 5.25. Risolvere l'equazione $|x^2-2x|=1$. Basterà risolvere le due equazioni $x^2-2x=\pm 1$ ottenendo $S=\left\{1-\sqrt{2},1,1+\sqrt{2}\right\}$.

Come già detto, torneremo con maggior dettaglio su questo tipo di equazioni nel prossimo capitolo.

5.2. Equazioni e sistemi in più incognite

Risolvere un'equazione in n incognite reali $x_1, x_2, \dots, x_n$ significa risolvere il seguente problema: date due funzioni, $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}$, trovare il sottoinsieme S di D in cui $f(x_1, x_2, \dots, x_n) = g(x_1, x_2, \dots, x_n)$. Una soluzione di una equazione di questo tipo è dunque una n-upla di reali.

Esempio 5.26. Alcune soluzioni dell'equazione in 2 incognite $x + y - 1 = 0$ sono le coppie $(0, 1)$,$(1, 0)$, $(-1, 2)$, $(3, 2)$,..., ed è chiaro che di coppie soluzioni ce ne sono infinite: basterà fissare arbitrariamente un valore per x e trovare il corrispondente valore per x

Esempio 5.27. Per l'equazione in due incognite $x^2 + y^2 = 0$ solo la coppia $(0, 0)$ è soluzione.

Esempio 5.28. Per l'equazione $x^2 + y^2 + 1 = 0$ non esiste alcuna coppia che sia soluzione.

Come nel caso delle equazioni in un'incognita diremo che l'equazione è in forma normale se g è la funzione nulla. Se f è una funzione razionale parleremo di equazione razionale e se f è addirittura un polinomio, di equazione razionale intera con grado uguale al grado del polinomio.

Nei casi che interessano, le equazioni in più di un'incognita hanno infinite soluzioni ed è particolarmente importante il problema di trovare le soluzioni comuni ad un insieme di equazioni: risolvere un *sistema di equazioni* significa proprio trovare le soluzioni comuni a due o più equazioni, cioè l'intersezione degli insiemi di soluzioni della varie equazioni. Il problema ha, di norma, poco interesse per il caso delle equazioni in un'incognita: esse hanno, nei casi più comuni, un insieme finito di soluzioni e abitualmente, ma non sempre, l'intersezione degli insiemi di soluzioni di più equazioni è vuota.

Esempio 5.29. Risolvere il sistema

$$\begin{cases} x^2 - 1 = 0 \\ x^2 - 3x + 2 = 0 \end{cases} .$$

La prima equazione ha le soluzioni ± 1, la seconda le soluzioni 1 e 2: il sistema ha la sola soluzione $x = 1$.

Nel seguito considereremo solo sistemi di equazioni in più incognite e ci occuperemo in particolare di sistemi di equazioni razionali intere: in questo caso si chiama *grado del sistema* il *prodotto* dei gradi delle singole equazioni. Un sistema di primo grado può dunque essere formato solo da equazioni di primo grado (e in questo caso si chiama anche un *sistema lineare*); un sistema di secondo grado solo da un'equazione di secondo grado e una o più di primo grado, ecc.

Per i sistemi si usa la seguente nomenclatura.

— Un sistema di dice *risolubile* o *compatibile* se ha almeno una soluzione, *irresolubile* o *incompatibile* se non ha alcuna soluzione. A volte si usano anche i termini possibile e impossibile al posto di risolubile e irresolubile.
— Un sistema risolubile si dice *determinato* se ha un numero finito di soluzioni, *indeterminato* se ha infinite soluzioni.
— Due o più sistemi si dicono *equivalenti* se hanno lo stesso insieme di soluzioni.
— Si chiama *combinazione lineare* di un insieme di equazioni di un sistema l'equazione che si ottiene sommando le equazioni dell'insieme ciascuna moltiplicata membro a membro per un numero reale, detto coefficiente della combinazione.

Un sistema di grado n, se è risolubile e determinato, non può avere più di n soluzioni.

Si tenga presente che, se un sistema è ad n incognite, non è detto che tutte le incognite figurino in tutte le equazioni: se in una delle equazioni mancano una o più incognite, si può sempre pensare che esse siano presenti con coefficiente 0. Per questo è bene che sia sempre precisato il numero di incognite di un sistema.

Esempio 5.30. Sia dato il sistema in tre incognite

$$\begin{cases} x+y=0 \\ 2x-y+z=0 \\ y+3z=1 \end{cases} :$$

nella prima e nella terza equazione si può sempre pensare che compaiano anche la z e la x rispettivamente, con coefficiente nullo.

$$\begin{cases} x+y+0z=0 \\ 2x-y+z=0 \\ 0x+y+3z=1 \end{cases}$$

Una delle tecniche più importanti per la risoluzione di un sistema si basa sull'uso delle combinazioni lineari: se a ciascuna equazione di un sistema si sostituisce una combinazione lineare, con almeno un coefficiente non nullo, si ottiene un sistema equivalente a quello dato.

5.2.1. Sistemi lineari

Come già osservato, un sistema lineare è tutto costituito da equazioni di primo grado. Esso si dice in forma normale se in ciascuna delle equazioni le incognite compaiono a primo membro, mentre gli eventuali termini noti compaiono a secondo membro. Il sistema dell'esempio precedente è in forma normale. Esistono tecniche standard per risolvere i sistemi lineari, che saranno studiate in dettaglio nei corsi universitari. Qui accenniamo solo alle più importanti.

Il metodo di sostituzione

Questa tecnica, applicabile, come vedremo, anche a sistemi di grado superiore, prevede di "ricavare" un'incognita da una delle equazioni per poi "sostituirla" nelle altre, ottenendo così un sistema con un minor numero di incognite. Illustriamo il metodo con un esempio.

Esempio 5.31. Per risolvere il sistema in tre incognite

$$\begin{cases} 2x+y-3z=-12 \\ x-2y+z=-1 \\ x+3y+z=9 \end{cases}$$

si può procedere come segue.

$$\begin{cases} y=-2x+3z-12 \\ x-2(-2x+3z-12)+z=-1 \\ x+3(-2x+3z-12)+z=9 \end{cases} \Rightarrow \begin{cases} y=-2x+3z-12 \\ x-z=-5 \\ x-2z=-9 \end{cases} \Rightarrow \begin{cases} y=-2x+3z-12 \\ x=z-5 \\ (z-5)-2z=-9 \end{cases} .$$

Dalla terza equazione dell'ultimo sistema si ricava ora $z=4$ che, sostituito nella seconda, porge $x=-1$; infine, per sostituzione di questi valori nella prima equazione si trova $y=2$. L'unica soluzione di questo sistema è dunque la terna $(-1,2,4)$.

Può naturalmente succedere che il sistema sia indeterminato, come nell'esempio che segue.

Esempio 5.32. Risolvere il sistema in tre incognite

$$\begin{cases} x+2y-3z=-4 \\ 2x-3y+z=-1 \\ 3x-y-2z=-5 \end{cases} \Rightarrow \begin{cases} x=-2y+3z-4 \\ 2(-2y+3z-4)-3y+z=-1 \\ 3(-2y+3z-4)-y-2z=-5 \end{cases} \Rightarrow \begin{cases} x=-2y+3z-4 \\ -y+z=1 \\ -y+z=1 \end{cases}.$$

A questo punto la seconda equazione è identica alla terza che dunque può essere eliminata: il sistema si riduce a due sole equazioni. Dalla seconda si ricava $y=z-1$ che, sostituito nella prima, porge $x=z-2$. Il valore di z rimane indeterminato, nel senso che, qualunque sia z, la terna $(z-2, z-1, z)$ è soluzione del sistema.

Può anche succedere che il sistema sia irresolubile, come mostra l'esempio che segue.

Esempio 5.33. Risolvere il sistema in tre incognite

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

L'ultimo sistema non ha palesemente soluzioni, in quanto le seconda e la terza equazione sono incompatibili.

Si noti che, nell'applicare il metodo di sostituzione conviene valutare accuratamente da quale equazione partire per ricavare un'incognita e quale incognita ricavare, onde evitare inutili calcoli.

Esempio 5.34. Si debba risolvere il sistema in tre incognite.

$$\begin{cases} 2x-4y+4z=1 \\ x+y-2z=1 \\ 3x+2y-4z=3 \end{cases}.$$

Conviene ricavare x (andrebbe bene anche y) dalla seconda equazione e sostituire nelle altre due:

$$\begin{cases} x=1-y+2z \\ -6y+8z=-1 \\ -y+2z=0 \end{cases}.$$

A questo punto conviene ricavare y dalla terza equazione e sostituire prima nella seconda e, successivamente, nella prima. Si ottiene facilmente la terna di soluzioni: $(1, 1/2, 1/4)$.

Non è affatto detto che il numero di incognite debba essere uguale a quello delle equazioni, anche se questo succede nella maggior parte dei casi a livello elementare.

Esempio 5.35. Risolvere il sistema in tre incognite

$$\begin{cases} 3x-y+6z=1 \\ 6x+3y+10z=3 \end{cases}.$$

Ricavando y dalla prima e sostituendo nella seconda si trova

$$\begin{cases} y = 3x + 6z - 1 = 0 \\ 15x + 28z = 6 \end{cases} .$$

A questo punto dalla seconda equazione si può ricavare x in funzione di z e sostituirlo nella prima; si trova

$$x = \frac{6 - 28z}{15}, \quad y = \frac{3 + 6z}{15},$$

mentre il valore di z rimane arbitrario. Il sistema ha dunque le seguenti, infinite, terne di soluzioni:

$$\left(\frac{6-28z}{15}, \frac{3+6z}{15}, z\right), \quad \forall z \in \mathbb{R}.$$

Se si fosse invece ricavato z in funzione di x si sarebbero trovate le seguenti, infinite terne di soluzioni:

$$\left(x, \frac{4-3x}{14}, \frac{6-15x}{28}\right), \quad \forall x \in \mathbb{R}.$$

Anche se le terne sono scritte in forma diversa, i due insiemi di terne coincidono. Per esempio la terna $(2/5, 1/5, 0)$ si ottiene per $z = 0$ dalla prima scrittura e per $x = 2/5$ dalla seconda.

Esempio 5.36. Risolvere il sistema in due incognite:

$$\begin{cases} 7x + 3y = 2 \\ x - 2y = -3 \\ 4x + 9y = 11 \end{cases} .$$

Ricavando $x = -3 + 2y$ dalla seconda equazione e sostituendo nelle altre due si ottiene:

$$\begin{cases} x = -3 + 2y \\ 17y = 23 \\ 17y = 23 \end{cases} ,$$

da cui l'unica coppia soluzione: $(-5/17, 23/17)$.

Il metodo di Cramer

Il metodo di sostituzione, soprattutto per sistemi con poche equazioni, è efficiente e rapido e, a nostro avviso, vale la pena di applicarlo sempre. Accenniamo tuttavia al metodo di Cramer, limitatamente al caso di due equazioni in due incognite, perché diventa importante per sistemi più grandi, come si apprenderà nei successivi corsi universitari, e come vedremo anche nel capitolo 15 della seconda parte. Occorre premettere alcune definizioni.

Definizione 5.6 (Matrice quadrata di ordine due e suo determinante)**.** *Dati quattro numeri reali* a, b, c, d *si chiama* matrice quadrata *di ordine* 2 *la tabella*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{5.8}$$

Si chiama determinante *della matrice* (5.8) *il numero* $ad - bc$*, che si indica con*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}. \tag{5.9}$$

Vale il seguente teorema.

Teorema 5.7 (Regola di Cramer)**.** *Sia dato il sistema lineare di due equazioni in due incognite*

$$\begin{cases} ax + by = \alpha \\ cx + dy = \beta \end{cases},$$

e poniamo

$$D = \begin{vmatrix} a & b \\ c & d \end{vmatrix}, \quad D_x = \begin{vmatrix} \alpha & b \\ \beta & d \end{vmatrix}, \quad D_y = \begin{vmatrix} a & \alpha \\ c & \beta \end{vmatrix}.$$

Allora:

— *se* $D \neq 0$ *il sistema ha la sola soluzione*

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D};$$

— *se* $D = D_x = D_y = 0$ *il sistema è indeterminato;*
— *se* $D = 0$*, ma* $D_x \neq 0$ *oppure* $D_y \neq 0$*, il sistema non ha soluzioni.*

Metodo delle combinazioni lineari

Anche questo metodo diventa di grande importanza per sistemi con un gran numero di equazioni e di incognite e viene formalizzato nella tecnica della "riduzione a scala" di Gauss. Molto grossolanamente si può dire che l'idea è quella di trasformare il sistema in uno equivalente in cui una equazione contenga una sola incognita, un'altra due, e così via. Senza insistere troppo, proponiamo solo un esempio, segnalando che anche questo argomento sarà approfondito nei successivi corsi universitari e sommariamente trattato anche nel capitolo 15 della seconda parte.

Esempio 5.37. Sia dato il sistema in tre incognite

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

Se sostituiamo la seconda equazione con la somma tra la prima e la seconda moltiplicata per -2 e successivamente la terza con la differenza tra la prima e la terza otteniamo il sistema equivalente:

$$\begin{cases} 2x + y - z = 1 \\ -y - 3z = 1 \\ -3z = -1 \end{cases}.$$

Da qui si ricava facilmente la terna di soluzioni $(5/3, -2, 1/3)$.

Esistono anche altre tecniche per risolvere i sistemi lineari, ma non ci pare opportuno insistere ulteriormente: la tecnica di sostituzione, nei casi di poche equazioni, funziona sempre ed è, a nostro avviso, da preferire.

Come già per le equazioni, qualche attenzione in più andrà posta nel caso di sistemi i cui coefficienti siano letterali. Proponiamo un esempio per chiarire il metodo.

Esempio 5.38. Risolvere il sistema nelle incognite[2] x, y, z:

$$\begin{cases} ax + by - 2z = 0 \\ ax + z = 2 \\ 2ax - by = 1 \end{cases} .$$

Se $a = b = 0$ dalla terza equazione si ricava $0 = 1$, dal che si deduce che il sistema non ha soluzioni. Se $a = 0 \land b \neq 0$, dalla seconda equazione si ricava $z = 2$, successivamente dalla prima $y = 4/b$ che non è compatibile con la terza: il sistema non ha soluzioni. Se $a \neq 0 \land b = 0$ dalla terza equazione si ricava $x = 1/2a$, successivamente dalla seconda $z = 3/4$ che non è compatibile con la prima: il sistema non ha soluzioni. Se si suppone $a \neq 0 \land b \neq 0$, si ottiene, partendo dalla seconda equazione e sostituendo nella terza e infine nella prima:

$$\begin{cases} ax = 2 - z \\ by = 2ax - 1 = 3 - 2z \\ 2 - z + 3 - 2z - 2z = 0 \end{cases} .$$

Dunque $z = 1$, $x = 1/a$, $y = 1/b$.

5.2.2. Sistemi di grado superiore al primo

Un sistema di secondo grado deve comprendere una sola equazione di secondo grado e le altre di primo grado. La tecnica normale di soluzione è quella di ricavare una o più incognite dalle equazioni di primo grado per poi sostituirle in quella di secondo che, nei casi di interesse, diventa un'equazione di secondo grado in una incognita, risolubile con la nota formula. Proponiamo un esempio nel caso di due equazioni in due incognite (è il caso di maggior interesse a livello di questo testo).

Esempio 5.39. Risolvere il sistema

$$\begin{cases} x - 2y = 1 \\ x^2 + y^2 = 2 \end{cases} \Rightarrow \begin{cases} x = 2y + 1 \\ (2y+1)^2 + y^2 = 2 \end{cases} \Rightarrow \begin{cases} x = 2y + 1 \\ 5y^2 + 4y - 1 = 0 \end{cases} .$$

Dalla seconda equazione si ricavano due valori per x che sostituiti nella prima porgono due valori per x. Si ottengono due coppie di soluzioni: $(-1, -1)$ e $(7/5, 1/5)$.

Per i sistemi di grado superiore al secondo non esistono tecniche specifiche ed occorre analizzare caso per caso: di solito la loro risoluzione è abbastanza complessa. Faremo solo qualche esempio, senza pretendere di essere esaustivi (cosa che, del resto, sarebbe impossibile).

[2] Abitualmente in un sistema le lettere x, y, x, t sono riservato alle incognite, le prime lettere dell'alfabeto ai cosidetti parametri, ma non è sempre così e sarebbe quindi sempre opportuno precisare quali sono le incognite, in caso di sistemi con coefficienti letterali.

Esempio 5.40. Risolvere il sistema in due incognite

$$\begin{cases} x+y=1 \\ x^3+y^3=7 \end{cases} .$$

Essendo $x^3+y^3=(x+y)(x^2-xy+y^2)$, si può riscrivere il sistema nella forma

$$\begin{cases} x+y=1 \\ x^2-xy+y^2=7 \end{cases} ,$$

che è di secondo grado e si può risolvere per sostituzione. Si può, alternativamente, osservare che $x^2+y^2=(x+y)^2-2xy$; la seconda equazione diventa allora $(x+y)^2-3xy=7$ e, tenendo di nuovo conto della prima $xy=-2$. Si potrebbe di nuovo operare per sostituzione, oppure osservare che, a questo punto, il sistema richiede di trovare due numeri reali che hanno per somma 1 e per prodotto -2. Tenendo conto della formula (5.6) nella pagina 137 si trova subito che si hanno due coppie soluzioni $(-1,2)$ e $(2,-1)$.

Esempio 5.41. Risolvere il sistema in due incognite

$$\begin{cases} x^2+3xy=-2 \\ y^2-xy=3 \end{cases} .$$

Sostituendo alla prima equazione la somma tra la prima e la seconda si ottiene il sistema equivalente

$$\begin{cases} x^2+y^2+2xy=1 \\ y^2-xy=3 \end{cases} \Rightarrow \begin{cases} (x+y)^2=1 \\ y^2-xy=3 \end{cases} \Rightarrow \begin{cases} x+y=\pm 1 \\ y^2-xy=3 \end{cases} .$$

Il sistema si spezza allora in due sistemi di secondo grado risolubili con la tecnica di sostituzione. Si ottengono le 4 coppie di soluzioni:

$$(-2,1),\ (2,-1),\ \left(\frac{1}{2},-\frac{3}{2}\right),\ \left(-\frac{1}{2},\frac{3}{2}\right).$$

Esempio 5.42. Risolvere il sistema in tre incognite:

$$\begin{cases} x^2+y^2+z^2=14 \\ z^2+2xy=-11 \\ x+y=1 \end{cases} .$$

Sottraendo la seconda equazione dalla prima, si ottiene: $x^2+y^2-2xy=25$, ovvero $(x-y)^2=25$, da cui $x-y=\pm 5$. Il sistema si spezza allora nei due sistemi:

$$\begin{cases} x-y=5 \\ x+y=1 \\ z^2+2xy=-11 \end{cases} \quad \text{e} \quad \begin{cases} x-y=-5 \\ x+y=1 \\ z^2+2xy=-11 \end{cases} .$$

La risoluzione è ora immediata; si ottengono le quattro terne di soluzioni

$$(3,-2,1),\ (3,-2,-1),\ (-2,3,1),\ (-2,3,-1).$$

5.3. Esercizi

Esercizio 5.1. *Risolvere l'equazione*

$$ax^2-(a^2+1)x+a=0.$$

Risoluzione. Se $a=0$ l'equazione è di primo grado e ha la sola soluzione $x=0$. Se $a\neq 0$ si ha

$$x_{1,2}=\frac{a^2+1\pm\sqrt{a^4+2a^2+1-4a^2}}{2a}=\frac{a^2+1\pm(a^2-1)}{2a}=\left\langle \begin{array}{l} a \\ 1/a \end{array}\right. .$$

Si noti che $\sqrt{a^4+2a^2+1-4a^2}=\sqrt{(a^2-1)^2}=|a^2-1|$. Essendoci però il doppio segno "$\pm$" davanti alla radice, abbiamo, lecitamente, omesso di indicare il valore assoluto. □

Esercizio 5.2. *Risolvere l'equazione*

$$x^4-x^2-12=0\,.$$

Risoluzione. Posto $x^2=t$ e risolvendo l'equazione di secondo grado che si ottiene, troviamo

$$t_1=4,\ t_2=-3\,.$$

Da $x^2=4$ troviamo $x=\pm2$; da $x^2=-3$ non troviamo invece alcuna soluzione reale. □

Esercizio 5.3. *Risolvere l'equazione*

$$x^3+x^2\sqrt{3}-x\sqrt{2}-\sqrt{6}=0\,.$$

Risoluzione. Il primo membro si può scomporre in fattori:

$$x^3+x^2\sqrt{3}-x\sqrt{2}-\sqrt{6}=x^2\big(x-\sqrt{3}\big)-\sqrt{2}\big(x-\sqrt{3}\big)=\big(x^2-\sqrt{2}\big)\big(x-\sqrt{3}\big).$$

Per la legge di annullamento del prodotto si ha allora:

$$x^2-\sqrt{2}=0\vee x-\sqrt{3}=0\,,$$

da cui

$$x_1=\sqrt[4]{2},\ x_2=-\sqrt[4]{2},\ x_3=\sqrt{3}.$$ □

Esercizio 5.4. *Risolvere l'equazione*

$$2x^5-3x^4-5x^3+5x^2+3x-2=0.$$

Risoluzione. L'equazione (che è anche reciproca di seconda specie di quinto grado) ha le radici razionali ±1, 2, $1/2$. Si può dunque scomporre il polinomio a primo membro mediante successive divisioni, per esempio con lo schema di Ruffini.

	2	−3	−5	5	4	−2
1		2	−1	−6	−1	2
	2	−1	−6	−1	+2	0
−1		−2	3	3	−2	
	2	−3	−3	2	0	
2		4	2	−2		
	2	1	−1	0		
1/2		1	1			
	2	2	0			

L'equazione si riscrive come segue:

$$(x-1)(x+1)(x-2)\left(x-\frac{1}{2}\right)(2x+2)=0\,.$$

Le sue radici sono allora le 4 indicate, e -1 è radice doppia. □

Esercizio 5.5. *Risolvere l'equazione*

$$2\sqrt{x+3}+\sqrt{x}=\sqrt{x+8}.$$

Risoluzione. Quadrando, semplificando e isolando il radicale residuo si ottiene

$$\sqrt{x^2+3x}=-1-x\,.$$

Quadrando di nuovo si ottiene $x=1$, che però non è soluzione dell'equazione data. Non ci sono dunque soluzioni. □

Esercizio 5.6. *Risolvere l'equazione*

$$\sqrt{2x+1}+2\sqrt{x}=\frac{21}{\sqrt{2x+1}}\,.$$

Risoluzione. Riducendo allo stesso numeratore e uguagliando i numeratori si ottiene:

$$\sqrt{(2x+1)^2}+2\sqrt{x(2x+1)}=21\,.$$

Poiché $2x+1$ deve essere strettamente positivo (per l'esistenza del radicale), $\sqrt{(2x+1)^2}=2x+1$. Semplificando si ottiene allora

$$\sqrt{2x^2+x}=10-x.$$

Quadrando e semplificando si ottiene

$$x^2+21x-100=0\,,$$

che ha come radici 4 e -25. Solo 4 soddisfa anche l'equazione di partenza. □

Esercizio 5.7. *Risolvere l'equazione*

$$\sqrt{x+2}+\sqrt{x-3}=\sqrt{3x-4}.$$

Risoluzione. Quadrando e riordinando si ottiene

$$2\sqrt{x^2-x-6}=x+5.$$

Quadrando di nuovo e semplificando si ottiene

$$3x^2-14x+49=0,$$

che ha le radici 7 e $-7/3$. Solo 7 è radice anche dell'equazione di partenza. □

Esercizio 5.8. *Risolvere l'equazione*

$$1+\sqrt{1+x\sqrt{x^2-1}}=x.$$

Risoluzione. Portando 1 a secondo membro, quadrando e riordinando si ottiene

$$x\sqrt{x^2-1}=x^2-2x.$$

Quadrando di nuovo e riordinando si ottiene

$$4x^3-5x^2=0, \quad \text{ovvero} \quad x^2(4x-5)=0,$$

che ha le radici 0 e $5/4$. Nessuna delle due soddisfa l'equazione di partenza, che dunque non ha alcuna soluzione. □

Esercizio 5.9. *Risolvere l'equazione*

$$\frac{\sqrt{x}+\sqrt{x-a}}{\sqrt{x}-\sqrt{x-a}}=\frac{a}{x-a}.$$

Risoluzione. Se si riduce allo stesso denominatore, si uguagliano i numeratori e si riordina, si ottiene,

$$x\sqrt{x-a}=(2a-x)\sqrt{x}.$$

Quadrando e semplificando si ottiene

$$3ax^2-4a^2x=0.$$

Se $a=0$ l'equazione di partenza perde di significato, dunque si può supporre $a\neq 0$. L'ultima equazione ha allora le due radici 0 e $4a/3$.

La radice $x=0$ è accettabile per l'equazione di partenza solo se $a<0$, la radice $x=4a/3$ è invece accettabile solo se $a>0$. □

Esercizio 5.10. *Risolvere il sistema nelle incognite x ed y*

$$\begin{cases} ax-by=2a-b \\ ax+4by=2a+4b \end{cases}.$$

Risoluzione. Se si ricava ax nella prima equazione e si sostituisce nella seconda si ottiene $5by = 5b$. Se $b = 0$ qualunque valore di y è soluzione e dalla prima equazione si ricava $ax = 2a$. Se anche $a = 0$ anche tutti i valori di x sono soluzione. Se invece $a \neq 0$, allora $x = 2$. Se $b \neq 0$ si trova invece $y = 1$ che, sostituito nella prima equazione dà ancora $ax = 2a$. Se $a = 0$ qualunque valore di x è soluzione, se invece $a \neq 0$, allora $x = 2$. Dunque:

— se $a = b = 0$ tutte le coppie (x, y) sono soluzione (come risulta evidente anche dal testo);
— se $a \neq 0 \wedge b = 0$ sono soluzione le coppie $(2, y)$ per ogni $y \in \mathbb{R}$;
— se $a = 0 \wedge b \neq 0$ sono soluzione le coppie $(x, 1)$ per ogni $x \in \mathbb{R}$;
— se $a \neq 0 \wedge b \neq 0$ solo la coppia $(2, 1)$ è soluzione. □

Esercizio 5.11. *Risolvere il sistema*

$$\begin{cases} x^2 + xy + y^2 = 3 \\ x + 2y = 0 \end{cases}$$

Risoluzione. Si ricava $x = -2y$ dalla seconda e si sostituisce nella prima; da qui si ricava facilmente $y = \pm 1$ e quindi $x = \mp = 2$. Dunque si hanno le due coppie soluzione $(2, -1)$ e $(-2, 1)$. □

Esercizio 5.12. *Risolvere il sistema*

$$\begin{cases} x + y - 2z = 0 \\ 5x^2 + 5y^2 = 2 \\ x - y + z = 0 \end{cases}$$

Risoluzione. Conviene ricavare $z = x - y$ dalla terza. Sostituendo nella prima e semplificando si trova $x = 3y$, che può essere sostituito nella seconda, ottenendo un'equazione nella sola y che ha come radici $\pm 1/5$. In corrispondenza di questi due valori per y si trovano i corrispondenti valori per x e z. Si trovano le due terne

$$\left(\frac{3}{5}, \frac{1}{5}, \frac{2}{5}\right), \quad \left(-\frac{3}{5}, -\frac{1}{5}, -\frac{2}{5}\right).$$

□

Esercizio 5.13. *Risolvere il sistema*

$$\begin{cases} x + y = 10 \\ x^3 + y^3 - x^2 - y^2 - 5xy = 207 \end{cases}.$$

Risoluzione. Conviene riscrivere la seconda equazione come

$$(x + y)^3 - 3xy(x + y) - (x + y)^2 + 2xy - 5xy = 207.$$

Se qui si sostituisce $x + y = 10$, si ottiene $xy = 21$. Il sistema con $x + y = 10$ è ora di immediata risoluzione per sostituzione. Si ottengono le coppie $(3, 7)$ e $(7, 3)$. □

6. Disequazioni

6.1. Disequazioni in un'incognita

Risolvere una disequazione[1], nell'incognita reale x, significa risolvere il seguente problema: date due funzioni reali di comune dominio D, $f, g\colon D \to \mathbb{R}$, trovare il sottoinsieme S di D in cui si ha $f(x) > g(x)$, oppure $f(x) \geq g(x)$, oppure $f(x) < g(x)$, o ancora $f(x) \leq g(x)$, . In termini grafici questo problema equivale a determinare i punti sull'asse x in corrispondenza dei quali il grafico di una delle due funzioni sta al di sopra, oppure al di sotto, dell'altro. Se $g(x) = 0$ la disequazione si dice *ridotta in forma normale* e, dal punto di vista dei grafici cartesiani, si riduce a determinare la posizione del grafico della funzione f rispetto all'asse delle x.

La variabile x si chiama *incognita* della disequazione e l'insieme S *insieme delle soluzioni*; ciascuno degli elementi di S si chiama una *soluzione*. Nella scrittura $f(x) \lesseqgtr g(x)$, $f(x)$ si chiama il *primo membro*, $g(x)$ il *secondo membro* della disequazione.

Nei casi pratici, di norma, l'insieme D non è assegnato e si conviene che esso sia il dominio naturale comune delle due funzioni f e g: si parla anche di *dominio della disequazione*. La ricerca del dominio della disequazione deve essere sempre fatta prima di eseguire qualsiasi operazione semplificativa sull'equazione data. Chiariamo questo fatto con un semplice esempio.

Esempio 6.1. Se nella disequazione

$$x^2 + \frac{1}{x} < 1 + \frac{1}{x}$$

si "semplifica" il termine $1/x$, si ottiene la disequazione $x^2 < 1$ che ha come soluzioni l'insieme $]-1, 1[$. La disequazione di partenza invece non può avere come soluzione il numero 0 che non fa parte del dominio.

Le disequazioni più importanti sono quelle che, ridotte in forma normale, hanno a primo membro un polinomio o un quoziente di polinomi (funzione razionale). In proposito si dà la seguente definizione.

Definizione 6.1. *Una disequazione che, ridotta in forma normale, ha a primo membro una funzione razionale si dice una* disequazione razionale*; se il primo membro è, in particolare, un polinomio, si dice* razionale intera *e il grado del polinomio si chiama anche grado dell'equazione.*

6.1.1. Segno delle funzioni

Strettamente connesso con il problema di risolvere una disequazione è quello di determinare il segno di una funzione (reale di variabile reale, per i casi che qui interessano). Data una funzione f si deve:

[1] Si noti che i paragrafi introduttivi sulle disequazioni hanno un contenuto quasi identico a quelli sulle equazioni: si sarebbero potuti trattare i due argomenti contemporaneamente. Abbiamo preferito, seguendo anche la tradizione, fare capitoli separati: su argomenti così importanti vale la pena a volte anche di essere prolissi!

1. trovare il dominio naturale $D \subseteq \mathbb{R}$ (se già non assegnato nel testo);
2. trovare il sottoinsieme degli x di D ove $f(x) = 0$;
3. trovare il sottoinsieme degli x di D ove $f(x) > 0$;
4. trovare il sottoinsieme degli x di D ove $f(x) < 0$.

La determinazione del segno prevede quindi, dopo aver trovato il dominio, la risoluzione di un'equazione e di due disequazioni. Dal punto di vista tecnico, come vedremo, in molti casi la determinazione del dominio e dell'insieme ove $f(x) > 0$ consente di determinare agevolmente anche l'insieme dei punti ove $f(x) = 0$ e quello ove $f(x) < 0$. La situazione non è però sempre così semplice e conviene prestare la massima attenzione. Chiariamo anche questo fatto con un semplice esempio. La funzione $f(x) = x + |x|$, che ha come dominio naturale $\mathbb{R}$, è strettamente positiva in $]0, +\infty[$ e si annulla in $]-\infty, 0]$, mentre non è mai negativa, come si può facilmente verificare; ricordando infatti la definizione di valore assoluto si ha

$$f(x) = x + |x| = \begin{cases} x + x, & \text{se } x \geq 0; \\ x + (-x), & \text{se } x < 0; \end{cases} = \begin{cases} 2x, & \text{se } x \geq 0; \\ 0, & \text{se } x < 0. \end{cases}$$

Il grafico di questa funzione, rappresentato nella figura 6.1, visualizza queste proprietà.

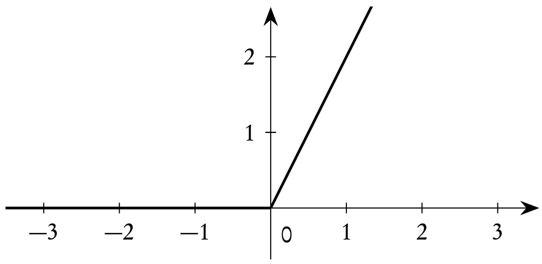


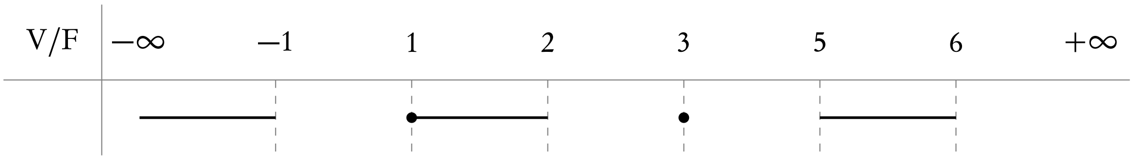
Figura 6.1.: *Grafico della funzione* $f(x) = x + |x|$

Nel seguito, trattando il problema di determinare il segno di una funzione, useremo spesso scritture coinvolgenti la funzione segno, definita nella formula (4.17), nella pagina 126. Si tratta in realtà di un uso leggermente improprio, in quanto la funzione segno assume i valori 0, 1, −1, mentre noi qui siamo interessati al fatto che la funzione in esame sia positiva, negativa o nulla. Tuttavia questo non dà luogo ad equivoci ed è largamente usato in pratica.

6.1.2. Convenzioni grafiche

Avremo spesso bisogno di rappresentare graficamente l'insieme delle soluzioni di una disequazione: nei casi che considereremo questi insiemi saranno costituiti dall'unione di intervalli di $\mathbb{R}$ ed eventualmente da punti isolati. Li rappresenteremo convenzionalmente utilizzando una linea continua per gli intervalli e un "pallino" per rappresentare i punti isolati o per evidenziare il fatto che gli estremi di un intervallo appartengono all'insieme delle soluzioni. Chiariamo tutto questo con un esempio.

Esempio 6.2. Se $S =]-\infty, -1[\cup [1, 2[\cup \{3\} \cup]5, 6[$ è l'insieme delle soluzioni di una certa disequazione, esso sarà rappresentato dal grafico che segue.



Questo tipo di rappresentazione consente di valutare velocemente l'intersezione di più insiemi di soluzioni, cosa indispensabile quando si devono risolvere sistemi di disequazioni. Si noti l'indicazione "V/F" (Vero/Falso) in alto a sinistra a ricordare che in questo grafico si rappresenta l'insieme in cui una disequazione è verificata.

Avremo anche bisogno di rappresentare graficamente i segni di una funzione e il suo dominio. Ci sono diverse tradizioni a questo proposito. Noi useremo i segni "+", "−", "0" per indicare gli intervalli (o i punti) dove una funzione è positiva, negativa, nulla; indicheremo inoltre con una "×" gli intervalli (o i punti) dove una funzione non è definita. Anche in questo caso chiariamo il tutto con un esempio.

Esempio 6.3. Sia f una funzione che

1. è definita per $x < -5 \lor -5 < x \leq 12$;
2. è strettamente positiva per $x < -5 \lor 1 \leq x < 3$;
3. è strettamente negativa per $-5 < x < 1 \lor 8 \leq x \leq 12$;
4. si annulla per $3 \leq x < 8$.

La rappresentazione grafica sarà la seguente.

+/−	$-\infty$		-5		1		3		8		12		$+\infty$
		+	×	−	+	+	0	0	−	−	−	×	

Questo tipo di rappresentazione consente di valutare rapidamente il segno di un prodotto o quoziente di due o più funzioni. Si noti l'indicazione "+/−" in alto a sinistra a ricordare che in questo grafico si rappresentano i segni (e il dominio) di una funzione.

In questi grafici, in entrambi i casi, non hanno alcuna importanza le misure degli intervalli, occorre solo essere certi che i cosiddetti *caposaldi* (i numeri che compaiono sulla prima riga) siano disposti in ordine crescente.

6.1.3. Principi di equivalenza

Definizione 6.2. *Due disequazioni si dicono* equivalenti *se hanno lo stesso insieme di soluzioni.*

La risoluzione di una disequazione richiede in genere una serie di passaggi preliminari atti a ridurre la disequazione stessa a una disequazione equivalente e appartenente ad una delle forme canoniche che saranno esaminate in seguito. Questi passaggi si basano sulle seguenti note proprietà delle disuguaglianze tra numeri:

1. Aggiungendo o sottraendo uno stesso numero ai due membri di una disuguaglianza, si ottiene una disuguaglianza che mantiene lo stesso valore di verità.

2. Moltiplicando o dividendo ambo i membri di una disuguaglianza per uno stesso numero *diverso da zero* si ottiene una disuguaglianza che mantiene lo stesso valore di verità.
3. Due disuguaglianze dello stesso verso possono essere sempre sommate membro a membro.
4. Due disuguaglianze dello stesso verso *non* possono essere moltiplicate membro a membro a meno che *tutti* i membri non siano *strettamente positivi*.

Da qui si ricava il seguente teorema.

Teorema 6.3. *Siano $A(x)$ e $B(x)$ due funzioni definite in un sottoinsieme D di $\mathbb{R}$.*

1. *La disequazione $A(x) < B(x)$ è equivalente alla disequazione $B(x) > A(x)$.*
2. *Qualunque sia la funzione $C(x)$, definita in D, la disequazione $A(x) < B(x)$ è equivalente alla disequazione $A(x) + C(x) < B(x) + C(x)$.*
3. *Qualunque sia la funzione $C(x)$, definita in D e ivi mai nulla, la disequazione $A(x) < B(x)$ è equivalente alla disequazione $A(x)C(x) < B(x)C(x)$ se $C(x) > 0$, alla disequazione $A(x)C(x) > B(x)C(x)$ se $C(x) < 0$.*
4. *Sia f una funzione* strettamente monotona *per cui esistano le funzioni composte $f\big(A(x)\big)$ e $f\big(B(x)\big)$. Allora la disequazione $A(x) < B(x)$ è equivalente alla disequazione $f\big(A(x)\big) < f\big(B(x)\big)$ se f è crescente, alla disequazione $f\big(A(x)\big) > f\big(B(x)\big)$ se f è decrescente.*

Il teorema rimane vero anche se i versi delle disuguaglianze sono invertiti o sostituiti con disuguaglianze in senso lato ($\leq$ o $\geq$ al posto di $<$ o $>$).

Il punto 1) del teorema, anche se può sembrare banale, serve in molte circostanze per evitare inutili passaggi.

Il punto 2) si può esprimere a parole dicendo che si può sempre sommare ad ambo i membri di una disequazione una stessa quantità, oppure "portare una quantità da un membro all'altro" cambiandola di segno (attenzione al dominio!).

Il punto 3) dice che si può moltiplicare (o dividere) ambo i membri di una disequazione per una stessa quantità strettamente positiva, mantenendo il verso; se invece si moltiplica per una quantità strettamente negativa basta cambiare il verso. Si noti che questa quantità può essere una funzione non costante (cioè "può contenere la x"): quello che conta è che abbia segno costante. Per esempio le disequazioni $x^3 + x > 0$ e $x > 0$ sono equivalenti: la prima è ottenuta dalla seconda moltiplicando per $x^2 + 1$ che è sempre strettamente positivo

Il punto 4) è molto importante e ci interesserà in particolare la possibilità di elevare ambo i membri di un'equazione ad una potenza, oppure prendere l'esponenziale, in una certa base, di ambo i membri, e cose simili. Vedremo i dettagli in seguito. Qui limitiamoci ad osservare i seguenti fatti.

1. Non è in genere possibile elevare al quadrato, o a una potenza pari, ambo i membri di una disequazione (la funzione "elevamento al quadrato" non è monotona). Per esempio da $x < -1$, elevando al quadrato si ottiene $x^2 < 1$, che è verificata, come vedremo, per $-1 < x < 1$, in contrasto con la disequazione di partenza. Del resto è ben noto che, mentre per esempio $-4 < 1$ è vera, elevando al quadrato si ottiene $16 < 1$ è falsa. Questo fatto complica abbastanza la risoluzione di disequazioni in cui figurano radicali quadratici. L'elevazione a una potenza pari è invece consentita, se ambo i membri sono positivi: l'elevazione di $A(x) < B(x)$, per esempio, al quadrato, equivale a moltiplicare membro a membro $A(x) < B(x)$ sempre per $A(x) < B(x)$, cosa lecita, come più sopra osservato, se tutti i membri sono positivi.

2. È sempre possibile elevare al cubo, o a una potenza dispari, ambo i membri di una disequazione (la funzione "elevamento al cubo" è monotona).

6.2. Il binomio di primo grado $f(x)=ax+b$

Una disequazione di primo grado in un'incognita, ridotta in forma normale, si può sempre scrivere in una delle forme

$$ax+b \lesseqgtr 0, \quad a, b \in \mathbb{R}, \quad a \neq 0. \tag{6.1}$$

La sua risoluzione è immediata: si porta b a secondo membro e si divide per a, facendo attenzione al suo segno. Vediamo alcuni esempi.

Esempio 6.4. $-2x+5>0 \Rightarrow -2x>-5 \Rightarrow x<\dfrac{5}{2}$.

Esempio 6.5. $\left(\sqrt{2}-\sqrt{3}\right)x-1<0 \Rightarrow x>\dfrac{1}{\sqrt{2}-\sqrt{3}}$.

La determinazione del segno di un binomio di primo grado è immediata, una volta che si sia risolta la disequazione $ax+b>0$ (oppure la disequazione $ax+b<0$). Con riferimento al secondo degli esempi sopra riportati si ha:

$+/-$	$-\infty$		$\dfrac{1}{\sqrt{2}-\sqrt{3}}$		$+\infty$
$(\sqrt{2}-\sqrt{3})x-1$		$+$	0	$-$	

Dal punto di vista grafico, risolvere le disequazioni $ax+b \lesseqgtr 0$ significa determinare l'insieme dei punti in corrispondenza dei quali la retta $y=ax+b$ sta sopra o sotto l'asse delle ascisse (avendo supposto $a \neq 0$, tale retta non può essere parallela all'asse delle ascisse). Per esempio, con riferimento alla prima delle due disequazioni sopra trattate, si ha il grafico della figura 6.2 da cui si ritrova facilmente il risultato già ottenuto per via algebrica.

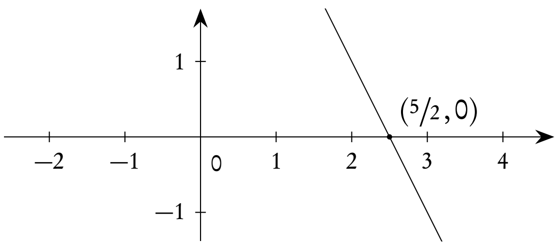


Figura 6.2.: *Grafico della retta* $y=-2x+5$

6.3. Il trinomio di secondo grado $f(x) = ax^2 + bx + c$

Una disequazione di secondo grado in un'incognita, ridotta in forma normale, si può sempre scrivere in una delle forme

$$ax^2 + bx + c \lesseqgtr 0, \quad a, b, c \in \mathbb{R}, \quad a \neq 0. \tag{6.2}$$

Per risolvere una disequazione di questo tipo, o per trovare il segno di un trinomio di secondo grado, *non* conviene imparare a memoria regole relative al segno di a, al segno del discriminante, ecc. Conviene piuttosto ricordare che la funzione a primo membro ha sempre come grafico una parabola con asse parallelo all'asse delle ordinate. La rappresentazione, anche *molto schematica*, di questa parabola, e la determinazione delle sue eventuali intersezioni con l'asse delle ascisse (con la nota formula risolutiva delle equazioni di secondo grado) permette di trarre immediatamente tutte le conclusioni volute. Mostreremo come procedere su alcuni esempi.

Esempio 6.6. Risolvere la disequazione $-2x^2 + 3x + 5 > 0$. Il trinomio a primo membro ha il grafico[2] riportato nella figura 6.3.

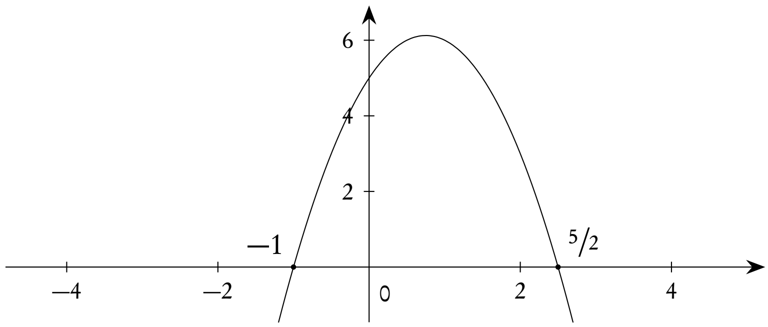


Figura 6.3.: *La parabola* $y = -2x^2 + 3x + 5$

Se ne deduce subito che la disequazione è verificata in $]-1, 5/2[$. Se il problema fosse la determinazione del segno del trinomio a primo membro, dallo stesso grafico potremmo dedurre il risultato seguente.

$+/-$	$-\infty$		-1		$\frac{5}{2}$		$+\infty$
$-2x^2 + 3x + 5$		$-$	0	$+$	0	$-$	

Esempio 6.7. Risolvere la disequazione $x^2 + x + 2 < 0$. Il trinomio a primo membro ha come grafico la parabola della figura 6.4.

Se ne deduce subito che la disequazione non è mai verificata, ovvero che l'insieme di soluzioni è vuoto. Per quanto riguarda il segno del trinomio a primo membro esso risulta sempre strettamente positivo.

[2] In questi grafici in realtà la posizione dell'asse delle ordinate non ha alcun interesse: la riportiamo solo per ragioni di completezza.

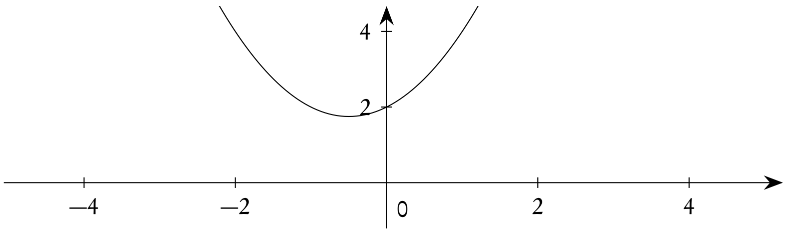


Figura 6.4.: *La parabola* $y = x^2 + x + 2$

+/−	$-\infty$		$+\infty$
$x^2 + x + 2$		+	

Esempio 6.8. Risolvere la disequazione $-x^2 + 2x - 1 < 0$. Tracciamo, nella figura 6.5, il grafico della parabola a primo membro.

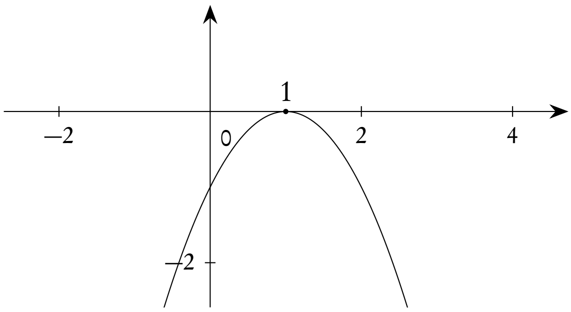


Figura 6.5.: *La parabola* $y = -x^2 + 2x - 1$

Se ne deduce che la disequazione data è verificata per $x \in \mathbb{R} \setminus \{1\}$, mentre il trinomio a primo membro ha il segno seguente.

+/−	$-\infty$		1		$+\infty$
$-x^2 + 2x - 1$		−	0	−	

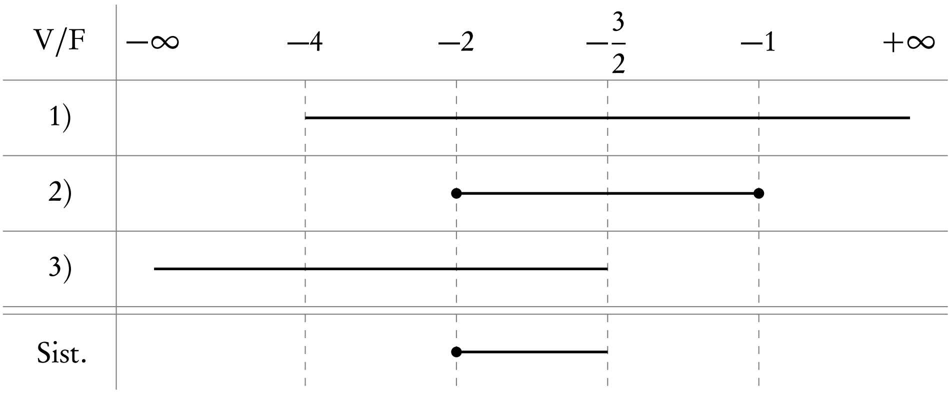
6.4. Sistemi di disequazioni in un'incognita

"Fare sistema" tra due o più disequazioni significa trovare le soluzioni comuni a tutte le disequazioni del sistema, ovvero fare l'intersezione degli insiemi delle soluzioni della varie disequazioni. La rappresentazione grafica che abbiamo adottato torna particolarmente utile e consente di ottenere facilmente l'insieme delle soluzioni del sistema, anche se in casi semplici questa rappresentazione può essere superflua. Utilizziamo un esempio per chiarire il metodo.

Esempio 6.9. Risolvere il sistema

$$\begin{cases} x+4>0 \\ x^2+3x+2\leq 0 \\ 2x+3<0 \end{cases} .$$

Le singole disequazioni sono di immediata risoluzione e si può costruire il seguente schema.



Il sistema è dunque verificato per $x \in [-2, -3/2[$.

6.5. Disequazioni fratte e scomponibili

Una disequazione, ridotta in forma normale, il cui primo membro sia una frazione con numeratore e denominatore scomponibili in fattori di cui si sappia trovare il segno, può essere agevolmente risolta usando la regola dei segni. La rappresentazione grafica del segno dei singoli fattori torna particolarmente utile. Chiariamo il metodo con alcuni esempi.

Esempio 6.10. Risolvere la disequazione $x^3-x^2-2x>0$. La disequazione può essere riscritta nel seguente modo: $x(x+1)(x-2)>0$. Si può agevolmente trovare il segno dei tre fattori e riportarlo nello schema seguente, dove abbiamo indicato anche il segno complessivo del prodotto.

+/−	$-\infty$		-1		0		2		$+\infty$
x		−	−	−	0	+	+	+	
$x+1$		−	0	+	+	+	+	+	
$x-2$		−	−	−	−	−	0	+	
Complessivo		−	0	+	0	−	0	+	

La disequazione è dunque verificata in $]-1,0[\cup]2,+\infty[$.

Esempio 6.11. Risolvere la disequazione

$$\frac{(4x^2-12x+9)x(x^2-3x+2)}{(x^4+1)(x^2+x)} \leq 0.$$

Nella risoluzione di una disequazione sarebbe preferibile trovare prima il dominio: finora non ce ne siamo occupati perché esso coincideva sempre con tutto $\mathbb{R}$. Nel caso di funzioni razionali fratte l'unico problema per il dominio può essere dato dall'eventuale annullarsi del denominatore: se si studia il segno di ciascuno dei fattori e si tiene poi conto che gli eventuali zeri di quelli che stanno al denominatore vanno esclusi, si può anche evitare di trovare il dominio preventivamente[3] e concludere correttamente sulla base della rappresentazione grafica.

In questo caso il primo membro risulta scomposto in fattori dei quali è facile trovare il segno. Si ottiene lo schema che segue.

$+/-$	$-\infty$		-1		0		1		2		$\frac{3}{2}$		$+\infty$
$4x^2-12x+9$		$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	0	$+$	
x		$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$	
x^2-3x+2		$+$	$+$	$+$	$+$	$+$	0	$-$	0	$+$	$+$	$+$	
x^4+1		$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	
x^2+x		$+$	0	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$	
Complessivo		$-$	$\times$	$+$	$\times$	$+$	0	$-$	0	$+$	0	$+$	

Si ha dunque $S =]-\infty, -1[\cup [1,2] \cup \{3/2\}$.

Si noti che nel trovare il segno del primo membro non ha avuto alcuna influenza il verso della disequazione: in un problema come questo, il verso della disequazione va considerato *solo* al termine dell'esercizio, *dopo* aver trovato il segno del primo membro.

6.6. Le funzioni irrazionali

Si chiamano irrazionali quelle funzioni in cui la variabile figura sotto il segno di radice. Per esempio è irrazionale $f(x) = \sqrt{x} + 2$, *non* è irrazionale $f(x) = \sqrt{2}\,x + \sqrt{3}$ (si tratta di un binomio di primo grado). Noi ci occuperemo in questo capitolo solo di funzioni irrazionali il cui radicando sia una funzione

[3]Naturalmente bisogna evitare qualunque semplificazione che possa modificare il dominio. Per esempio se è data la disequazione $x^2+x/x > 0$ e si semplifica il primo membro fino a ottenere $x+1>0$, si conclude con $x>-1$, che è errato, in quanto 0 non fa parte del dominio e va escluso.

razionale, intera o fratta. Tuttavia le tecniche che useremo rimangono valide anche per altri tipi di funzioni irrazionali.

È importante mettere subito in evidenza il fatto che *non* esistono regole generali per risolvere le disequazioni irrazionali o per trovare il segno di funzioni irrazionali. In molti testi sono proposte alcune regole che si applicano in casi semplici, e che riporteremo in seguito solo per ragioni di completezza: tuttavia, a nostro avviso, non vale la pena di memorizzare queste regole o altre simili, proprio perché di applicazione limitata (e anche perché di difficile memorizzazione). Consiglieremo una strategia che funziona sempre, anche se, lo ripetiamo, non è possibile garantire in ogni caso la risolubilità di questo tipo di problemi coinvolgenti le funzioni irrazionali.

6.6.1. Disequazioni irrazionali - la via algebrica

La strategia risolutiva si basa sull'idea di trasformare la disequazione irrazionale in un'altra, ad essa equivalente, e che sia razionale, da risolvere con i metodi già considerati. Per fare questo occorrerà, di norma, elevare ambo i membri ad una opportuna potenza. Purtroppo, come già osservato, l'elevazione a potenza pari non è in genere consentita: bisognerà valutare caso per caso. Inoltre l'elevazione a potenza può complicare il problema anziché risolverlo. Chiariamo questo fatto con un esempio.

Esempio 6.12. Sia data la disequazione $\sqrt{x}-2>x$. Se, dopo avere trovato il dominio (cioè $x\geq 0$) si eleva al quadrato (trascurando, per il momento, i problemi connessi con questa operazione) si ottiene $x-4\sqrt{x}+4>x^2$, che contiene ancora un radicale ed è più complessa della precedente. Se invece si riscrive la disequazione nella forma $\sqrt{x}>x+2$ e poi si eleva al quadrato (ancora trascurando, per il momento, i problemi connessi con questa operazione) si ottiene $x>x^2+4x+4$, che è di secondo grado e quindi di facile risoluzione

La strategia risolutiva standard

Consigliamo di procedere nel seguente modo.

1. Trovare il dominio.
2. Valutare la miglior forma possibile della disequazione in modo che l'elevazione ad una opportuna potenza semplifichi i calcoli.
3. Se si deve elevare ad una potenza dispari procedere direttamente. Se si deve elevare ad una potenza pari, esaminare il segno dei due membri.
   — Se sono entrambi positivi, si può elevare ambo i membri allo stesso esponente pari.
   — Se sono entrambi negativi, cambiare il segno e quindi il verso e poi elevare ambo i membri allo stesso esponente pari.
   — Se sono uno negativo e uno negativo, si può concludere facilmente (un numero positivo è sempre maggiore di una negativo).

Alcuni esempi chiariranno il metodo.

Esempio 6.13. Risolvere la disequazione $\sqrt{x^2-9x+14}-x+8>0$.

1. Dominio: $x\leq 2 \vee x\geq 7$.
2. Chiaramente conviene riscrivere la disequazione nella forma $\sqrt{x^2-9x+14}>x-80$: una elevazione al quadrato trasforma la disequazione in una razionale (di secondo grado).

3. Consideriamo dunque disequazione $\sqrt{x^2-9x+14} > x-8$. Il primo membro, quando ha senso, è sempre maggiore o uguale a 0, il secondo può essere minore di 0 oppure maggiore o uguale a 0: nel primo caso la disequazione è sicuramente vera, nel secondo caso si può elevare al quadrato. Si dovranno dunque distinguere due casi.
   — $x-8<0$, cioè $x<8$. In questo caso il primo membro (non negativo) è sicuramente maggiore del secondo, per cui la disequazione è verificata (all'interno del dominio!).
   — $x-8 \geq 0$, cioè $x \geq 8$. Elevando al quadrato e semplificando si ottiene $7x-50>0$. Si può dunque considerare il sistema
   $$\begin{cases} x \geq 8 \\ 7x-50>0 \end{cases},$$
   che ha come soluzione $x \geq 8$.

   Riunendo le soluzioni dei due casi, e tenendo conto del dominio, si trova $S =]-\infty, 2] \cup [7, +\infty[$.

Esempio 6.14. Risolvere la disequazione $-\sqrt{3-2x}+6+x>0$.

1. Dominio $x \leq 3/2$.
2. Chiaramente non conviene elevare al quadrato direttamente. Si può trasformare in $-\sqrt{3-2x} > -6-x$ o, ancora meglio, in $\sqrt{3-2x} < 6+x$. La seconda forma è da preferire perché in essa il primo membro, quando ha senso, è sempre positivo.
3. Consideriamo dunque la disequazione $\sqrt{3-2x} < 6+x$. Il primo membro, come già osservato, è sempre maggiore o uguale a zero, quando ha senso; il secondo può essere <0, oppure ≥ 0. Se <0 la disequazione non è certamente vera. Si può dunque considerare un unico caso:
   — $6+x \geq 0$. Elevando al quadrato e semplificando si ottiene $x^2+14x+33>0$. Si può considerare il sistema[4]
   $$\begin{cases} x+6 \geq 0 \\ x^2+14x+33>0 \end{cases},$$
   che ha come soluzione $x>-3$.

   Tenendo conto del dominio, si trova l'insieme di soluzioni $S =]-3, 3/2]$.

Esempio 6.15. Risolvere la disequazione $\sqrt{x+\sqrt{x-2}} > \sqrt{x-1}$.

1. Per il dominio si devono imporre le tre condizioni seguenti:
   $$\begin{cases} x-2 \geq 0 \\ x+\sqrt{x-2} \geq 0 \\ x-1 \geq 0 \end{cases}.$$

   Per risolvere questo sistema, anziché partire subito con i calcoli, conviene osservare che dalla prima si ottiene $x>2$; quindi la seconda disequazione è la somma tra un numero ≥ 2 e un numero

[4] Alcuni testi pretendono che si ripeta la condizione trovata per il dominio ad ogni passaggio: anche se la cosa è naturalmente corretta, la riteniamo un'inutile perdita di tempo: basta tenerne conto al momento opportuno, e non scrivere cose errate: in questo esempio la disequazione *non* è equivalente al sistema che abbiamo scritto, ma la risoluzione del sistema, tenendo conto del dominio, consente comunque di risolvere la disequazione. È per questo che abbiamo scritto: "si può considerare il sistema" e non "la disequazione è equivalente al sistema". Ma forse stiamo troppo cercando il pelo nell'uovo!

≥ 0, per cui è sicuramente vera; se poi $x \geq 2$ anche la terza disequazione è vera. Dunque il dominio è $x \geq 2$.
2. Quando i due membri hanno senso essi sono positivi e l'elevazione al quadrato semplifica la disequazione, dunque la forma scritta è la migliore possibile per elevare al quadrato.
3. Trattandosi di una disequazione fra numeri positivi si può elevare al quadrato; semplificando si ottiene: $\sqrt{x-2} > -1$, banalmente vera perché il primo membro, quando ha senso, è non negativo, il secondo è negativo. Si ha dunque $S = [2, +\infty[$.

Esempio 6.16. Risolvere la disequazione $\sqrt{x+2} - \sqrt{x-3} > 1$.

1. Dominio: $x \geq 3$.
2. In qualunque forma si scriva la disequazione una elevazione al quadrato non elimina tutti i radicali, ma in ogni caso dopo l'elevazione ne resta uno solo. Si può però osservare che nella forma assegnata non è immediato constatare che il primo membro è sempre positivo e, in ogni caso, che nel doppio prodotto compaiono termini in x^2, sotto la radice residua. Se si scrive invece nella forma $\sqrt{x+2} > \sqrt{x-3} + 1$, è immediato constatare che i due membri, quando hanno senso, sono non negativi e inoltre, nell'elevazione al quadrato, non compaiono termini di secondo grado. È questa la forma da preferire.
3. Considerata dunque la disequazione $\sqrt{x+2} > \sqrt{x-3} + 1$, quadrando e semplificando si ottiene $\sqrt{x-3} < 7$, che può essere nuovamente risolta quadrando. Si ottiene $x < 7$, tenendo conto del dominio se ne deduce che $S = [3, 7[$.

Esempio 6.17. Risolvere la disequazione $\sqrt[3]{x-2} \leq -x$. In questo caso non ci sono problemi né con il dominio, né con l'elevazione al cubo (che è sempre lecita: x^3 è una funzione strettamente monotona). Si ottiene $x^3 + x - 2 \leq 0$. Scomponendo in fattori il primo membro si ottiene $(x-1)(x^2 + x + 2) \leq 0$. Poiché $x^2 + x + 2$ è strettamente positivo, può essere semplificato ottenendo facilmente $x - 1 \leq 0$ e quindi $S =]-\infty, 1]$.

6.6.2. Disequazioni irrazionali - risoluzione grafica

Per le disequazioni irrazionali la risoluzione per via grafica è particolarmente utile, sia come metodo proprio, sia per controllare i risultati ottenuti per via algebrica (dove non è difficile sbagliare i calcoli). Nei casi che interessano i grafici relativi possono essere sempre ottenuti con i metodi dell'analisi (ma questo esula dagli scopi del presente testo) e, spesso, anche con metodi elementari, anche se è normalmente richiesta la conoscenza della geometria analitica, che tratteremo nel capitolo 8. Solo per un controllo dei risultati[(5)] si può naturalmente usare uno dei numerosi software di calcolo disponibili (anche gratuitamente).

Esempio 6.18. Riprendiamo in esame la disequazione già considerata $\sqrt{x^2 - 9x + 14} - x + 8 > 0$. Il grafico della funzione a primo membro permette di ritrovare subito lo stesso risultato sopra trovato, vedi la figura 6.6.

[5] Fin quando l'uso di questi software non sarà consentito nei test e nelle prove d'esame!

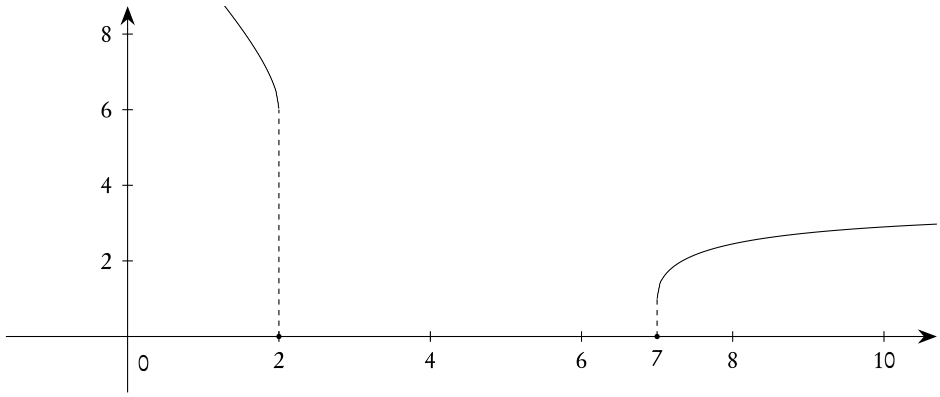


Figura 6.6.: *Grafico di* $f(x) = \sqrt{x^2 - 9x + 14} - x + 8$

Se si vogliono usare metodi elementari, riscriviamo la disequazione nella forma $\sqrt{x^2 - 9x + 14} > x - 8$, possiamo considerare il seguente sistema

$$\begin{cases} y_1 = \sqrt{x^2 - 9x + 14} \\ y_2 = x - 8 \\ y_1 > y_2 \end{cases} .$$

Risolvere la disequazione equivale a trovare per quali x il primo grafico sta sopra il secondo. Il secondo grafico è una retta, tracciabile per via elementare. Per quanto riguarda il primo si può osservare che $y = \sqrt{x^2 - 9x + 14}$ è equivalente a

$$\begin{cases} y \geq 0 \\ y^2 = x^2 - 9x + 14 \end{cases}, \quad \text{ovvero} \quad \begin{cases} y \geq 0 \\ x^2 - y^2 - 9x + 14 = 0 \end{cases}$$

La seconda equazione rappresenta dunque la parte della conica $x^2 - y^2 - 9x + 14 = 0$ che sta sopra l'asse delle ascisse. Si tratta (e non entriamo qui nei dettagli) di un'iperbole equilatera di centro $(9/2, 0)$ e semiassi di lunghezza $5/2$. Il grafico della figura 6.7 fornisce ancora una volta il risultato già noto.

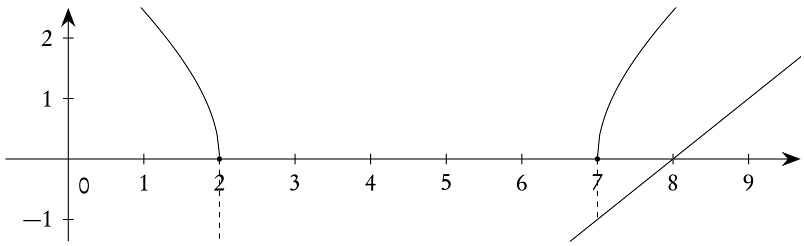


Figura 6.7.: *Grafici di* $y = \sqrt{x^2 - 9x + 14}$ *e* $y = x - 8$

6.6.3. Regole per due casi standard

Riportiamo qui le regole per due casi standard e frequenti nelle applicazioni: le ricaveremo dalla strategia generale indicata e, come già detto, non riteniamo utile una loro memorizzazione indipendente.

La disequazione $\sqrt{f(x)} > g(x)$

Consideriamo una disequazione del tipo

$$\sqrt{f(x)} > g(x), \tag{6.3}$$

e applichiamo la regola generale per le disequazioni irrazionali.

1. Intanto deve essere $f(x) \geq 0$ per il dominio.
2. La forma della disequazione è tale che un'elevazione al quadrato elimina la radice ed è dunque la migliore possibile.
3. Poiché il primo membro, quando ha senso, è sempre non negativo, basterà distinguere due casi, a seconda che il secondo membro sia < 0 o ≥ 0. Nel primo caso la disequazione è sicuramente vera, nel secondo caso si potrà elevare al quadrato. Ripetendo anche la condizione per il dominio, si trova che la risoluzione della disequazione può essere ricondotta all'unione delle soluzioni dei seguenti due sistemi

$$\begin{cases} f(x) \geq 0 \\ g(x) < 0 \end{cases} \cup \begin{cases} f(x) \geq 0 \\ g(x) \geq 0 \\ f(x) > g^2(x) \end{cases}. \tag{6.4}$$

La prima equazione del secondo sistema nella (6.4) è sovrabbondante, in quanto contenuta nella terza. Possiamo concludere che

$$\sqrt{f(x)} > g(x) \quad \Leftrightarrow \quad \begin{cases} f(x) \geq 0 \\ g(x) < 0 \end{cases} \cup \begin{cases} g(x) \geq 0 \\ f(x) > g^2(x) \end{cases}. \tag{6.5}$$

Per la disequazione

$$\sqrt{f(x)} \geq g(x) \tag{6.6}$$

si conclude, senza bisogno di ulteriori considerazioni, che

$$\sqrt{f(x)} \geq g(x) \quad \Leftrightarrow \quad \begin{cases} f(x) \geq 0 \\ g(x) < 0 \end{cases} \cup \begin{cases} g(x) \geq 0 \\ f(x) \geq g^2(x) \end{cases}. \tag{6.7}$$

Si noti che il passaggio dalla disuguaglianza stretta ("$>$") a quella larga ("$\geq$") comporta una variazione solo nella seconda disequazione del secondo sistema.

Può anche essere utile memorizzare le formule (6.5) e (6.7), ma secondo noi è meglio memorizzare la strategia usata per ottenerle, strategia che è quella generale per risolvere le disequazioni irrazionali.

La disequazione $\sqrt{f(x)} < g(x)$

Consideriamo una disequazione del tipo

$$\sqrt{f(x)} < g(x) \tag{6.8}$$

e applichiamo la regola generale per le disequazioni irrazionali.

1. Intanto deve essere $f(x) \geq 0$ per il dominio.
2. La forma della disequazione è tale che un'elevazione al quadrato elimina la radice ed è dunque la migliore possibile.
3. Poiché il primo membro, quando ha senso, è sempre non negativo, basterà distinguere due casi, a seconda che il secondo membro sia < 0 o ≥ 0. Nel primo caso la disequazione è sicuramente falsa, nel secondo caso si potrà elevare al quadrato. Ripetendo anche la condizione per il dominio, si trova che

$$\sqrt{f(x)} < g(x) \quad \Leftrightarrow \quad \begin{cases} f(x) \geq 0 \\ g(x) \geq 0 \\ f(x) < g^2(x) \end{cases} . \tag{6.9}$$

Per la disequazione

$$\sqrt{f(x)} \leq g(x) \tag{6.10}$$

si conclude, senza bisogno di ulteriori considerazioni, che

$$\sqrt{f(x)} \leq g(x) \quad \Leftrightarrow \quad \begin{cases} f(x) \geq 0 \\ g(x) \geq 0 \\ f(x) \leq g^2(x) \end{cases} . \tag{6.11}$$

Come prima, segnaliamo che può anche essere utile memorizzare le formule (6.9) e (6.11), ma secondo noi è meglio memorizzare la strategia usata per ottenerle, strategia che è quella generale per risolvere le disequazioni irrazionali.

6.6.4. Il segno di una funzione irrazionale

Per determinare il segno di una funzione irrazionale si deve seguire il metodo generale indicato per trovare il segno di una funzione. Bisogna prestare particolare attenzione con le funzioni irrazionali: non è affatto detto che la risoluzione della disequazione $f(x) > 0$ sia sufficiente per concludere, come spesso succede nel caso delle funzioni razionali. Per chiarire il problema proponiamo un esempio.

Esempio 6.19. Trovare il segno della funzione $f(x) = x^2 + \sqrt{x^4 - 2x^2 + 1} - 1$. È facile provare che la funzione ha come dominio tutto $\mathbb{R}$: $x^4 - 2x^2 + 1 = (x^2 - 1)^2 \geq 0$. Se si risolve la disequazione $f(x) > 0$ si trova $x < -1 \lor x > 1$. Per concludere bisogna anche risolvere l'equazione $f(x) = 0$ che fornisce $-1 \leq x \leq 1$. Dunque la funzione data ha il segno rappresentato nel seguente grafico.

$+/-$	$-\infty$		-1		1		$+\infty$
$f(x)$		$+$	0	0	0	$+$	

Il grafico di f conferma il risultato trovato, vedi la figura 6.8.

In realtà, questa funzione può essere scritta, facendo uso del valore assoluto, senza utilizzare radicali (e il suo grafico può essere tracciato in via elementare), ma nulla cambia nella sostanza.

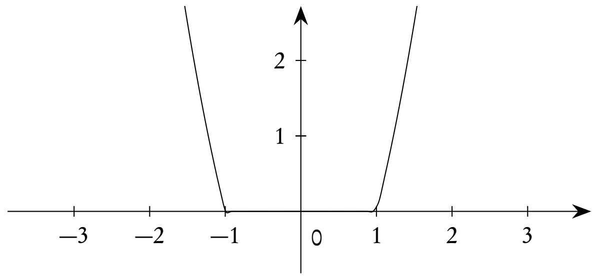


Figura 6.8.: *Grafico di* $f(x) = x^2 + \sqrt{x^4 - 2x^2 + 1} - 1$

6.7. La funzione valore assoluto

La soluzione di disequazioni con valori assoluti e la determinazione del segno di funzioni contenenti valori assoluti *non* richiede *nessuna* nuova strategia rispetto a quelle finora considerate: basterà solo ricordare la definizione di valore assoluto data nella sezione 4.8.3, nella pagina 120 e distinguere tutti i casi che si possono presentare. Per fare questo occorrerrà valutare il segno dell'argomento dei valori assoluti presenti, con le tecniche che abbiamo via via considerato (e con quelle che considereremo in futuro!). Può essere di aiuto anche un grafico del tipo di quelli che abbiamo usato per determinare il segno di un prodotto o quoziente, anche se naturalmente lo scopo sarà diverso! Vediamo un esempio (volutamente un po' complesso!) per chiarire come procedere.

Esempio 6.20. Sia data la funzione $f(x) = |x^2 - 4| + |x + 3| - |x^3 - 3x^2 + 2|$. Si chiede di scrivere la funzione senza l'uso del valore assoluto, distinguendo i vari casi che si possono presentare.

Determiniamo il segno dell'argomento di ciascuno dei tre valori assoluti presenti.

1. $x^2 - 4$. Trattandosi di un polinomio di secondo grado non ci sono problemi. Si ha il grafico seguente.

$+/-$	$-\infty$		-2		2		$+\infty$
$x^2 - 4$		$+$	0	$-$	0	$+$	

2. $x + 2$. Anche in questo caso non ci sono problemi, trattandosi di un binomio di primo grado. Si ha il grafico seguente.

$+/-$	$-\infty$		-3		$+\infty$
$x + 3$		$-$	0	$+$	

3. $x^3 - 3x^2 + 2$. In questo caso si deve scomporre il polinomio di terzo grado: tenendo conto che ha la radice 1, si trova facilmente $x^3 - 3x^2 + 2 = (x - 1)(x^2 - 2x - 2)$. Si può ora trovarne il segno con la nota strategia.

$+/-$	$-\infty$		$1-\sqrt{3}$		1		$1+\sqrt{3}$		$+\infty$
$x-1$		$-$	$-$	$-$	0	$+$	$+$	$+$	
x^2-2x-2		$+$	0	$-$	$-$	$-$	0	$+$	
Complessivo		$-$	0	$+$	0	$-$	0	$+$	

Possiamo ora raggruppare i risultati in un unico grafico, in cui abbiamo messo in alto a sinistra l'indicazione "abs", proprio per segnalare che vogliamo solo "spezzare" i valori assoluti.

abs	$-\infty$		-3		-2		$1-\sqrt{3}$		1		2		$1+\sqrt{3}$		$+\infty$
x^2-4		$+$	$+$	$+$	0	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	
$x+3$		$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	
x^3-3x^2+2		$-$	$-$	$-$	$-$	$-$	0	$+$	0	$-$	$-$	$-$	0	$+$	

Questo grafico ci porta a distinguere i seguenti casi.

1. $-\infty < x < -3$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (x^2-4) + (-x-3) - (-x^3+3x^2-2)$.
2. $-3 \le x < -2$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (x^2-4) + (x+3) - (-x^3+3x^2-2)$.
3. $-2 \le x < 1-\sqrt{3}$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (-x^2+4) + (x+3) - (-x^3+3x^2-2)$.
4. $1-\sqrt{3} \le x < 1$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (-x^2+4) + (x+3) - (x^3-3x^2+2)$.
5. $1 \le x < 2$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (-x^2+4) + (x+3) - (-x^3+3x^2-2)$.
6. $2 \le x < 1+\sqrt{3}$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (x^2-4) + (x+3) - (-x^3+3x^2-2)$.
7. $1+\sqrt{3} \le x < +\infty$. $f(x) = |x^2-4| + |x+3| - |x^3-3x^2+2| = (x^2-4) + (x+3) - (x^3-3x^2+2)$.

Si noti che, nel distinguere i vari casi, abbiamo messo la disuguaglianza larga sempre in modo che si legga $x \ge \dots$; non cambierebbe nulla se scegliessimo la convenzione contraria: una volta sceltane una, però, è bene seguirla.

Si noti ancora che "l'eliminazione del valore assoluto" consiste, in sostanza, nella sua sostituzione con parentesi tonde, racchiudenti o il suo argomento o il suo opposto, a seconda che questi sia positivo o negativo.

Si noti, infine, che i casi 2) e 6) sono identici, come pure il 3) e il 5).

Proponiamo ora alcuni esempi di risoluzione di disequazioni e studio del segno.

Esempio 6.21. Risolvere la disequazione $x + |x^2-1| < 0$. Tenendo conto del segno di x^2-1, si devono distinguere due casi.

1. $x < -1 \vee x \geq 1$. La disequazione diventa $x + (x^2 - 1) < 0$ che ha come soluzioni $-1-\sqrt{5}/2 < x < -1+\sqrt{5}/2$. Dunque in questo primo caso l'insieme delle soluzioni è
$$S_1 = \left]\frac{-1-\sqrt{5}}{2}, -1\right[.$$
2. $-1 \leq x < 1$. La disequazione diventa $x + (-x^2 + 1) < 0$ che ha come soluzioni $x < (1-\sqrt{5})/(2) \vee x > 1-\sqrt{5}/2$. Dunque in questo secondo caso l'insieme delle soluzioni è
$$S_2 = \left[-1, < \frac{1-\sqrt{5}}{2}\right[.$$

Facendo l'*unione* delle soluzioni ottenute nei due casi si trova
$$S = S_1 \cup S_2 = \left]\frac{-1-\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right[.$$

Il grafico della funzione $f(x) = x + |x^2 - 1|$, che si può tracciare in via elementare essendo composto da archi di parabola, conferma il risultato ottenuto: vedi la figura 6.9.

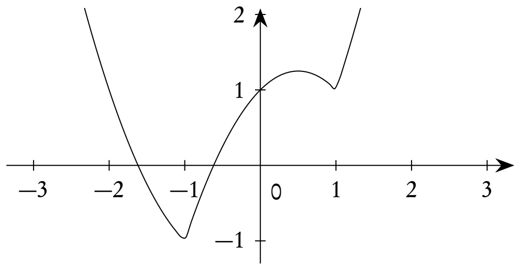


Figura 6.9.: *Il grafico della funzione* $f(x) = x + |x^2 - 1|$

Esempio 6.22. Risolvere la disequazione $\big|1 - |x+2|\big| + x - 1 \leq 0$. Esaminando prima il segno dell'argomento del valore assoluto "più interno", si conclude che si devono esaminare due casi.

1. $x < -2$. La disequazione diventa $|1 - (-x-2)| + x - 1 \leq 0$, ovvero $|x+3| + x - 1 \leq 0$. Si devono di nuovo distinguere due casi.
   a) $x < -3$. La disequazione si riduce a $-4 \leq 0$ che è sempre vera.
   b) $-3 \leq x < -2$. La disequazione si riduce a $2x + 2 \leq 0$, ovvero $x \leq -1$ che, nell'intervallo considerato, è sempre vera.
2. $x \geq -2$. La disequazione si riduce a $|1 - (x+2)| + x - 1 \leq 0$, ovvero $|-x-1| + x - 1 \leq 0$. Si devono di nuovo distinguere due casi.
   a) $-2 \leq x < -1$. La disequazione si riduce a $-2 \leq 0$ che è sempre vera.
   b) $x \geq -1$. La disequazione si riduce a $(x+1) + x - 1 \leq 0$, ovvero $x \leq 0$ che, nell'intervallo considerato, fornisce $-1 \leq x \leq 0$.

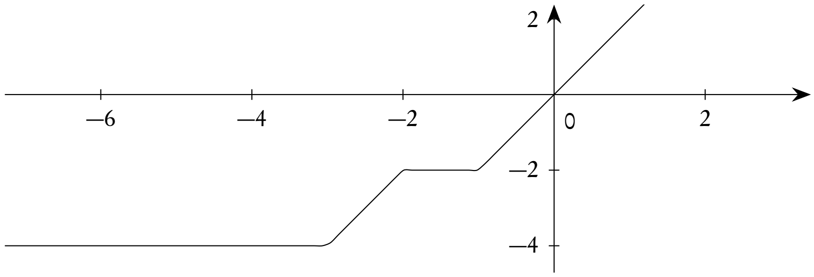


Figura 6.10.: *Grafico della funzione* $f(x) = \big|1 - |x+2|\big| + x - 1$

Facendo l'unione dei risultati trovati nei quattro casi si ottiene $S =]-\infty, 0]$. Il grafico della funzione $f(x) = \big|1-|x+2|\big| + x - 1$, costituito da segmenti e semirette, è tracciabile per via elementare e conferma il risultato trovato: vedi la figura 6.10.

Esempio 6.23. Trovare il segno della funzione $f(x) = x|x| - 2x - 1 + (x+1)|x+1|$. Con considerazioni ormai standard si conclude che si devono distinguere 3 casi.

1. $x < -1$, dove $f(x) = -2x^2 - 4x - 2$.
2. $-1 \le x < 0$, dove $f(x) = 0$.
3. $x \ge 0$, dove $f(x) = 2x^2$.

È ora immediato concludere con il segno riportato nel grafico che segue.

$+/-$	$-\infty$		-1		0		$+\infty$
$f(x)$		$-$	0	0	0	$+$	

Il grafico della funzione $f(x) = x|x| - 2x - 1 + (x+1)|x+1|$, sempre tracciabile per via elementare, conferma il risultato trovato: vedi la figura 6.11.

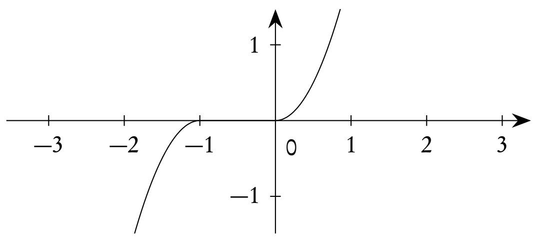


Figura 6.11.: *Il grafico della funzione* $f(x) = x|x| - 2x - 1 + (x+1)|x+1|$

Si noti, cosa del resto già vista in precedenti esempi, come la funzione in questione si annulli su un intero intervallo di numeri reali, a differenza di quanto succede per i polinomi che possono avere al massimo tanti zeri quant'è il loro grado.

6.7.1. Due casi semplici

Le disequazioni del tipo $|f(x)| < a$ (o anche $\leq a$) e $|f(x)| > a$ (o anche $\geq a$) si possono risolvere in maniera più semplice, sempre tenendo conto della definizione e delle proprietà del valore assoluto. Precisamente

1. $|f(x)| < a$ e $|f(x)| \leq a$, con $a < 0$ non hanno alcuna soluzione;
2. $|f(x)| < 0$ non ha nessuna soluzione; $|f(x)| \leq 0$ equivale a $f(x) = 0$;
3. $|f(x)| < a$, con $a > 0$ equivale a $-a < f(x) < a$; $|f(x)| \leq a$, con $a > 0$ equivale a $-a \leq f(x) \leq a$;
4. $|f(x)| > a$ e $|f(x)| \geq a$, con $a < 0$ hanno come soluzioni i punti del dominio di f;
5. $|f(x)| > 0$ equivale a $f(x) \neq 0$; $|f(x)| \geq 0$ ha come soluzioni i punti del dominio di f;
6. $|f(x)| > a$, con $a > 0$ equivale a $f(x) < -a \vee f(x) > a$; $|f(x)| \geq a$, con $a > 0$ equivale a $f(x) \leq -a \vee f(x) \geq a$.

Anche se pare complesso memorizzare tutti questi casi, in realtà, come si può dedurre dall'esempio che segue, basta solo ricordare che il modulo di un numero reale non è mai negativo e che, sulla retta reale, esso rappresenta la distanza del numero stesso dall'origine.

Esempio 6.24. La risoluzione della disequazione $|x-3| < 2$ equivale alla risoluzione del sistema $-2 < x-3 < 2$, ovvero $1 < x < 5$. Infatti dire che il modulo di $x-3$ deve essere minore di 2 equivale a dire che il numero $x-3$ deve distare dall'origine meno di 2, dunque deve essere compreso tra -2 e 2.

6.8. Un esempio complesso

Proponiamo in questo paragrafo un esempio di disequazione contenente funzioni razionali, irrazionali e con valori assoluti, L'esempio è svolto in dettaglio con lo scopo di proporre un modello utile per affrontare le varie situazioni che si possono presentare. È molto importante, in un problema appena un po' complesso, essere molto precisi e chiari e magari fare uno schema di risoluzione. Naturalmente il metodo che proponiamo non è l'unico possibile: quello che conta è costruirsi un metodo e seguire un ordine preciso, senza confusioni.

Il testo dell'esercizio

Risolvere la disequazione

$$\frac{1-x+\sqrt{x^2+x+2}}{3-\left|\dfrac{2x+1}{x}\right|} \geq 0.$$

Ricerca del dominio

La ricerca del dominio, o, meglio, del dominio naturale, deve essere sempre la prima preoccupazione nel risolvere una disequazione: come noto si tratta di trovare il massimo sottoinsieme dei reali dove le operazioni indicate hanno senso. Questo si traduce usualmente nella risoluzione di un sistema di disequazioni in un'incognita. Nel nostro caso le due operazioni che possono essere "soggette a condizioni" sono l'estrazione di radice quadrata (il radicando deve essere non negativo) e la divisione (il denominatore deve essere non nullo). Si ottiene

$$\begin{cases} x^2+x+2 \geq 0 & \text{(radicando non negativo)} \\ 3-\left|\dfrac{2x+1}{x}\right| \neq 0 & \text{(primo denominatore non nullo)} \\ x \neq 0 & \text{(secondo denominatore non nullo)} \end{cases} .$$

La prima disequazione è sempre verificata (il primo membro è una parabola con la concavità verso l'alto e che non interseca mai l'asse delle ascisse). La terza disequazione è banale. Occupiamoci della seconda. Essa può essere risolta con il metodo standard per trattare disequazioni con valori assoluti, valutando il segno della frazione $(2x+1)/x$ e distinguendo i due casi che si possono presentare. Si costruisce facilmente il grafico seguente.

$+/-$	$-\infty$		$-\frac{1}{2}$		0		$+\infty$
$2x+1$		$-$	0	$+$	$+$	$+$	
x		$-$	$-$	$-$	0	$+$	
Complessivo		$+$	0	$-$	$\times$	$+$	

I due casi da considerare sono allora $x < -1/2 \vee x > 0$ e $-1/2 \leq x < 0$. Si conclude con facili calcoli che

$$D = \mathbb{R} \setminus \left\{ -\frac{1}{5}, 0, 1 \right\} .$$

La seconda disequazione poteva anche essere risolta più semplicemente (vedi il paragrafo 6.7.1 nella pagina 174).

Ricerca del segno

Naturalmente terremo conto del dominio già trovato, senza ripetere inutili calcoli.

1. Segno del numeratore. Dovremo risolvere, intanto, la disequazione, $1-x+\sqrt{x^2+x+2} > 0$, che conviene riscrivere nella forma $\sqrt{x^2+x+2} > x-1$. Essendo il primo membro sempre positivo, dovremo considerare solo il segno del secondo membro, il che ci porta a distinguere due casi.
   a) $x \geq 1$. Quadrando e semplificando si ottiene $3x+1 > 0$, cioè $x > -1/3$. Tenendo conto della condizione $x \geq 1$ si conclude che, qui, la disequazione è sempre verificata.
   b) $x < 1$. In questo caso il primo membro è positivo, il secondo è negativo, per cui la disequazione è verificata senza problemi.

   Ne segue che la disequazione è sempre verificata per cui il numeratore risulta sempre maggiore di 0 e quindi non si annullerà mai né, tantomeno, potrà essere negativo.

2. Segno del denominatore. Dobbiamo risolvere la disequazione
$$3-\left|\frac{2x+1}{x}\right|>0.$$
Tenendo conto delle osservazioni già fatte a proposito della determinazione del dominio, dovremo risolvere la disequazione
$$\left|\frac{2x+1}{x}\right|<3 \quad\Leftrightarrow\quad -3<\left|\frac{2x+1}{x}\right|<3.$$
Si ottiene $x \le -1/5 \vee x > 1$. Per questi valori il denominatore è maggiore di 0.I valori che annullano il denominatore sono già stati esclusi dal dominio.

Possiamo concludere con il grafico seguente.

+/−	$-\infty$		$-\frac{1}{5}$		0		1		$+\infty$
Num.		+	+	+	+	+	+	+	
Den.		+	×	−	×	−	×	+	
Complessivo		+	×	−	×	−	×	+	

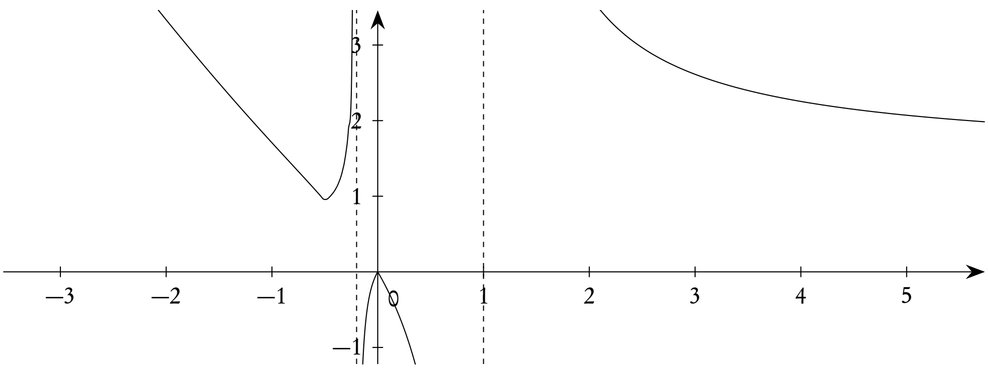


Figura 6.12.: *Grafico della funzione* $f(x)=\left(1-x+\sqrt{x^2+x+2}\right)/(3-|(2x+1)/x|)$

L'insieme delle soluzioni è dunque
$$S=\left]-\infty,-\frac{1}{5}\right[\ \cup\ ]1,+\infty[\,.$$

Il grafico della funzione a primo membro della disequazione fornisce un utile e immediato riscontro ai risultati trovati: vedi la figura 6.12. È chiaro che, se dovessimo tracciare a mano questo grafico, il gioco non varrebbe la candela, ma poiché è possibile usare numerosi software gratuiti in grado di farlo per noi, perché[(6)] non usarli?

[6] Ancora peggio: perché non ne viene consentito l'uso nei temi d'esame, nemmeno alla maturità?

6.9. Esercizi

Esercizio 6.1. *Risolvere la seguente disequazione.*

$$\frac{\left(\sqrt{x-1}-2\right)\left(\sqrt{x-1}-\sqrt{2x-3}\right)}{x-|x|} \geq 0$$

Risoluzione. Troviamo innanzitutto il dominio. Si deve considerare il sistema

$$\begin{cases} x-1 \geq 0 \\ 2x-3 \geq 0 \\ x-|x| \neq 0 \end{cases}.$$

Poiché $x-|x| = x-(-x) = 2x$, se $x<0$ e $x-|x| = x-(x) = 0$, se $x>0$, la terza condizione porta alla conclusione $x<0$. Tenendo conto anche della prima e seconda si conclude (senza bisogno di grafici!) che il primo dominio è vuoto: non ci sono soluzioni della disequazione. □

Esercizio 6.2. *Risolvere la seguente disequazione.*

$$\frac{\left(\sqrt{x-1}-2\right)\left(\sqrt{x-1}-\sqrt{2x-3}\right)}{x+|x|} \geq 0$$

Risoluzione. Si tratta di una piccola variante dell'esercizio 6.1. Per trovare il dominio il sistema diventa ora

$$\begin{cases} x-1 \geq 0 \\ 2x-3 \geq 0 \\ x+|x| \neq 0 \end{cases}.$$

Poiché $x+|x| = x+(-x) = 0$, se $x<0$ e $x+|x| = x+(x) = 2x$, se $x>0$, la terza condizione porta alla conclusione $x>0$. Tenendo conto anche della prima e seconda si conclude (senza bisogno di grafici!) che il dominio è $x \geq 1$. La discussione fatta sul denominatore porta a concludere che esso risulta, nel dominio trovato, sempre positivo e dunque può essere eliminato (moltiplicando ambo i membri per $x+|x|$).

Troviamo allora il segno dei due fattori del numeratore. Per il primo si deve risolvere la disequazione $\sqrt{x-1}-2>0$, ovvero $\sqrt{x-1}>2$. Sempre tenendo conto del dominio trovato, si può tranquillamente elevare al quadrato (i due membri sono non negativi), ottenendo $x>5$. Se ne conclude facilmente che questo primo fattore sarà negativo per $1 \leq x < 5$, nullo per $x=5$, positivo per $x>5$. Per trovare il segno del secondo risolviamo la disequazione $\sqrt{x-1}-\sqrt{2x-3}>0$, che può essere riscritta come $\sqrt{x-1}>\sqrt{2x-3}$. Anche qui si può tranquillamente elevare al quadrato (sempre tenendo conto del dominio!), ottenendo $x<2$. Se ne conclude che questo secondo fattore è positivo per $1 \leq x < 2$, nullo per $x=2$, negativo per $x>2$. Riportiamo i risultati nel solito grafico (anche se in un caso così semplice si potrebbe farne a meno).

$+/-$	$-\infty$		1		2		5		$+\infty$
1)		$\times$	$-$	$-$	$-$	$-$	0	$+$	
2)		$\times$	$+$	$+$	0	$-$	$-$	$-$	
Complessivo		$\times$	$-$	$-$	0	$+$	0	$-$	

La disequazione è dunque verificata per $2 \leq x \leq 5$, ovvero si ha $S = [2,5]$.

□

Esercizio 6.3. *Risolvere la disequazione*

$$\frac{\left(\sqrt{x^2-2x+3}-x+1\right)\left(\sqrt[3]{x^3+26x}-x-2\right)}{x^2+|x-1|-2} \geq 0.$$

Risoluzione. Determiniamo innanzitutto il dominio. Si deve avere

$$\begin{cases} x^2-2x+3 \geq 0 \\ x^2+|x-1|-2 \neq 0 \end{cases}.$$

La prima disequazione è sempre verificata (parabola con concavità verso l'alto e che non interseca l'asse delle ascisse). Per risolvere la seconda distinguiamo due casi:

1. $x < 1$. La disequazione diventa $x^2+(-x+1)-2 \neq 0$, ovvero $x^2-x-1 \neq 0$, che comporta $x \neq (1-\sqrt{5})/2$.
2. $x \geq 1$. La disequazione diventa $x^2+(x-1)-2 \neq 0$, ovvero $x^2+x-3 \neq 0$, che comporta $x \neq (-1+\sqrt{13})/2$. Dunque

$$D = \mathbb{R} \setminus \left\{ \frac{1-\sqrt{5}}{2}, \frac{-1+\sqrt{13}}{2} \right\}.$$

Troviamo poi il segno dei due fattori del numeratore e di quello del denominatore (senza nuovamente preoccuparci del dominio che abbiamo già trovato).

1. $\sqrt{x^2-2x+3}-x+1$. Risolviamo la disequazione $\sqrt{x^2-2x+3}-x+1 > 0$, cioè $\sqrt{x^2-2x+3} > x-1$. Se $x < 1$ essa è sicuramente vera; se $x \geq 1$ si può quadrare e la disequazione si riduce a $3 > 1$ che è sicuramente vera. Questo fattore è dunque strettamente positivo e può essere semplificato.
2. $\sqrt[3]{x^3+26x}-x-2$. Risolviamo la disequazione $\sqrt[3]{x^3+26x}-x-2 > 0$, ovvero $\sqrt[3]{x^3+26x} > x+2$. Si può tranquillamente elevare al cubo: dopo semplificazione si ottiene $3x^2-7x+4 < 0$, che è verificata per $1 < x < 4/3$. Questo fattore è quindi positivo per $1 < x < 4/3$, si annulla per $x = 1 \vee x = 4/3$, è negativo per $x < 1 \vee x > 4/3$.
3. $x^2+|x-1|-2$. Tenendo conto dei calcoli già fatti per il dominio, concludiamo che il denominatore è positivo per $x < (1-\sqrt{5})/2 \vee x > (-1+\sqrt{13})/2$, negativo tra questi due valori.

Riportiamo i risultati sul solito grafico.

+/−	$-\infty$		$\frac{1-\sqrt{5}}{2}$		1		$\frac{-1+\sqrt{13}}{2}$		$\frac{4}{3}$		$+\infty$
1)		$-$	$\times$	$-$	0	$+$	$\times$	$+$	0	$-$	
2)		$+$	$\times$	$-$	$-$	$-$	$\times$	$+$	$+$	$+$	
Complessivo		$-$	$\times$	$+$	0	$-$	$\times$	$+$	0	$-$	

Si conclude che

$$S = \left]\frac{1-\sqrt{5}}{2}, 1\right] \cup \left]\frac{-1+\sqrt{13}}{2}, \frac{4}{3}\right].$$

Non avendo a disposizione una calcolatrice, non è immediato concludere che $4/3$ è maggiore di $(-1+\sqrt{13})/(2)$. Si può seguire il seguente ragionamento, di carattere generale. Supponiamo che $4/3 < (-1+\sqrt{13})(2)$ allora

$$\frac{4}{3} < \frac{-1+\sqrt{13}}{2} \Leftrightarrow 8 < -3+3\sqrt{13} \Leftrightarrow 11 < 3\sqrt{13} \Leftrightarrow 121 < 117.$$

Poiché l'ultima disuguaglianza è falsa , se ne conclude che è falsa anche la prima, cioè che

$$\frac{4}{3} > \frac{-1+\sqrt{13}}{2}.$$

□

Esercizio 6.4. *Risolvere la disequazione*

$$|5-x| < |2x-3|.$$

Risoluzione. Invece di seguire la solita strategia, distinguendo i vari casi che si possono presentare, osserviamo che i due membri non sono mai negativi, dunque si può elevare al quadrato senza problemi, tenendo anche conto del fatto che $|t|^2 = t^2$. Si ottiene, dopo semplificazioni,

$$3x^2 - 2x - 16 > 0,$$

il cui insieme di soluzioni è

$$S =]-\infty, -2[\cup \left]\frac{8}{3}, +\infty\right[.$$

Il tutto funziona, naturalmente, perché elevando al quadrato non si ottengono potenze superiore al 2 per l'incognita. □

Esercizio 6.5. *Risolvere la disequazione*

$$\big||x|-2\big| > 3.$$

Risoluzione. Anche qui si può procedere come fatto con l'esercizio 6.4, elevando al quadrato ambo i membri. Si ottiene, dopo semplificazioni,

$$x^2 - 4|x| - 5 > 0,$$

che richiede due casi.

1. Se $x < 0$ si ha $x^2 + 4x - 5 > 0$, da cui $x < -5$.
2. Se $x \geq 0$ si ha $x^2 - 4x - 5 > 0$, da cui $x > 5$.

Dunque $S =]-\infty, -5[\cup]5, +\infty[$. Si poteva anche osservare che la disequazione ottenuta dopo elevamento al quadrato era simmetrica rispetto al cambio di x con $-x$, dunque una volta risolta per $x \geq 0$, basta considerare le soluzioni simmetriche rispetto all'origine.

Presentiamo anche la risoluzione ottenuta con metodo standard. Cominciano dal valore assoluto più interno, si devono distinguere due casi.

1. $x < 0$, da cui $|-x-2| > 3$, che richiede ancora due casi.
   a) $x \geq -2$, da cui $x + 2 > 3$, che non ha soluzioni nell'intervallo in considerazione.
   b) $x < -2$, da cui $-x - 2 > 3$, ovvero $x < -5$.
2. $x \geq 0$, da cui $|x-2| > 3$, che richiede ancora due casi.
   a) $x < 2$, da cui $-x + 2 > 3$, che non ha soluzioni nell'intervallo considerato.
   b) $x \geq 2$, da cui $x - 2 > 3$, ovvero $x > 5$.

Si ritrovano gli stessi risultati di prima.

Si poteva ancora osservare che, scambiando x con $-x$ la disequazione non cambia. Basta dunque risolvera per $x \geq 0$, e poi considerare le soluzioni simmetriche rispetto all'origine.

Come si vede, anche se la strategia standard produce sempre risultati, nel risolvere un esercizio si possono spesso individuare strategie semplificative. □

Esercizio 6.6. *Trovare il dominio della seguente funzione*

$$f(x) = \sqrt{\frac{|x+2| - |x-1|}{1 - \sqrt[3]{x^2 - 1}}}.$$

Risoluzione. Si tratta di risolvere il sistema

$$\begin{cases} \dfrac{|x+2| - |x-1|}{1 - \sqrt[3]{x^2-1}} \geq 0 \\ 1 - \sqrt[3]{x^2-1} \neq 0 \end{cases}.$$

Dalla seconda, scrivendola nella forma $1 \neq \sqrt[3]{x^2-1}$ ed elevando al cubo, si trova $x \neq \pm\sqrt{2}$. Per la prima troviamo il segno di numeratore e denominatore,

1. Numeratore. Risolviamo $|x+2| - |x-1| > 0$, ovvero $|x+2| > |x-1|$. Procedendo come nell'esercizio 6.4, si trova facilmente $x > -1/2$.

2. Denominatore. Risolviamo $1-\sqrt[3]{x^2-1}>0$, ovvero $1>\sqrt[3]{x^2-1}$. Elevando al cubo e semplificando si trova facilmente $-\sqrt{2}<x<\sqrt{2}$.

Si può ora costruire il solito grafico.

+/−	$-\infty$		$-\sqrt{2}$		$-\frac{1}{2}$		$\sqrt{2}$		$+\infty$
Num.		−	×	−	0	+	×	+	
Den.		−	×	+	+	+	×	−	
Complessivo		+	×	−	0	+	×	−	

Il dominio richiesto è dunque

$$S=\left]-\infty,-\sqrt{2}\right[\cup\left[-\frac{1}{2},\sqrt{2}\right[. \qquad \square$$

Esercizio 6.7. *Risolvere la disequazione*

$$\sqrt{|2x+1|-1}\geq x-3.$$

Risoluzione. Troviamo innanzitutto il dominio. Si deve avere $|2x+1|-1\geq 0$, ovvero $|2x+1|\geq 1$. Trattandosi di una disequazione di tipo elementare si ottiene $2x+1\leq -1 \vee 2x+1\geq 1$, da cui $x\leq -1 \vee x\geq 0$.

A questo punto dobbiamo elevare al quadrato, e la forma proposta è la migliore possibile. Ci sono due casi.

1. $x<3$. Il primo membro è positivo, il secondo negativo: la disequazione è sempre verificata (nel dominio trovato!).
2. $x\geq 3$. Ambo i membri sono positivi, per cui l'elevazione al quadrato non dà problemi. Inoltre, se $x\geq 3$, $|2x+1|=2x+1$. Quadrando e semplificando si trova $3\leq x<4+\sqrt{7}$.

Si ha dunque

$$S=]-\infty,-1]\cup\left[0,4+\sqrt{7}\right]. \qquad \square$$

7. Potenze, esponenziali, logaritmi

Le definizioni e proprietà delle potenze sono già state definite nei capitoli precedenti e sono qui riportate per ragioni di completezza e per chiarire il perché di certe definizioni.

7.1. Potenze con esponente intero

7.1.1. Esponente intero ≥ 2

Se $a \in \mathbb{R}$ e $m \in \mathbb{N}$, $m \geq 2$, si pone, per definizione

$$a^m = \underbrace{a \cdot a \dots a}_{m \text{ volte}}. \tag{7.1}$$

Il numero reale a^m si chiama *potenza* di *base* a ed *esponente* m. In sostanza questa definizione appare, per ora, solo come una scrittura abbreviata di un prodotto con tanti fattori tutti tra di loro uguali.

Valgono le seguenti proprietà di immediata verifica, dette *proprietà formali delle potenze*, dove a e b sono reali (diversi da zero se appaiono al denominatore) ed m, n naturali ≥ 2.

$$a^m \cdot a^n = a^{m+n}; \quad (a \cdot b)^m = a^m \cdot b^m; \tag{7.2a}$$

$$\frac{a^m}{a^n} = a^{m-n}, \quad \text{se} \quad m \geq n+2; \quad \left(\frac{a}{b}\right)^m = \frac{a^m}{b^m}; \tag{7.2b}$$

$$(a^m)^n = a^{mn}. \tag{7.2c}$$

7.1.2. Esponente 0 ed esponente 1

La definizione data si estende, senza alcuna difficoltà, al caso che l'esponente sia 1, ponendo, per definizione

$$a^1 = a. \tag{7.3}$$

Con una certa cautela è possibile estendere la definizione di potenza anche al caso che l'esponente sia 0, occorre solo introdurre una limitazione per le basi possibili, escludendo lo zero. Precisamente si ha, per definizione,

$$a^0 = 1, \quad a \in \mathbb{R}, a \neq 0. \tag{7.4}$$

È lecito chiedersi il perché delle scelte effettuate con le definizioni (7.3) e (7.4). Se si può ritenere, in un certo senso, "naturale" la definizione (7.3), altrettanto non può dirsi della (7.4).

Alcuni pretendono di "dimostrare" che, se $x \in \mathbb{R}$, $x \neq 0$ allora $x^0 = 1$. Il ragionamento è, all'incirca, questo: se $x \neq 0$ e $m \in \mathbb{N}$, $m \geq 2$, allora

$$1 = \frac{x^m}{x^m} = x^{m-m} = x^0 . \tag{7.5}$$

Questo ragionamento contiene un grave errore logico: il simbolo x^0 non è stato definito e non si può pretendere che rappresenti alcunché, finché non si sa che cosa sia. È ben vero che $x^m/x^m = 1$, se $x \neq 0$, ma non si può pretendere che sia anche uguale a x^0, visto che non si sa cosa sia x^0. Il discorso per x^1 è simile.

Questo tipo di argomentazione, tuttavia, è importante, *non* perché dimostra che $x^1 = x$ oppure $x^0 = 1$, ma perché dimostra che, se si vogliono estendere le proprietà formali (7.2) delle potenze, è indispensabile dare le definizioni (7.3) e (7.4). È proprio questa la linea che si sceglie per estendere la definizione di potenza, in modo da consentire la scelta degli esponenti in insiemi via via più ampi, magari con qualche sacrificio per la possibilità di scelta delle basi: si dà una definizione tale che anche con i nuovi esponenti le proprietà formali delle potenze continuino ad essere valide. Si giunge così, infine, ad un concetto di potenza in cui il punto di partenza (prodotto di fattori tutti uguali) è poco più che un lontano ricordo[1].

7.1.3. Esponente intero negativo

Proseguendo con la strategia che ormai abbiamo adottato, se vogliamo che $a^{m+n} = a^m \cdot a^n$ continui a rimanere valida anche con interi negativi, dovrà essere

$$1 = a^0 = a^{m-n} = a^m \cdot a^{-n}, \tag{7.6}$$

da cui

$$a^{-n} = \frac{1}{a^n}, \quad a \in \mathbb{R}, a \neq 0, \quad n \in \mathbb{N}, n > 0. \tag{7.7}$$

Non è difficile provare che anche le altre proprietà formali delle potenze rimangono valide.

7.2. Potenze con esponente razionale

Cominciamo con la seguente osservazione di grande importanza, già in parte anticipata in un esempio nella pagina 70.

Osservazione 7.1. Non è possibile definire la potenza con esponente razionale e base negativa in modo da conservare le proprietà formali delle potenze e, contemporaneamente, mantenere la compatibilità con il caso degli esponenti interi. Sappiamo infatti che, qualunque sia il significato che vogliamo attribuire al

[1] L'argomentazione contenuta nella (7.5) non è l'unica possibile. Si può anche osservare che

$$a^n = a^{n+0} = a^n \cdot a^0,$$

da cui $a^0 = 1$, se $a \neq 0$. Ancora una volta siamo "costretti" a definire $a^0 = 1$, se vogliamo mantenere la proprietà $a^{m+n} = a^m \cdot a^n$.

simbolo $(-2)^{1/2}$, il suo quadrato deve essere positivo. Se ora vogliamo che rimangano valide le proprietà delle potenze e le definizioni che abbiamo già dato per le potenze ad esponente intero, dobbiamo avere

$$-2 = (-2)^1 = (-2)^{2/2} = (-2)^{1/2\times 2} = \left((-2)^{1/2}\right)^2 > 0\,,$$

cosa palesemente assurda.

Nel caso di base nulla è chiaro che porremo $0^r = 0$, per ogni numero razionale $r > 0$, mentre, come già fatto con gli esponenti interi, non daremo alcun significato al simbolo 0^r con $r \leq 0$. Ci resta da vedere come possiamo dare significato alle potenze con esponente razionale e base > 0, applicando di nuovo la strategia già usata con profitto per esponenti interi minori di 2. Poiché deve essere, per $a > 0$,

$$a = a^1 = a^{n/n} = a^{(1/n)\times n} = \left(a^{1/n}\right)^n\,, \tag{7.8}$$

utilizzando la definizione di radicale aritmetico dovremo porre, per definizione,

$$a^{1/n} = \sqrt[n]{a}\,. \tag{7.9}$$

Sempre per le proprietà delle potenze dovremo poi avere, sempre per $a > 0$,

$$a^{m/n} = \left(a^{1/n}\right)^m = \left(\sqrt[n]{a}\right)^m = \sqrt[n]{a^m}. \tag{7.10}$$

Tenendo conto delle proprietà dei radicali aritmetici concludiamo che

$$m/n = p/n \quad \Rightarrow \quad a^{m/n} = a^{p/q}, \tag{7.11}$$

come *deve* essere.

Non è difficile provare, tenendo conto delle proprietà dei radicali, che la definizione (7.10) soddisfa tutte le proprietà formali delle potenze.

Osservazione 7.2. La convenzione che le potenze ad esponente razionale siano definite solo con basi positive non è universalmente accettata: anche molte delle calcolatrici in commercio non forniscono indicazioni di errore con calcoli del tipo $(-1)^{1/3}$, mentre non accettano calcoli del tipo $(-1)^{0.333\ldots}$, anche se, a livello dei numeri utilizzati dalla calcolatrice i numeri $1/3$ e $0.333\ldots$ sono considerati identici. In effetti si potrebbe dare una buona definizione di potenza con base negativa ed esponente razionale, almeno per certi esponenti, ma si perderebbero in ogni caso le proprietà formali e i vantaggi che si otterrebbero sarebbero di gran lunga compensati dagli svantaggi: meglio evitare rischi.

Per concludere questa discussione sulle potenze ad esponente razionale, proponiamo un esempio da cui si deduce che, in ogni caso, le potenze con base negativa vanno trattate con la massima cautela.

Esempio 7.1. Che legame c'è tra

$$\left((-1)^{\frac{3}{2}}\right)^2, \quad \left((-1)^2\right)^{\frac{3}{2}}, \quad (-1)^3\,?$$

Se le potenze con base negativa fossero sempre definite e valessero le proprietà formali, le tre espressioni dovrebbero avere lo stesso valore. Invece

1. $\left((-1)^{\frac{3}{2}}\right)^2$ non è definito, perché -1 non può essere elevato a $3/2$;
2. $\left((-1)^2\right)^{\frac{3}{2}}$ vale 1, in quanto si *deve* calcolare prima $(-1)^2 = 1$ e poi elevare a $3/2$, ottenendo appunto 1;
3. $(-1)^3 = -1$, come è ovvio.

7.3. Potenze con esponente reale irrazionale

L'ultima estensione del concetto di potenza a cui siamo interessati è quella di considerare esponenti reali anche irrazionali. Purtroppo questo problema è estremamente complesso e la sua trattazione rigorosa (per la quale esistono diverse strategie) esula dagli scopi di questo testo. Ci limiteremo solo a dare un'idea "intuitiva", che comunque chiarisce come si possa procedere.

Dato $a > 0$, $a \in \mathbb{R}$, e $\gamma \in \mathbb{R}$ γ irrazionale, si consideri la rappresentazione decimale di γ, che sarà illimitata e non periodica: $\gamma = \alpha, \beta_1\beta_2\beta_3\dots$. Si consideri la successione

$$a^{\alpha},\, a^{\alpha,\beta_1},\, a^{\alpha,\beta_1\beta_2},\, a^{\alpha,\beta_1\beta_2\beta_3},\, \dots : \tag{7.12}$$

si tratta di una successione di numeri reali positivi, crescente se $a > 1$, decrescente se $0 < a < 1$, costantemente uguale a 1 se $a = 1$. Ebbene, man mano che l'esponente "si avvicina" al numero γ, questa successione numerica "si avvicina" ad un numero reale ben definito che si chiama *potenza di base a ed esponente* γ. Per completare la definizione si pone poi, se $\gamma > 0$, $0^{\gamma} = 0$. Si dimostra poi che valgono le proprietà formali delle potenze (dimostrazione non banale!).

Si noti che, come già annunciato, questo "concetto esteso di potenza" ha ormai ben poco a che fare con il primitivo concetto di prodotto di fattori tra di loro uguali.

7.4. Considerazioni conclusive sulle potenze

A conclusione di questa trattazione del concetto di potenza proponiamo, perché molto utile nelle applicazioni un quadro riassuntivo della portata delle definizioni date.

Il simbolo a^{γ} ha senso, nell'ambito reale:

1. per ogni $a \in \mathbb{R}$, se γ è un intero > 0;
2. per tutti gli $a \in \mathbb{R}$ escluso lo zero, se γ è un intero ≤ 0;
3. per $a \geq 0$, se γ è non intero e > 0;
4. per $a > 0$, se γ è non intero e < 0.

È molto importante nelle applicazioni saper confrontare le potenze. I problemi che si possono presentare sono di tre tipi:

1. confronto tra potenze con la stessa base;
2. confronto tra potenze con lo stesso esponente;
3. confronto tra potenze con base ed esponenti diversi.

È abbastanza facile risolvere i problemi dei primi due tipi; i problemi del terzo tipo sono invece molto difficili e vanno trattati caso per caso. È già difficile, in generale, il problema del confronto tra a^b e b^a: si osservi[2] che, per esempio, si ha

$$2^3 < 3^2, \quad 2^4 = 4^2, \quad 2^5 > 5^2.$$

[2]Il problema di trovare le soluzioni razionali di $a^b = b^a$ è molto interessante. Tra gli interi l'unica possibilità è $2^4 = 4^2$, ma nei razionali ce ne sono, invece, infinite: una di queste è

$$\left(\frac{9}{4}\right)^{27/8} = \left(\frac{27}{8}\right)^{9/4}$$

7.5. Funzioni esponenziali

Abbiamo già introdotto, nel capitolo 4, le funzioni potenza, occupiamo ora delle funzioni esponenziali, sempre definite utilizzando il concetto di potenza. Si tratta di funzioni in cui "la base è fissa e l'esponente variabile". Precisamente si dà la seguente definizione.

Definizione 7.3. *Per ogni reale $a > 0$ si dice* funzione esponenziale di base a, *la funzione* $f: \mathbb{R} \to \mathbb{R}$, *data da* $f(x) = a^x$.

Si dimostra il seguente fondamentale teorema.

Teorema 7.4. *Tutte le funzioni esponenziali sono positive, sono strettamente crescenti se $a > 1$, strettamente decrescenti se $0 < a < 1$, costanti per $a = 1$. Se $a > 0$ e $a \neq 1$ le funzioni esponenziali assumono tutti i valori reali > 0 (e una sola volta).*

Le proprietà delle potenze implicano inoltre le seguenti proprietà delle funzioni esponenziali, che è utile tenere ben presenti.

$$a^0 = 1; \tag{7.13a}$$

$$a^1 = a; \tag{7.13b}$$

$$a^{x_1 + x_2} = a^{x_1} a^{x_2}. \tag{7.13c}$$

I grafici delle funzioni esponenziali hanno l'andamento indicato nella figura 7.1.

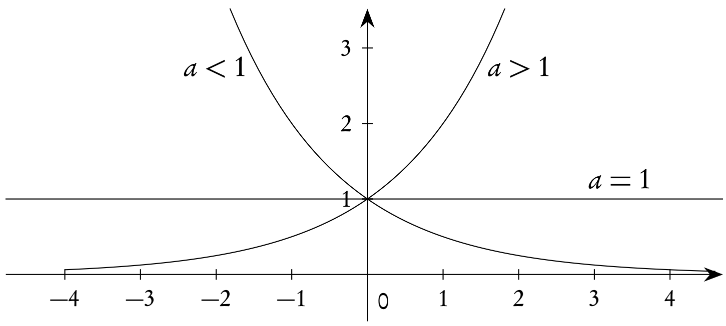


Figura 7.1.: *Grafici delle funzioni esponenziali*

Per le funzioni esponenziali si usa anche la notazione seguente

$$f(x) = a^x = \exp_a(x). \tag{7.14}$$

Quando non c'è possibilità di equivoci si omettono le parentesi tonde che racchiudono la variabile, scrivendo semplicemente $\exp_a x$. Nel caso particolare in cui $a = \mathrm{e}$, cioè quando la base è il numero di Nepero, si parla semplicemente di funzione esponenziale e si omette l'indicazione della base:

$$f(x) = \mathrm{e}^x = \exp(x). \tag{7.15}$$

Questo è legato al fatto che la base "e" per le funzioni esponenziali è, in un certo senso, la base *naturale*: il motivo lo si vedrà nei futuri corsi di analisi, ma possiamo anticipare che è legato al fatto che, con questa base, il coefficiente angolare della tangente al grafico coincide con l'ordinata del punto di tangenza.

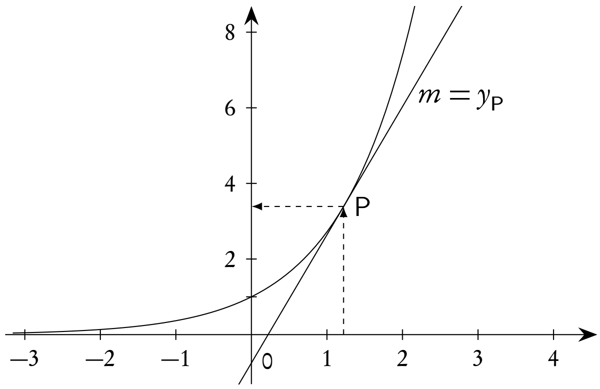


Figura 7.2.: *La funzione $f(x) = e^x$ e il coefficiente angolare della tangente in* P

Le proprietà di monotonia delle funzioni esponenziali hanno le seguenti conseguenze, importanti per la risoluzioni di equazioni e disequazioni coinvolgenti esponenziali. Se a è un reale positivo diverso da 1, per ogni coppia di reali x e y si ha

(7.16a) $$x = y \Leftrightarrow a^x = a^y;$$

(7.16b) $$x < y \Leftrightarrow a^x < a^y, \quad \text{se} \quad a > 1;$$

(7.16c) $$x < y \Leftrightarrow a^x > a^y, \quad \text{se} \quad a < 1.$$

È utile un confronto fra funzioni esponenziali con diversa base. Si veda la figura 7.3.

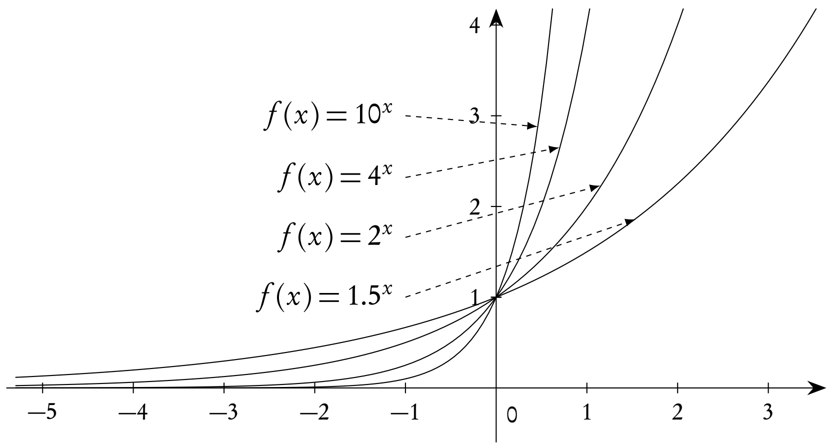


Figura 7.3.: *Confronto tra diverse funzioni esponenziali*

È anche utile un confronto tra funzioni potenza ed esponenziali, per esempio tra 2^x e x^2. In questo caso è molto più significativo il confronto a mezzo di una tabella, piuttosto che quello grafico: la tabella 7.1, riferita solo ad alcuni valori positivi della variabile, mostra come, per piccoli valori di x, x^2 e 2^x siano confrontabili, mentre successivamente 2^x prevale drasticamente su x^2.

È proprio questa caratteristica di rapidissima crescita, per le funzioni esponenziali di base maggiore di 1, che interessa in molte applicazioni.

Tabella 7.1.: *Confronto tra funzioni potenza ed esponenziali*

x	0	1	2	3	4	5	6	7	8	9	10	10^2	10^3
x^2	0	1	4	9	16	25	36	49	64	81	100	10^3	10^6
2^x	1	2	4	8	16	32	64	128	256	512	1024	$\approx 10^{30}$	$\approx 10^{301}$

7.6. I logaritmi

7.6.1. Definizioni e proprietà

Come già detto (vedi il teorema 7.4) le funzioni esponenziali con base $a \neq 1$, ma naturalmente $a > 0$, sono biunivoche tra $\mathbb{R}$ e $\mathbb{R}_{>0}$. Ciò significa che, se a e b sono reali strettamente positivi e $a \neq 1$, l'equazione

$$a^x = b, \tag{7.17}$$

nell'incognita x ha una e una sola soluzione, e si dà la seguente definizione.

Definizione 7.5. *Dati a, $b \in \mathbb{R}$, $a > 0$, $a \neq 1$, $b > 0$, l'unica soluzione dell'equazione (7.17) si chiama* logaritmo in base a si b *e si indica con la scrittura* $\log_a(b)$, *o anche* $\log_a b$.

La definizione data implica che

$$a^{\log_a b} = b, \quad a > 0, a \neq 1, b > 0. \tag{7.18}$$

La (7.18) si usa leggere nel seguente modo: "il logaritmo in base a, strettamente positiva e diversa da 1, del numero reale b, strettamente positivo, è l'esponente da dare ad a per ottenere b".

Esempio 7.2. Si ha $\log_3 81 = 4$, perché $3^4 = 81$.

Esempio 7.3. Si ha $\log_2 \dfrac{1}{8} = -3$, perché $2^{-3} = \dfrac{1}{8}$.

Esempio 7.4. Si ha $\log_5 \sqrt[3]{25} = \dfrac{2}{3}$, perché $5^{2/3} = \sqrt[3]{5^2} = \sqrt[3]{25}$.

Si presti particolare attenzione alle limitazioni a cui sono sottoposti i numeri a e b per la validità della definizione di logaritmo. Per esempio, pur essendo $(-2)^3 = -8$ (cioè pur essendo 3 l'esponente da dare a -2 per ottenere -8), *non* ha senso dire che $\log_{-2}(-8) = 3$, in quanto non si è data alcuna definizione per i logaritmi con base negativa.

In termini di funzioni inverse, la definizione 7.5 equivale a dire che l'inversa della funzione di $\mathbb{R}$ in $\mathbb{R}_{>0}$, data da $f(x) = a^x$, è la funzione

$$\log_a : \mathbb{R}_{>0} \to \mathbb{R}, \quad x \mapsto \log_a x, \quad a > 0, a \neq 1. \tag{7.19}$$

Le funzioni logaritmo hanno dunque l'andamento grafico indicato nella figura 7.4.

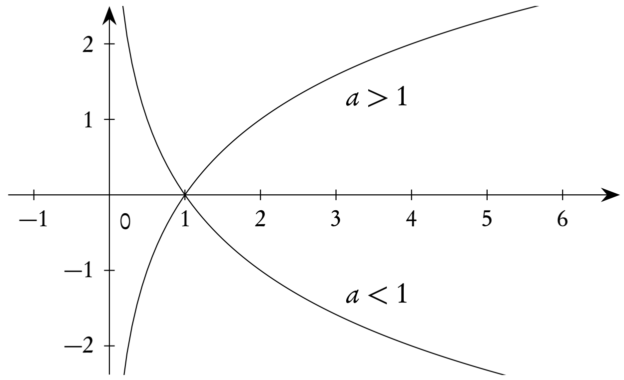


Figura 7.4.: *Andamento delle funzioni logaritmo*

Come già per le funzioni esponenziali, anche per i logaritmi la base più importante è il numero di Nepero "e". In questo caso si omette l'indicazione della base e si usa[3] la scrittura ln (leggi "logaritmo naturale") al posto di $\log_e$. Il motivo di questo fatto risulterà chiaro nei successivi corsi di analisi, ma possiamo anticipare il fatto che, con questa base, il coefficiente angolare della tangente al grafico coincide con il reciproco dell'ascissa del punto di tangenza.

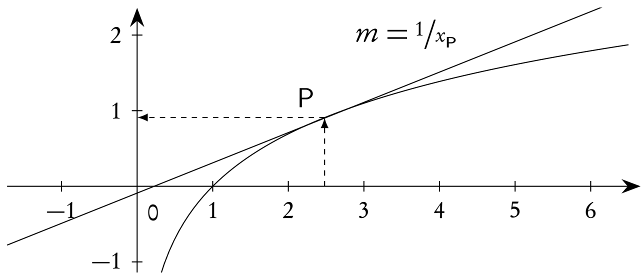


Figura 7.5.: *La funzione* $f(x) = \ln(x)$ *e il coefficiente angolare della tangente in* P

È interessante un confronto tra le figure 7.2 e 7.5. Per comodità abbiamo tracciato i grafici delle due funzioni e delle tangenti in punti "corrispondenti" nella figura 7.6.

Essendo le funzioni e^x ed $\ln x$ inverse una dell'altra, i loro grafici sono simmetrici rispetto alla bisettrice del primo e terzo quadrante. Le tangenti nei due punti P e P' sono anch'esse simmetriche rispetto a quella bisettrice e dunque i due angoli α e β, evidenziati nella figura 7.6, sono complementari. Come mostreremo nel capitolo sulla geometria analitica, ciò implica che i coefficienti angolari rispettivi siano uno il reciproco dell'altro. Poiché, ovviamente, $y_\mathsf{P} = x_{\mathsf{P}'}$, ne segue che se il coefficiente angolare della tangente in P ad e^x è y_P, quello della tangente al grafico di $\ln x$ in P′ deve essere ${}^1/_{y_\mathsf{P}} = {}^1/_{x_{\mathsf{P}'}}$, come sopra accennato.

È molto importante, in particolare per risolvere equazioni e disequazioni, imparare a "leggere" con disinvoltura i grafici delle funzioni esponenziali e logaritmiche, tenendo conto che le funzioni sono inverse una dell'altra. Discorso analogo si verificherà per le funzioni trigonometriche. Un accurato

[3] Anche se quella da noi adottata è la notazione ufficialmente prevista dalle norme ISO in vigore, non è l'unica possibile: si veda quanto scritto nel capitolo sulle notazioni utilizzate.

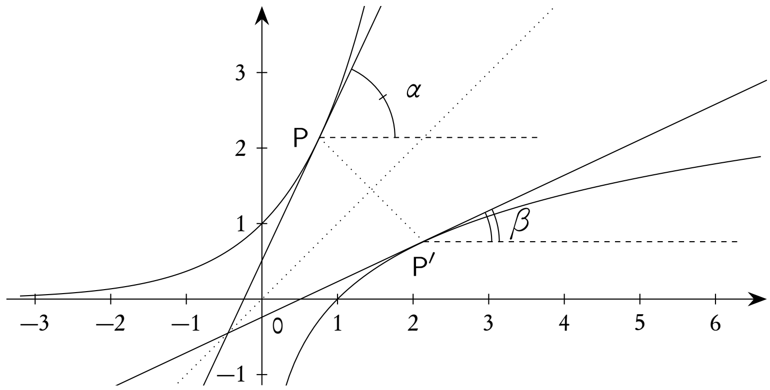


Figura 7.6.: *Le funzioni* e^x, $\ln x$ *e le tangenti in punti corrispondenti*

esame della figura 7.7, relativa al caso di 2^x e $\log_2 x$, aiuterà a chiarire come questi grafici debbano essere usati.

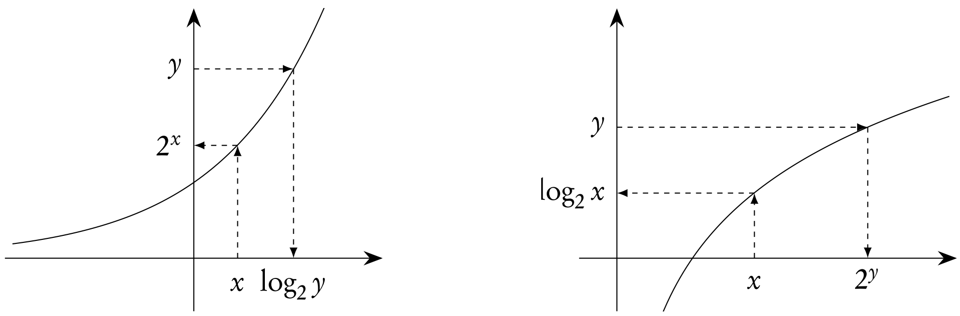


Figura 7.7.: *Grafici di* $f(x) = 2^x$ *e* $g(x) = \log_2 x$

I legami tra una funzione e la sua inversa implicano che, se $a > 0$, $a \neq 1$,

(7.20a) $$\log_a a^x = x, \quad \forall x \in \mathbb{R};$$

(7.20b) $$a^{\log_a x} = x, \quad \forall x \in \mathbb{R},\ x > 0 \text{ (si veda la formula (7.18))};$$

(7.20c) $$\log_a 1 = 0;$$

(7.20d) $$\log_a a = 1.$$

Si noti come per le ultime due delle (7.20) sia fondamentale che $a \neq 1$. Se infatti si ponesse $a = 1$, dalla terza si otterrebbe $\log_1 1 = 0$, dalla quarta $\log_1 1 = 1$, palesemente impossibile.

Le proprietà di monotonia delle funzioni logaritmiche sono conseguenze delle analoghe proprietà delle funzioni esponenziali. In perfetta analogia con le (7.16), si ha allora, se x e y sono due reali strettamente positivi,

(7.21a) $$x = y \iff \log_a x = \log_a y;$$

(7.21b) $$x < y \iff \log_a x < \log_a y, \quad \text{se} \quad a > 1;$$

(7.21c) $$x < y \iff \log_a x > \log_a y, \quad \text{se} \quad a < 1.$$

Dalle proprietà delle potenze discendono le seguenti due proprietà dei logaritmi.

(7.22a) $$\log_a(\alpha \cdot \beta) = \log_a \alpha + \log_a \beta, \quad \alpha > 0, \beta > 0;$$

(7.22b) $$\log_a \alpha^\beta = \beta \log_a \alpha, \quad \alpha > 0.$$

La dimostrazione di queste proprietà per via grafica è particolarmente istruttiva. Cominciamo dalla prima delle (7.22). Esaminiamo la figura 7.8.

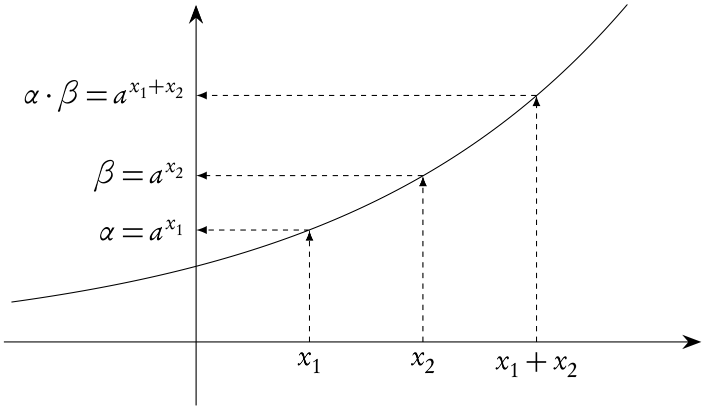


Figura 7.8.: *Dimostrazione della formula* $\log_a(\alpha \cdot \beta) = \log_a \alpha + \log_a \beta, \quad \alpha > 0, \beta > 0$

Per concludere basta solo osservare, "leggendo il grafico in senso inverso", che

$$x_1 = \log_a(\alpha), \quad x_2 = \log_a(\beta), \quad x_1 + x_2 = \log_a(\alpha \cdot \beta).$$

Passiamo ora alla seconda delle (7.22). Esaminiamo la figura 7.9.

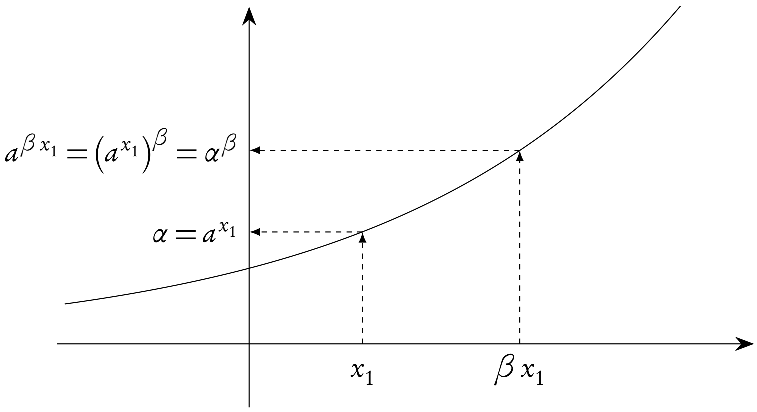


Figura 7.9.: *Dimostrazione della formula* $\log_a \alpha^\beta = \beta \log_a \alpha, \quad \alpha > 0$

Per concludere basta solo osservare, "leggendo il grafico in senso inverso", che

$$x_1 = \log_a \alpha, \quad \beta x_1 = \log_a \left(\alpha^\beta\right).$$

Si presti particolare attenzione all'uso delle (7.22), per la cui validità occorre che $\alpha > 0$ e $\beta > 0$. La situazione è analoga a quella dei radicali (si vedano le equazioni (2.27) e (2.28)). Chiariamo la situazione con due esempi, che mettono in luce le difficoltà ed evidenziano le possibili strategie risolutive.

Esempio 7.5. Si ha $\log_a\big((-2)(-3)\big) = \log_a 2 + \log_a 3$. In generale $\log_a(\alpha\beta) = \log_a|\alpha| + \log_a|\beta|$, se α e β sono entrambi negativi.

Esempio 7.6. Si ha $\log_a(-5)^2 = 2\log_a 5$. In generale $\log_a x^2 = 2\log_a|x|$.

Sono anche importanti per le applicazioni le formule dei *cambiamenti di base* nelle potenze e nei logaritmi.

$$a^\alpha = b^{\alpha\log_b(a)}, \quad \log_a b = \frac{\log_c b}{\log_c a}. \tag{7.23}$$

Entrambe le (7.23) sono una semplice conseguenza della definizione di logaritmo. Per la prima si ha:

$$b^{\alpha\log_b(a)} = b^{\log_b(a^\alpha)} = a^\alpha.$$

Per la seconda si ha:

$$a^{\log_a b} = b \Rightarrow \log_c\left(a^{\log_a b}\right) = \log_c b \Rightarrow \left(\log_a b\right)\left(\log_c a\right) = \log_c b \Rightarrow \log_a b = \frac{\log_c b}{\log_c a}.$$

Dalla seconda delle (7.23), se $c = b$, si trae

$$\log_a b = \frac{1}{\log_b a}. \tag{7.24}$$

7.6.2. Le funzioni del tipo $f(x)^{g(x)}$

Consideriamo una funzione del tipo $f(x)^{g(x)}$ e chiediamoci qual è il suo dominio naturale. Come è noto, il dominio naturale di una funzione di variabile reale è l'insieme di tutti i reali x per cui i calcoli necessari per ottenere da x il valore della funzione hanno senso. Se indichiamo dunque con E il dominio dell'esponente $g(x)$, per trovare il dominio D di una funzione del tipo $f(x)^{g(x)}$ dobbiamo fare l'unione dei seguenti tre insiemi.

$$D_1 = \{\, x \mid f(x) > 0 \,\} \cap E; \tag{7.25a}$$
$$D_2 = \{\, x \mid f(x) = 0 \,\} \cap \{\, x \mid g(x > 0 \,\}; \tag{7.25b}$$
$$D_3 = \{\, x \mid f(x) < 0 \,\} \cap \{\, x \mid g(x) \in \mathbb{Z} \,\}. \tag{7.25c}$$

Proponiamo alcuni esempi.

Esempio 7.7. La funzione $f(x) = x^x$ ha come dominio naturale $\{\, x \in \mathbb{Z} \mid x < 0 \,\} \cup]0, +\infty[$.

Esempio 7.8. La funzione $f(x) = \left(x^2\right)^x$ ha come dominio naturale $\mathbb{R} \setminus \{0\}$.

Esempio 7.9. La funzione $f(x) = x^{(2x)}$ ha $\left\{\, \dots, -3, -\frac{5}{2}, -2, -\frac{3}{2}, -1, -\frac{1}{2} \,\right\} \cup]0, +\infty[$ come dominio naturale, ovvero i reali positivi e, tra i negativi, solo quelli per cui $2x$ è intero, ovvero quelli del tipo $x = -{}^n/_2$, con $n \in \mathbb{N}$, $n \geq 1$.

Si noti come le funzioni del secondo e terzo degli esempi proposti siano diverse, in quanto a dominio naturale.

Tuttavia quando si considerano funzioni di questo tipo, di solito si conviene di accettare come dominio solo l'insieme D_1 dei tre della (7.25). Per questi x, e *solo* per questi x, la prima delle (7.23) consente di ottenere la comoda formula

$$f(x)^{g(x)} = a^{g(x)\log_a f(x)}, \text{ se } f(x) > 0 \text{ e } a > 0,\ a \neq 1. \tag{7.26}$$

L'importanza della formula è dovuta al fatto che trasforma una funzione del tipo $f(x)^{g(x)}$ in una normale funzione esponenziale (cioè con base fissa). Se si assume, come si fa di solito, come base il numero "e", la (7.26) diventa

$$f(x)^{g(x)} = \mathrm{e}^{g(x)\ln f(x)}, \text{ se } f(x) > 0. \tag{7.27}$$

7.7. Equazioni e disequazioni esponenziali e logaritmiche

Sono *equazioni esponenziali* quelle in cui l'incognita figura ad esponente; sono *equazioni logaritmiche* quelle in cui l'incognita figura come argomento di logaritmi. La risoluzione di equazioni di questo tipo può essere molto complessa: per esempio l'equazione $x + 2^x = 0$ sicuramente non può avere soluzioni maggiori o uguali a zero (quando il primo membro è la somma di due numeri non negativi), ma può averne di negative (quando il primo membro è la somma di un numero negativo e di uno positivo non superiore ad 1), anzi le proprietà della funzione (che si studieranno nei corsi di analisi) implicano che ne ha esattamente una, compresa tra -1 e 0; tuttavia non esistono strategie elementari per trovarla, anche se esistono tecniche per trovarne un valore approssimato con qualsivoglia numero di cifre decimali esatto, cosa molto importante per le applicazioni: anche queste tecniche saranno oggetto dei corsi di analisi. Qui ci limitiamo a proporre, nella figura 7.10, il grafico della funzione a primo membro, $f(x) = x + 2^x$, che fornisce una conferma dell'affermazione fatta: il grafico interseca l'asse delle ascisse solo nel punto P.

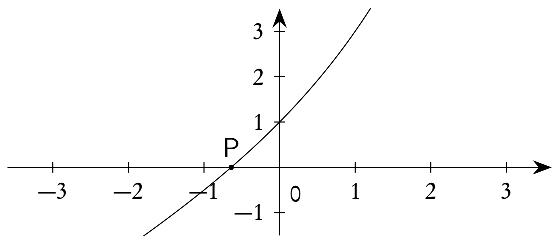


Figura 7.10.: *Grafico della funzione* $f(x) = x + 2^x$

Analogo discorso per le disequazioni. Noi ci occuperemo solo di alcuni casi semplici.

7.7.1. Equazioni e disequazioni elementari

Sono dette *elementari* le equazioni, esponenziali o logaritmiche, del tipo

$$a^x = b; \tag{7.28a}$$

$$\log_a x = b. \tag{7.28b}$$

Per quelle esponenziali si nota che, se $b \leq 0$, non hanno soluzioni; se invece $b > 0$ si risolvono prendendo il logaritmo, di solito in base a, di ambo i membri. Per quelle logaritmiche si comincia con lo scrivere la condizione per l'esistenza del logaritmo, $x > 0$, dopodiché si prende l'esponenziale, in base a, di ambo i membri.

Sono *disequazioni elementari* le disequazioni del tipo

$$a^x \gtreqless b; \tag{7.29a}$$

$$\log_a x \gtreqless b. \tag{7.29b}$$

cioè quelle in cui, nelle (7.28), il segno di $=$ è sostituito da un segno di disuguaglianza. Per risolverle bisogna innanzitutto ricordare che le funzioni esponenziali sono sempre strettamente positive per qualsiasi valore dell'esponente e che, se $a > 1$, sia le funzioni esponenziali che quelle logaritmiche sono strettamente crescenti, se invece $0 < a < 1$ sono strettamente decrescenti. Poi, se del caso, bisognerà prendere il logaritmo o l'esponenziale di ambo i membri, prestando attenzione al fatto che se $0 < a < 1$ bisognerà cambiare il verso, se $a > 1$ invece mantenere il verso.

Vediamo alcuni esempi, sia per equazioni che per disequazioni.

Esempio 7.10. Risolvere l'equazione $3^x = 5$. Si ha $\log_3 3^x = \log_3 5$, da cui $x = \log_3 3 = \log_3 5$ e, infine, $x = \log_3 5$. Si sarebbero potuti usare anche i logaritmi in base e (soprattutto per il fatto che, nelle calcolatrici tascabili, i logaritmi in base 3 non sono previsti in modo standard). Procedendo nell stesso modo si sarebbe ottenuto $x = {}^{\ln 5}/_{\ln 3}$, che è lo stesso di prima (tenendo conto della formula di cambiamento di base).

Per risolvere la disequazione $3^x < 5$ si procede sempre nello stesso modo, ottenendo ($3 > 1$!), $x < \log_3 5$.

Esempio 7.11. Risolvere l'equazione $\left({}^1/_2\right)^x = 3$. Procedendo come nell'esempio precedente si ottiene facilmente $x = \log_{1/2} 3$.

Per risolvere la disequazione $\left({}^1/_2\right)^x > 3$ bisogna osservare che ${}^1/_2 < 1$, da cui $x < \log_{1/2} 3$ (si noti il cambio di verso).

Esempio 7.12. Risolvere l'equazione $3^x = -2$. Poiché il primo membro è positivo e il secondo è negativo, non ci sono soluzioni.

Per la disequazione $3^x < -2$ si possono trarre le stesse conclusioni: nessuna soluzione. Per la $3^x > -2$, invece, con la stessa osservazione si trae invece che l'insieme di soluzioni è tutto $\mathbb{R}$.

Esempio 7.13. Risolvere l'equazione $\log_2 x = 5$. Prendendo l'esponenziale in base 2 di ambo i membri si ha $2^{\log_2 x} = 2^5$, ovvero $x = 32$.

Per la disequazione $\log_2 x < 5$ si trova, con lo stesso metodo, $0 < x < 32$: si tenga conto che l'argomento del logaritmo *deve* essere strettamente positivo.

Esempio 7.14. Risolvere l'equazione $\log_{1/2} x = -2$. Prendendo l'esponenziale in base ${}^1/_2$ di ambo i membri si trova $x = 4$.

Per la disequazione $\log_{1/2} x < -2$, con la stessa tecnica si trova $x > 4$. Invece per la disequazione $\log_{1/2} x > -2$ si può applicare la stessa tecnica, ma bisogna poi ricordare che l'argomento del logaritmo deve essere strettamente positivo, da cui si trova $0 < x < 4$.

Come si vede non esistono sostanziali differenze nelle tecniche da applicare per risolvere le equazioni e le disequazioni: l'unica cosa da tenere presente è che nel caso delle equazioni il fatto che le funzioni siano crescenti o decrescenti non ha importanza, mentre questo fatto è cruciale nel caso delle disequazioni.

7.7.2. Equazioni e disequazioni non elementari

Per le equazioni e disequazioni non elementari non ci sono strategie standard. L'unica idea di base è quella di cercare di ricondursi a una o più equazioni elementari. Proporremo una serie di esempi guida che possono dare un'idea di quali siano le strategie più comuni.

Esempio 7.15.

Risolvere l'equazione $2^{x^2-1}=8$. Si tratta, in realtà, sempre di una equazione di tipo elementare: prendendo il logaritmo in base 2 di ambo i membri, e ricordando che $\log_2 8=\log_2 2^3=3$, si ottiene $x^2-1=3$, da cui $x=\pm 2$. In maniera equivalente si poteva osservare che $8=2^3$ e concludere subito che doveva essere $x^2-1=3$.

Esempio 7.16.

Risolvere l'equazione $\log_5(x+|x|+1)=0$. Anche questa è, in realtà, una equazione elementare. Il dominio del primo membro è $\mathbb{R}$, come si può facilmente verificare. Prendendo l'esponenziale in base 5 di ambo i membri si trova poi $x+|x|+1=1$, che ha come soluzioni $]-\infty,0]$.

Esempio 7.17. Risolvere l'equazione

$$2^{x+1}=5-4^{x-1}.$$

L'equazione può essere trasformata come segue

$$2^x\times 2^1=5-4^x\times 4^{-1}\Rightarrow 2^x\times 2=5-2^{2x}\times 4^{-1}\Rightarrow 2^x\times 2=5-\left(2^x\right)^2\times 4^{-1}.$$

A questo punto, posto $2^x=t$, si ottiene, dopo semplificazione, $t^2+8t-20=0$, da cui $t=-10\vee t=2$. L'equazione data è dunque equivalente a

$$2^x=-10\vee 2^x=2.$$

La prima equazione non ha soluzioni, la seconda ha la soluzione $x=1$. Si ha dunque $S=\{1\}$.

Nello stesso modo si possono risolvere tutte le equazioni in cui compare solo a^x, oppure $\log_a x$, purché la sostituzione $t=a^x$ oppure $t=\log_a x$ porti a un'equazione risolubile. Esempi di questo tipo, che invitiamo a risolvere, sono

— $9^x-2\times 3^x-5=0$;
— $\log_5^2 x+\log_5 x-6=0$;
— $\log_3^2 x+\log_3\left(x^2\right)-6=0$.

Osservazione 7.6. Si noti, nelle equazioni sopra proposte, la scrittura $\log_5^2 x$. Si tratta di una scrittura un po' ambigua, ma ormai entrata nell'uso comune. Essa deve essere interpretata come $\left(\log_5 x\right)^2$. A voler essere assolutamente pignoli si dovrebbe addirittura scrivere $\left(\log_5(x)\right)^2$, ma l'uso è quello di scrivere meno parentesi possibili, quando non ci sono equivoci. Si noti, altresì, che

$$\log_3^2 x\neq\log_3 x^2\ : \tag{7.30}$$

la differenza sta nell'ordine in cui le operazioni devono essere eseguite.

— In $\log_3^2 x$ si calcola *prima* il logaritmo in base 3 del numero x e *successivamente* si eleva al quadrato.

— In $\log_3 x^2$ si calcola *prima* il quadrato di x e *successivamente* si prende il logaritmo del risultato.

Se risolviamo l'equazione

$$\log_3^2 x = \log_3 x^2,$$

otteniamo $(\log_3 x)^2 = 2\log_3 |x|$. Poiché dal dominio otteniamo $x > 0$, si ha $|x| = x$. Posto $\log_3 x = t$ troviamo $t = 0 \vee t = 2$, da cui $x = 1 \vee x = 9$. Tutto questo trova conferma nei grafici delle due funzioni che compaiono nella (7.30): la figura 7.11 evidenzia la sostanziale differenza tra le due situazioni e conferma che l'uguaglianza vale solo per $x = 1 \vee x = 9$.

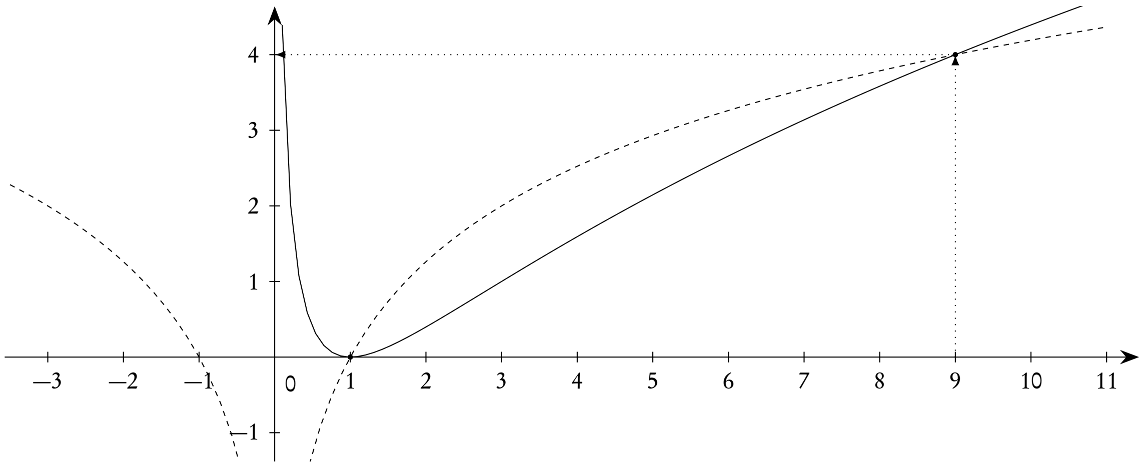


Figura 7.11.: *Grafici di $f(x) = \log_3^2 x$ (linea continua) e $g(x) = \log_3 x^2$ (in tratteggio)*

Esempio 7.18. Risolvere l'equazione

$$\log_5(x-3) - \log_5(5-x) = 0.$$

Troviamo innanzitutto il dominio

$$\begin{cases} x-3>0 \\ 5-x>0 \end{cases} \quad \Rightarrow \quad 3<x<5.$$

L'equazione si può riscrivere come segue.

$$\log_5 \frac{x-3}{5-x} = 0,$$

da cui, prendendo l'esponenziale in base 5 di ambo i memebri,

$$\frac{x-3}{5-x} = 1 \Rightarrow x-3 = 5-x \Rightarrow x = 4.$$

Si noti che abbiamo "eliminato il denominatore" perché, in base al dominio esso è strettamente positivo. Dunque $S = \{4\}$.

Si noti altresì che siamo passati alla scrittura

$$\log_5 \frac{x-3}{5-x}$$

solo *dopo* aver trovato il dominio. In questo caso non sarebbe cambiato nulla: le due funzioni

$$\log_5(x-3)-\log_5(5-x) \quad \text{e} \quad \log_5 \frac{x-3}{5-x}$$

hanno lo stesso dominio, ma la cosa non è vera in generale. Per esempio delle due funzioni

$$\log_5(x-3)-\log_5(x-5) \quad \text{e} \quad \log_5 \frac{x-3}{x-5}$$

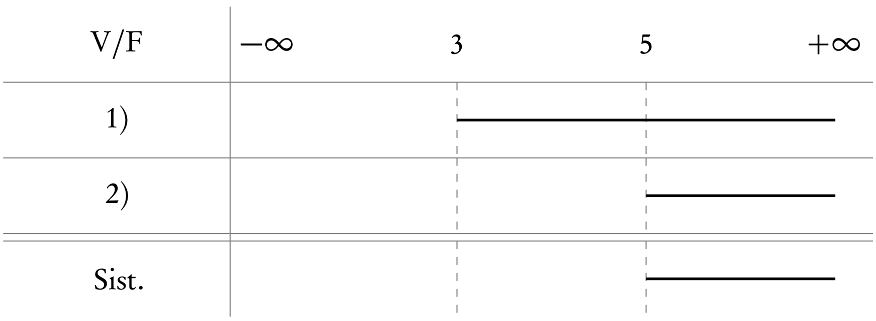
la prima ha come dominio $x > 5$, la seconda $x < 3 \vee x > 5$. Infatti per la prima si deve risolvere il sistema

$$\begin{cases} x-3>0 \\ x-5>0 \end{cases},$$

per la seconda la disequazione fratta

$$\frac{x-3}{x-5} > 0.$$

Il grafico relativo al sistema è il seguente:



Il grafico relativo alla disequazione fratta è, invece, il seguente.

+/−	$-\infty$		3		5		$+\infty$
$x-3$		−	0	+	+	+	
$x-5$		−	−	−	0	+	
Complessivo		+	0	−	0	+	

Esempio 7.19. Risolvere la disequazione

$$3^{x+2}-\frac{9}{5}\times 5^{x+1}<3^x-5^x\,.$$

Si possono considerare i passaggi seguenti.

$$3^x\times 3^2-\frac{9}{5}\times 5^x\times 5<3^x-5^x\Rightarrow 8\big(3^x<5^x\big)<0\Rightarrow 3^x<5^x\Rightarrow\left(\frac{3}{5}\right)^x<1.$$

L'ultima è una disequazione elementare, la cui soluzione è $x>0$: abbiamo cambiato il verso perché $3/5<1$.

Esempio 7.20. Risolvere la disequazione

$$\ln(2^x-2)>\ln(2\times 2^{2x}+2)-\ln(2^x-1).$$

Troviamo anzitutto il dominio.

$$\begin{cases} 2^x-2>0\\ 2\times 2^{2x}+2>0\\ 2^x-1>0\end{cases}.$$

La prima e la terza disequazione sono elementari e hanno come soluzioni, rispettivamente, $x>1$ e $x>0$; la seconda disequazione è sempre vera (il primo membro è la somma di due numeri strettamente positivi). Il dominio è dunque $x>1$.

Posto ora $2^x=t$ e applicando le proprietà dei logaritmi, si ottiene facilmente

$$\ln\frac{(t-2)(t-1)}{2t^2+2}>0\Rightarrow\frac{(t-2)(t-1)}{2t^2+2}>1\Rightarrow(t-2)(t-1)>2t^2+2\,,$$

dove abbiamo "eliminato" il denominatore perché strettamente positivo. L'ultima disequazione, che è di secondo grado, ha come soluzioni $-3<t<0$, da cui $-3<2^x<0$ che non ha soluzioni (2^x è sempre strettamente positivo).

Esempio 7.21. Risolvere la disequazione

$$\frac{2^{\sqrt{3-x}}}{(x-1)\ln(x+2)}>0.$$

Troviamo innanzitutto il dominio.

$$\begin{cases} 3-x\geq 0\Rightarrow x\leq 3\\ x+2>0\Rightarrow x>-2\\ (x-1)\ln(x+2)\neq 0\Rightarrow x\neq -1\wedge x\neq 1\end{cases}$$

Il dominio è dunque

$$]-2,-1[\,\cup\,]-1,1[\,\cup\,]1,3].$$

A questo punto il numeratore della frazione è strettamente positivo è può essere semplificato. Si deve trovare il segno dei due fattori al denominatore. Si ottiene il grafico seguente, nel quale, in ogni linea, abbiamo tenuto conto del dominio complessivo trovato, anche.

$+/-$	$-\infty$		-2		-1		1		3		$+\infty$
$x-1$		$\times$	$\times$	$-$	$\times$	$-$	$\times$	$+$	$+$	$\times$	
$\ln(x+2)$		$\times$	$\times$	$-$	$\times$	$+$	$\times$	$+$	$+$	$\times$	
Complessivo		$\times$	$\times$	$+$	$\times$	$-$	$\times$	$+$	$+$	$\times$	

L'insieme di soluzioni è dunque

$$S =]-2,-1[\,\cup\,]1,3].$$

7.7.3. Qualche considerazione grafica

Nel risolvere le disequazioni elementari (alle quali tutte si riconducono) occorre prestare la massima attenzione alla base degli esponenziali e dei logaritmi e si deve inoltre tenere conto del fatto che l'argomento dei logaritmi deve essere strettamente positivo. Per evitare inopportune dimenticanze è sempre utile fare riferimento alle rappresentazioni grafiche, come mostrato nei seguenti esempi.

Esempio 7.22. Risolvere la disequazione

$$\log_{2/3} x > -2.$$

La soluzione per via algebrica è immediata, ma la visualizzazione grafica rende molto meglio conto di quanto succede ed evita pericolose dimenticanze:

$$S = \left]0, \left(\frac{2}{3}\right)^{-2}\right[= \left]0, \frac{9}{4}\right[.$$

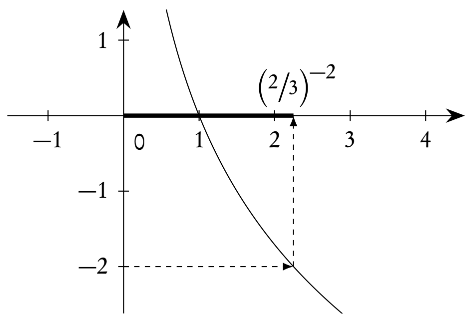


Figura 7.12.: *La disequazione* $\log_{2/3} x > -2$

Esempio 7.23. Risolvere la disequazione

$$\log_3 x < 2$$

Anche in questo caso la risoluzione per via algebrica è immediata, ma la visualizzazione grafica è oltremodo istruttiva.

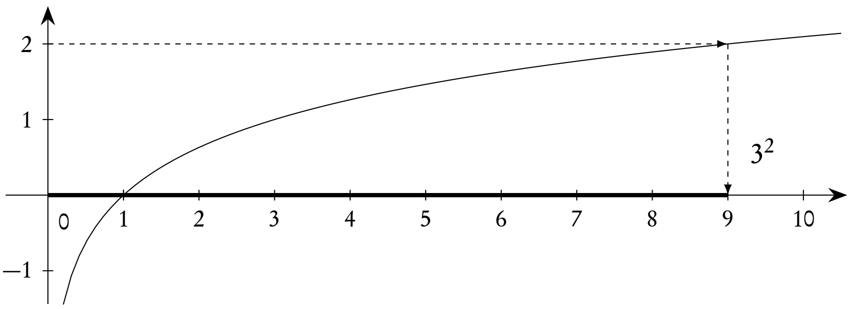


Figura 7.13.: *La disequazione* $\log_3 x < 2$

$$S =]0, 3^2[=]0, 9[.$$

Esempio 7.24. Risolvere il sistema di disequazioni

$$2 < 3^x < 4.$$

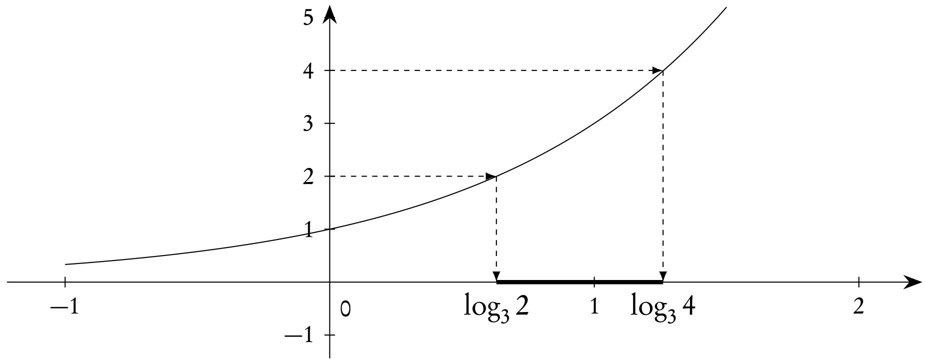


Figura 7.14.: *Il sistema di disequazioni* $2 < 3^x < 4$

Dunque

$$S =]\log_3 2, \log_3 4[.$$

7.8. Esercizi

Esercizio 7.1. *Calcolare il logaritmo in base* 2 *dei seguenti numeri:*

$$2,\ 4,\ 8,\ 16,\ \sqrt{2},\ \sqrt[3]{2},\ \frac{1}{2},\ \frac{1}{16},\ 4\sqrt{2},\ \frac{\sqrt{2}}{2},\ \frac{\sqrt{2}}{4}.$$

Risoluzione. I numeri dati possono tutti essere scritti come potenze di 2:

$$2^1, 2^2, 2^3, 2^4, 2^{1/2}, 2^{1/3}, 2^{-1}, 2^{-4}, 2^2 \times 2^{1/2} = 2^{5/2}, 2^{1/2} \times 2^{-1} = 2^{-1/2}, 2^{1/2} \times 2^{-2} = 2^{-3/2}.$$

I logaritmi richiesti sono allora

$$1, 2, 3, 4, \frac{1}{2}, \frac{1}{3}, -1, -4, \frac{5}{2}, -\frac{1}{2}, -\frac{3}{2}. \qquad \square$$

Esercizio 7.2. *Dire quale delle seguenti uguaglianze sono vere e quali sono false, e perché (la base dei logaritmi può essere arbitraria).*

1. $\log^2 x = \log x^2$.
2. $\log^2 x = \log\log x$.
3. $\log x^4 = 2\log x^2$.
4. $\log x^3 = 3\log x$.

Risoluzione. La 1. è falsa: il primo membro significa $(\log x)^2$, il secondo $\log(x^2)$. È dunque l'ordine delle operazioni che è diverso; si noti, tra l'altro, che il primo membro ha dominio $x > 0$, il secondo $x \neq 0$.

La 2. è falsa. Il primo membro significa $(\log x)^2 = (\log x)(\log x)$, il secondo no. Si noti, tra l'altro, che il primo membro ha dominio $x > 0$, il secondo $x > 1$ (si deve infatti imporre $x > 0 \wedge \log x > 0$, che comporta $x > 0 \wedge x > 1$, ovvero $x > 1$).

La 3. è vera. Il passaggio dal primo al secondo membro si fa osservando che

$$\log x^4 = \log(x^2)^2 = 2\log x^2,$$

e l'ultimo passaggio è lecito perché non modifica il dominio.

La 4. è vera per le proprietà dei logaritmi, applicando le quali non si modifica il dominio. $\square$

Esercizio 7.3. *Risolvere la disequazione*

$$\log_x(x^2 + 2x) > 0.$$

Risoluzione. Troviamo innanzitutto il dominio:

$$\begin{cases} x^2 + x > 0 \\ x > 0 \\ x \neq 1 \end{cases} :$$

la prima condizione è quella richiesta dall'argomento di un logaritmo (che deve essere strettamente positivo), le altre due sono quelle richieste dalla base di un logaritmo (che deve essere strettamente positiva e diversa da 1). La soluzione del sistema è elementare e porge

$$D =]0, 1[\cup]1, +\infty[.$$

A questo punto conviene riscrivere la disequazione cambiando la base del logaritmo, per esempio nella base e.

$$\frac{\ln(x^2 + x)}{\ln x} > 0.$$

Troviamo ora il segno di numeratore e denominatore, segnalando che terremo conto del dominio al momento della costruzione del grafico.

$$\ln(x^2+x)>0 \Rightarrow x^2+x>1 \Rightarrow x<-1-\sqrt{2} \vee x>-1+\sqrt{2}.$$
$$\ln x>0 \Rightarrow x>1.$$

Si ottiene il grafico seguente.

$+/-$	$-\infty$		0		$-1+\sqrt{2}$		1		$+\infty$
N		$\times$	$\times$	$-$	0	$+$	$\times$	$+$	
D		$\times$	$\times$	$-$	$-$	$-$	$\times$	$+$	
Complessivo		$\times$	$\times$	$+$	0	$-$	$\times$	$+$	

Si ha dunque

$$S=\left]0,-1+\sqrt{2}\right]\cup]1,+\infty[.$$ □

Esercizio 7.4. *Risolvere la disequazione*

$$3^{3x}-2\times 3^{3x/2}-3\leq 0.$$

Risoluzione. Posto $3^{3x/2}=t$, si ottiene

$$t^2-2t-3\leq 0 \Rightarrow -1\leq t\leq 3 \Rightarrow -1\leq 3^{3x/2}\leq 3.$$

La prima disequazione è sempre verificata, la seconda fornisce

$$\frac{3x}{2}\leq 1 \Rightarrow x\leq\frac{2}{3}.$$ □

Esercizio 7.5. *Risolvere la disequazione*

$$\ln\big(\ln(x+2)\big)>0.$$

Risoluzione. Dominio:

$$\begin{cases} x+2>0 \Rightarrow x>-2 \\ \ln(x+2)>0 \Rightarrow x+2>1 \Rightarrow x>-1 \end{cases}.$$

Dunque $D=]-1,+\infty[$.

A questo punto prendendo l'esponenziale in base e di ambo i membri si ottiene $\ln(x+2)>1$ e successivamente, prendendo nuovamente l'esponenziale in base e di ambo i membri, $x+2>\mathrm{e}$, ovvero $x>\mathrm{e}-2$. □

Esercizio 7.6. *Risolvere l'equazione*

$$3^{x+2}+3^{2-x}=82.$$

Risoluzione. Si possono eseguire i seguenti passaggi:

$$3^x \times 3^2 + 3^2 \times 3^{-x} = 82 \Rightarrow 9 \times 3^x + 9\frac{1}{3^x} = 82.$$

Posto $3^x = t$, si ottiene $9t^2 - 82t + 9 = 0$, da cui $t_1 = 9$, $t_2 = 1/9$ e, infine $x_1 = 2$, $x_2 = -2$. □

Esercizio 7.7. *Risolvere la seguente disequazione*

$$25^{\sqrt{x-2}} - 5 \times 5^{\sqrt{x-2}} - 500 < 0.$$

Risoluzione. Per il dominio si deve avere $x \geq 2$. Posto poi $5^{\sqrt{x-2}} = t$ si trova

$$t^2 - 5t - 500 < 0 \Rightarrow -20 < t < 250 \Rightarrow -20 < 5^{\sqrt{x-2}} < 25.$$

La prima disequazione, nel dominio, è sempre verificata, per la seconda si deve avere $\sqrt{x-2} < 2$, ovvero $x < 6$. Si ha dunque

$$S = [2, 6[.$$ □

Esercizio 7.8. *Risolvere l'equazione*

$$x^x = x^3.$$

Risoluzione. Preoccupiamoci intanto del dominio. Nessun problema col secondo membro. Il primo membro ha senso[(4)] per $x > 0$ e per $x \in \mathbb{Z}$, $x < 0$. È facile convincersi che l'unica soluzione intera negativa è -1. Se $x > 0$ si deve avere

$$e^{x \ln x} = e^{3 \ln x} \Rightarrow x \ln x = 3 \ln x \Rightarrow x = 3 \vee \ln x = 0 \Rightarrow x = 3 \vee x = 1.$$

Si ha dunque

$$S = \{-1, 1, 3\}.$$ □

Esercizio 7.9. *Risolvere l'equazione*

$$x^{\log x} = 96\, x^{\log \sqrt{x}} + 400,$$

dove log *indica il logaritmo in base* 10.

Risoluzione. Il dominio è costituito dai reali strettamente positivi. Si ha poi

$$\log \sqrt{x} = \frac{1}{2} \log x.$$

Posto allora

$$x^{\log \sqrt{x}} = t,$$

si ottiene $t^2 - 96t - 400 = 0$, da cui $t_1 = -4$ e $t_2 = 100$. Da

$$x^{\log \sqrt{x}} = -4$$

[4] In genere con funzioni del tipo $f(x)^{g(x)}$ si conviene, vedi il paragrafo 7.6.2, che il dominio contenga solo i reali per cui $f(x) > 0$, ma in un problema come questo può avere interesse anche la ricerca di eventuali soluzioni negative.

non si ottiene alcuna soluzione. Si ha invece

$$x^{\log\sqrt{x}} = 100 \Rightarrow 10^{\log\sqrt{x}\,\log x} = 10^2 \Rightarrow \log\sqrt{x}\,\log x = 2 \Rightarrow \frac{1}{2}\log^2 x = 2 \Rightarrow \log x = \pm 2,$$

da cui, infine,

$$S = \left\{ 100,\ \frac{1}{100} \right\}.$$

□

Esercizio 7.10. *Trovare il dominio della seguente funzione*

$$f(x) = \log\log\log x.$$

Risoluzione. Si deve considerare il sistema

$$\begin{cases} x > 0 \\ \log x > 0 \Rightarrow x > 1 \\ \log\log x > 0 \Rightarrow \log x > 1 \Rightarrow x > 10 \end{cases}.$$

Dunque $D =]10, +\infty[$.

□

Esercizio 7.11. *Risolvere la disequazione*

$$4x^2 + 3^{\sqrt{x}+1} + x\cdot 3^{\sqrt{x}} < 2x^2\cdot 3^{\sqrt{x}} + 2x + 6.$$

Risoluzione. La sola speranza di risolvere una disequazione come quella proposta è di ridurla alla forma $f(x) < 0$, con $f(x)$ scomponibile in fattori di cui si sappia trovare il segno: nella forma data nessuna tecnica elementare da sola può produrre risultati.

Osserviamo intanto che il dominio è $x \geq 0$. Successivamente si ha

$$3^{\sqrt{x}+1} = 3 \times 3^{\sqrt{x}}.$$

Portando tutto a primo membro e raccogliendo $3^{\sqrt{x}}$ dove è presente, si ottiene

$$2(2x^2 - x - 3) - 3^{\sqrt{x}}(2x^2 - x - 3) < 0 \Rightarrow \left(2 - 3^{\sqrt{x}}\right)(2x^2 - x - 3) < 0.$$

Studiamo ora il segno dei due fattori.

$$2 - 3^{\sqrt{x}} > 0 \Rightarrow 3^{\sqrt{x}} < 2 \Rightarrow \sqrt{x} < \log_3 2 \Rightarrow x < \log_3^2 2.$$
$$2x^2 - x - 3 > 0 \Rightarrow x < -1 \vee x > 3/2.$$

Tenendo conto del dominio si ottiene il grafico seguente, dove abbiamo tenuto conto che $\log_3^2 2 < 3/2$, in quanto il primo numero è minore di 1, il secondo è maggiore di 1.

$+/-$	$-\infty$		0		$\log_3^2 2$		$\frac{3}{2}$		$+\infty$
1°		×	+	+	0	−	−	−	
2°		×	−	−	−	−	0	+	
Complessivo		×	−	−	0	+	0	−	

Si conclude che
$$S = \left[0, \log_3^2 2\right[\cup \left]\frac{3}{2}, +\infty\right[. \qquad \square$$

8. Geometria analitica

Abbiamo già utilizzato le coordinate cartesiane principalmente per rappresentare grafici di funzioni reali. In questo capitolo vogliamo affrontare il problema in maniera più sistematica.

8.1. Generalità

Cominciamo con il riepilogare il concetto di ascisse sulla retta, già trattato nel paragrafo 2.5.2. Sappiamo che è possibile stabilire una corrispondenza biunivoca tra i reali e i punti di una retta, introducendo quello che si chiama un *sistema di ascisse* sulla retta stessa, che in ragione di questo fatto viene chiamata *asse delle ascisse*. Sostanzialmente si procede nel seguente modo: su una retta r si fissa un punto O, detto *origine* e un punto U. Se si assume la lunghezza del segmento $\overline{\mathsf{OU}}$ come unità di misura, ad esso si può far corrispondere il numero reale 1; ad ogni altro punto P della semiretta OU si può far corrispondere il numero reale positivo che rappresenta la lunghezza del segmento $\overline{\mathsf{OP}}$; ad ogni punto P dell'altra semiretta di origine O si può far corrispondere l'opposto del numero che rappresenta la lunghezza del segmento $\overline{\mathsf{OP}}$. In questo modo si viene ad associare ad ogni numero reale uno e un sol punto della retta su cui si sia fissato il sistema di ascisse, e viceversa, potendo addirittura identificare i numeri reali con i punti della retta stessa: in molti contesti si parla addirittura (e anche noi lo faremo spesso) di "punti" come sinonimo di "numeri reali". Si parla spesso di *retta reale* per intendere l'insieme dei numeri reali, che è comunemente immaginato proprio come una retta orientata.

Passiamo ora al piano. Consideriamo su un piano π una coppia di rette distinte (r,s) incidenti in un punto O e su ciascuna di esse introduciamo un sistema di ascisse. È indispensabile considerare una coppia di rette (cioè una *prima* retta e una *seconda* retta) e non semplicemente un insieme di due rette: l'ordine gioca un ruolo cruciale in quanto diremo. Sia ora P un punto del piano. Da esso tracciamo le parallele alle due rette date, ottenendo un punto $\mathsf{P_r} \in \mathsf{r}$ e un punto $\mathsf{P_s} \in \mathsf{s}$, ciascuno dei quali ha un'ascissa. In conclusione al punto P risulta associata la coppia di numeri reali costituita dalle ascisse di $\mathsf{P_r}$ e $\mathsf{P_s}$. Viceversa data una coppia di numeri reali è possibile individuare un punto $\mathsf{P_r} \in \mathsf{r}$ e un punto $\mathsf{P_s} \in \mathsf{s}$ aventi come ascissa rispettivamente il primo e il secondo elemento della coppia. Tracciate per questi due punti le parallele a r e a s, si individua un punto P che si può pensare associato alla coppia data. In questo modo è possibile stabilire una corrispondenza biunivoca tra punti del piano e coppie di numeri reali: il primo elemento della coppia lo chiameremo *ascissa* del punto P, il secondo elemento *ordinata* del punto P e la coppia stessa sarà detta coppia delle *coordinate* del punto P. Abitualmente indicheremo con x o x_P l'ascissa e con y o y_P l'ordinata e scriveremo $\mathsf{P}(x,y)$ o $\mathsf{P} = (x,y)$. Il sistema delle due rette che ci permette questa costruzione lo indicheremo con la notazione $\mathsf{O}xy$ e lo chiameremo un *sistema di coordinate cartesiane*.

Se le due rette r e s sono ortogonali, il sistema di coordinate è detto *ortogonale*. Se i due segmenti $\overline{\mathsf{OU}}$ sulle due rette sono congruenti, ovvero se si è scelta la stessa unità di misura, il sistema si dice *monometrico*. Useremo normalmente solo sistemi ortogonali e nella maggior parte dei casi sistemi monometrici. Non sempre però la scelta di sistemi monometrici è efficiente: se per esempio si vuol rappresentare in un grafico l'andamento annuale dei profitti di una impresa è chiaro che si useranno diverse unità di misura.

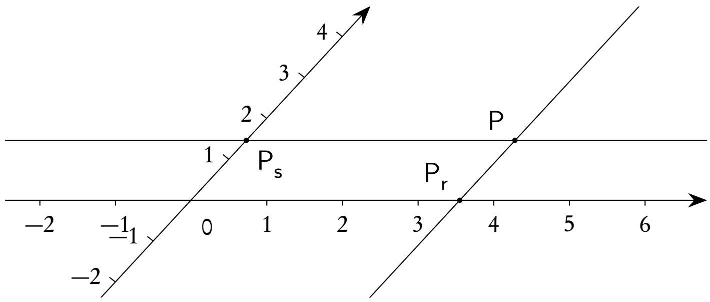


Figura 8.1.: *Coordinate cartesiane di un punto in un piano*

Tuttavia anche in casi molto comuni in matematica si usano sistemi non monometrici, e spesso, purtroppo, senza farne esplicita menzione: il caso più eclatante è quello delle funzioni trigonometriche con angoli misurati in gradi. Riservandoci di ritornare sull'argomento, segnaliamo qui che in un grafico come quello della figura 8.2 è evidente che l'unità di misura sui due assi è diversa (e di molto!) ed è questo l'aspetto che siamo abituati a immaginare quando pensiamo alla funzione seno.

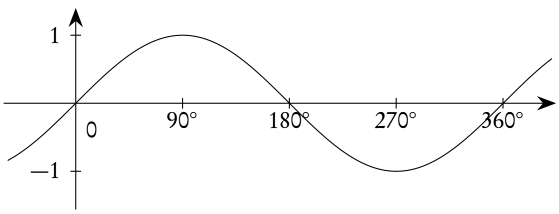


Figura 8.2.: *La funzione $f(x°) = \sin(x°)$, in un sistema non monometrico*

La stessa funzione in un sistema monometrico apparirebbe come nella figura 8.3.

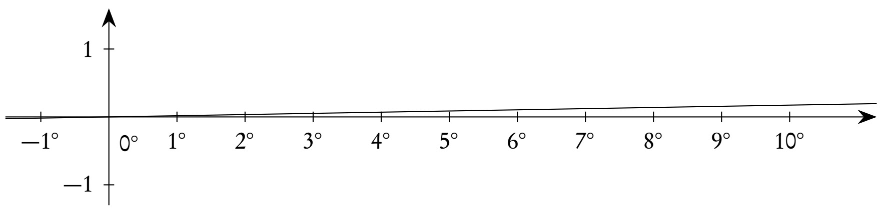


Figura 8.3.: *La funzione $f(x°) = \sin(x°)$, in un sistema monometrico*

La funzione seno ha "l'aspetto tradizionale" anche in un sistema monometrico *solo* se gli angoli sono misurati in radianti, come mostra la figura 8.4, ove si deve tenere conto che $\pi \approx 3.14$.

È estremamente importante, analizzando un grafico cartesiano, controllare sempre quali sono le unità di misura sui due assi, per leggere correttamente il grafico stesso. Si veda, a questo proposito, la figura 4.8, nella pagina 107.

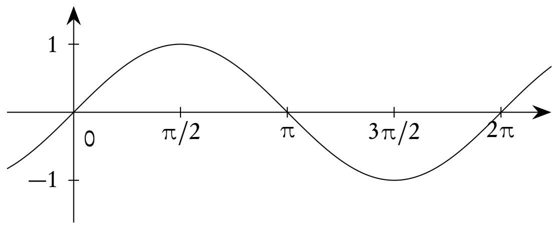


Figura 8.4.: *La funzione $f(x) = \sin(x)$, in un sistema monometrico*

Anche se in questo capitolo non avremo occasione di servircene, accenniamo brevemente alle coordinate nello spazio. Occorrerà introdurre una terna (ordinata) di rette non complanari, a due a due distinte, r, s, t, e concorrenti in un punto O. Su ciascuna di esse si introduce un sistema di ascisse con origine in O, abitualmente indicate, nell'ordine, con x, y, z. Dato un punto P dello spazio, mandiamo per esso il piano parallelo al piano yz: questo incontra l'asse x in un punto P_r; operiamo in modo analogo con i piani per P e paralleli xz e a xy, ottenendo altri due punti P_s e P_t. Le ascisse dei tre punti così ottenuti possono essere associate al punto P e viceversa ad ogni terna di numeri reali si può associare un unico punto P dello spazio, come fatto nel piano. In questo modo è possibile stabilire una corrispondenza biunivoca tra punti dello spazio ordinario e terne di numeri reali, che saranno chiamate le tre coordinate del punto P.

Questo passaggio, assolutamente "indolore", tra le coppie di numeri reali e le terne di numeri reali suggerisce possibili estensioni: se, dal punto di vista geometrico, non è possibile pensare a più di tre rette distinte, a due a due non complanari e incidenti in un punto O, dal punto di vista algebrico è invece possibile pensare a quaterne, ecc. di numeri reali e immaginare spazi a un numero qualunque di dimensioni, ma questo sarà oggetto dei futuri corsi universitari.

8.2. Le formule fondamentali

8.2.1. Punto medio di un segmento, baricentro di un triangolo

Dati due punti $P = (x_1, y_1)$ e $Q = (x_2, y_2)$, il punto medio M del segmento $\overline{PQ}$ che li unisce ha coordinate

$$x_M = \frac{x_1 + x_2}{2}, \quad y_M = \frac{y_1 + y_2}{2}, \tag{8.1}$$

e la stessa formula vale per la terza coordinata del punto medio nel caso dello spazio: le coordinate del punto medio sono la media delle coordinate degli estremi.

Dati poi tre punti $A = (x_1, y_1)$, $B = (x_2, y_2)$ e $C = (x_3, y_3)$, le coordinate[1] del *baricentro geometrico* G del triangolo ABC sono la media delle coordinate degli estremi.

$$x_G = \frac{x_1 + x_2 + x_3}{2}, \quad y_G = \frac{y_1 + y_2 + y_3}{2}, \tag{8.2}$$

[1] È molto importante distinguere tra il baricentro geometrico e il baricentro fisico di un triangolo. Considerando tre punti di massa uguale, il baricentro geometrico e quello fisico del triangolo individuato dai tre punti coincidono, ma la cosa non è più vera se i tre punti non hanno la stessa massa. Inoltre, cosa ben più inaspettata, se si considera un profilo triangolare omogeneo (cioè tre aste omogenee saldate a formare un triangolo), il baricentro geometrico e quello fisico *non* coincidono, a meno che il triangolo non sia equilatero.

e formula analoga per la terza coordinata nel caso di un triangolo nello spazio.

La dimostrazione sia della formula (8.1) che della (8.2) si basa sul teorema di Talete e costituisce un utile esercizio.

8.2.2. Distanza tra due punti

Siano dati due punti $\mathsf{P}=(x_1,y_1)$ e $\mathsf{Q}=(x_2,y_2)$ nel piano. Se si introduce un sistema cartesiano ortogonale monometrico la distanza $d(\mathsf{A},\mathsf{B})$ tra i due punti è data da

$$d(\mathsf{A},\mathsf{B})=|\overline{\mathsf{AB}}|=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}, \tag{8.3}$$

e, nel caso dello spazio, è sufficiente aggiungere come argomento della radice la quantità $(z_1-z_2)^2$.

Si tenga ben presente che, mentre le formule (8.1) e (8.2) per il punto medio e il baricentro di un triangolo valgono anche per sistemi non ortogonali e monometrici, la (8.3) vale *solo* per sistemi ortogonali monometrici.

La (8.3), nel caso di segmenti paralleli all'asse x (cioè con $y_1=y_2$) o all'asse y (cioè con $x_1=x_2$) assume, rispettivamente, la forma

$$y_1=y_2 \Rightarrow d(\mathsf{A},\mathsf{B})=|x_1-x_2|, \quad x_1=x_2 \Rightarrow d(\mathsf{A},\mathsf{B})=|y_1-y_2|. \tag{8.4}$$

Esempio 8.1. Si verifichi che il triangolo di vertici $\mathsf{A}(1,1)$, $\mathsf{B}(0,3)$ e $\mathsf{A}(3,2)$ è isoscele e rettangolo in A.

Calcoliamo le lunghezze dei tre lati, applicando la (8.3).

$$|\overline{\mathsf{AB}}|=\sqrt{(1-0)^2+(1-3)^2}=\sqrt{5};$$
$$|\overline{\mathsf{AC}}|=\sqrt{(1-3)^2+(1-2)^2}=\sqrt{5};$$
$$|\overline{\mathsf{BC}}|=\sqrt{(0-3)^2+(3-2)^2}=\sqrt{10}.$$

Dunque il triangolo è isoscele; per verificare che è rettangolo basta provare che vale il teorema di Pitagora, ovvero che

$$|\overline{\mathsf{AB}}|^2+|\overline{\mathsf{AC}}|^2=|\overline{\mathsf{BC}}|^2,$$

e questo è banalmente verificato perché $5+5=10$. Si veda la figura 8.5.

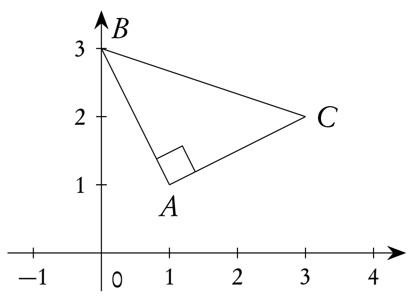


Figura 8.5.: *Triangolo rettangolo isoscele*

8.3. Luoghi geometrici e rappresentazione analitica

Si usa dare la seguente definizione.

Definizione 8.1. *Si chiama* luogo geometrico *un sottoinsieme di punti del piano o dello spazio definito mediante una proprietà caratteristica.*

Il nome *luogo geometrico* è entrato nell'uso comune: a nostro avviso non sarebbe strettamente necessario. Tuttavia la locuzione rende subito chiaro che l'oggetto di cui si tratta è un insieme di punti (del piano o dello spazio) e non un insieme generico e inoltre mette in evidenza che, di norma, siamo interessati a certe proprietà "geometriche" di questo insieme.

Alcuni esempi classici sono i seguenti.

— L'*asse* di un segmento è l'insieme dei punti del piano equidistante dagli estremi del segmento stesso.
— Una *circonferenza* è l'insieme dei punti del piano aventi da un punto fisso detto centro distanza uguale a un dato numero r.
— Un *cerchio* è l'insieme dei punti del piano aventi da un punto fisso detto centro distanza minore o uguale a un dato numero r.
— Una *superficie sferica* è l'insieme dei punti dello spazio aventi distanza assegnata da un punto fisso detto centro.
— La *bisettrice* di un angolo è l'insieme dei punti del piano equidistanti dai lati dell'angolo.
— Una *parabola* è l'insieme dei punti del piano equidistanti da un punto fisso F e da una retta fissa r, non passante per il punto F.

Se nel piano o nello spazio si è introdotto un sistema di coordinate cartesiane, i punti possono essere pensati come coppie o terne di numeri reali: si dice che il luogo geometrico è *rappresentato analiticamente* da un'equazione, una disequazione, o un sistema di equazioni e/o disequazioni se le coordinate di tutti e soli i punti del luogo soddisfano a quella equazione o disequazione o sistema. Tratteremo, in questo capitolo, alcuni dei luoghi più importanti nelle applicazioni. Alcuni esempi elementari, nel piano, sono i seguenti.

— Il semiasse positivo delle ascisse è rappresentato dal sistema

$$\begin{cases} y = 0 \\ x > 0 \end{cases},$$

in quanto tutti e soli i punti dell'asse delle ascisse hanno ordinata nulla e, per essere nel semiasse positivo devono avere ascissa positiva.

— I punti del primo quadrante sono rappresentati dal sistema

$$\begin{cases} x \geq 0 \\ y \geq 0 \end{cases}.$$

— I punti della bisettrice del primo e terzo quadrante (sistema ortogonale monometrico) sono rappresentati dall'equazione

$$x = y,$$

come si deduce facilmente tenendo conto della proprietà sopra ricordata della bisettrice di un angolo.

Una volta determinate le equazioni di un luogo sarà facile verificare se un punto appartiene o no al luogo: basterà controllare se le sue coordinate soddisfano o no le equazioni trovate.

In generale i luoghi che considereremo saranno costituiti da curve (nel senso intuitivo del termine) del piano rappresentate da una sola equazione in due incognite che scriveremo genericamente come

$$f(x,y)=0. \tag{8.5}$$

In questo caso, se un punto appartiene al luogo si dice che la curva passa per il punto: casi importanti saranno la retta, la parabola, la circonferenza e in genere le coniche. Sarà anche possibile (e di grande interesse applicativo) trovare i punti che appartengono contemporaneamente a due o più luoghi. Vediamo alcuni esempi.

Esempio 8.2. Dato il luogo di equazione $x^2+y^2-1=0$, il punto $(3/5, 4/5)$ appartiene al luogo, il punto $(1,1)$ non vi appartiene.

Esempio 8.3. Dati i luoghi di equazione $x^2+y^2-1=0$ e $x^2+y=0$ per trovare i punti comuni si risolvere il sistema

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

Si ottiene

$$\mathrm{P}_1=\left(\sqrt{\frac{\sqrt{5}-1}{2}}, \frac{1-\sqrt{5}}{2}\right), \mathrm{P}_2=\left(-\sqrt{\frac{\sqrt{5}-1}{2}}, \frac{1-\sqrt{5}}{2}\right)$$

La rappresentazione grafica di questi luoghi, che studieremo a breve, e delle due intersezioni trovate è proposta nella figura 8.6.

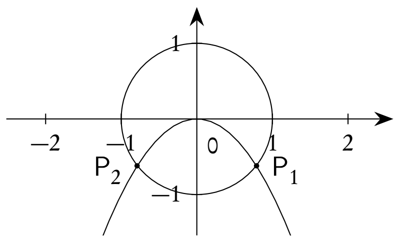


Figura 8.6.: *Luoghi geometrici e loro intersezione*

8.4. Cambiamenti di coordinate

Ci limiteremo a considerare, nell'ambito dei sistemi cartesiani ortogonali monometrici, solo le traslazioni, le rotazioni e le rototraslazioni, senza cambio dell'unità di misura sugli assi.

Siano $\mathrm{O}xy$, $\mathrm{O}'x'y'$ due sistemi di coordinate, cartesiane ortogonali monometriche, nel piano; li chiameremo "vecchio" e "nuovo" sistema di coordinate rispettivamente e indicheremo con $\mathrm{O}'(a,b)$ le coordinate di O' rispetto ad $\mathrm{O}xy$, e (anche se non le useremo nelle formule che utilizzeremo) con $\mathrm{O}(a',b')$ le coordinate di O rispetto ad $\mathrm{O}'x'y'$. Se P è un punto del piano, indicheremo con (x,y) e con (x',y') le coordinate di P rispetto ai due sistemi di coordinate. Se è compresa anche una rotazione degli assi, indicheremo con α l'angolo individuato dai due semiassi positivi x e x', angolo che indica di quanto il nuovo sistema è ruotato rispetto al vecchio. Potremo ritenere $0 \le \alpha < 2\pi$, ma nulla cambia se invece pensiamo $-\pi \le \alpha < \pi$.

8.4.1. Traslazione degli assi

È facile provare che le formule per la traslazione degli assi sono

$$\begin{cases} x = x' + a \\ y = y' + b \end{cases}, \tag{8.6}$$

che esprimono le vecchie coordinate in funzione delle nuove, oppure le loro inverse

$$\begin{cases} x' = x - a \\ y' = y - b \end{cases}, \tag{8.7}$$

che esprimono le nuove coordinate in funzione delle vecchie.

Non dimostreremo queste formule per le traslazioni, segnalando comunque che il ragionamento per ottenerle è elementare e può essere ricavato semplicemente considerando la figura 8.7.

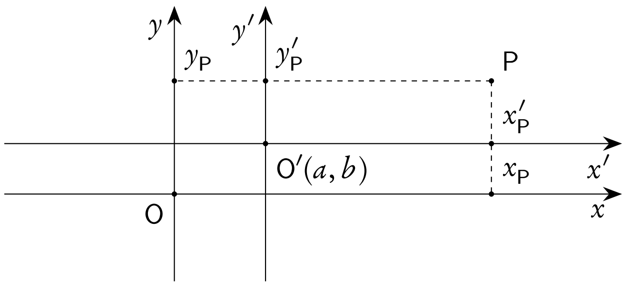


Figura 8.7.: *Traslazione degli assi cartesiani*

Le formule (8.6) servono per ottenere le vecchie coordinate di un punto P, note che siano le nuove e, soprattutto, per passare dall'equazione di una curva data nelle vecchie coordinate all'equazione della stessa curva espressa nelle nuove coordinate. Le formule (8.7) servono per ottenere le nuove coordinate di un punto P, note che siano le vecchie, oppure per passare dall'equazione di una curva data nelle nuove coordinate all'equazione della stessa curva espressa nelle vecchie coordinate.

Esempio 8.4. Data la curva di equazione cartesiana

$$x^2 - 2y^2 - 2x + 4y - 3 = 0,$$

trovare l'equazione della stessa curva in un nuovo sistema di coordinate di centro $O'(1,1)$.

Le (8.6) si scrivono, in questo caso,

$$\begin{cases} x = x' + 1 \\ y = y' + 1 \end{cases}.$$

Sostituendo questi valori nell'equazione della curva si ottiene

$$(x'+1)^2 - 2(y'+1)^2 - 2(x'+1) + 4(y'+1) - 3 = 0 \quad \Rightarrow \quad (x')^2 - 2(y')^2 - 2 = 0.$$

La figura 8.8 mostra la curva e i due sistemi di coordinate. Come vedremo si tratta di un'iperbole, e la traslazione di assi trasforma l'equazione nella sua forma canonica.

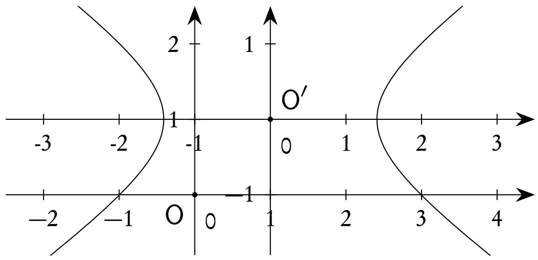


Figura 8.8.: *Una curva "vista" da due diversi sistemi di coordinate*

8.4.2. Rotazione degli assi

Le formule per la rotazione degli assi sono

$$\left\{\begin{array}{l} x = x'\cos\alpha - y'\sin\alpha \\ y = x'\sin\alpha + y'\cos\alpha \end{array}\right. , \tag{8.8}$$

che forniscono le vecchie coordinate in funzione delle nuove e le

$$\left\{\begin{array}{l} x' = x\cos\alpha + y\sin\alpha \\ y' = -x\sin\alpha + y\cos\alpha \end{array}\right. , \tag{8.9}$$

che forniscono le nuove coordinate in funzione delle vecchie. Si noti come il passaggio dalle (8.8) alle (8.9) possa avvenire, oltreché per risoluzione diretta[(2)], anche semplicemente per simmetria, scrivendo (x', y') al posto di (x, y) e $-\alpha$ al posto di α.

Queste formule possono essere ricavate esaminando la figura 8.9 e utilizzando i teoremi sui triangoli rettangoli, che saranno presentati nel capitolo 11 sulla trigonometria.

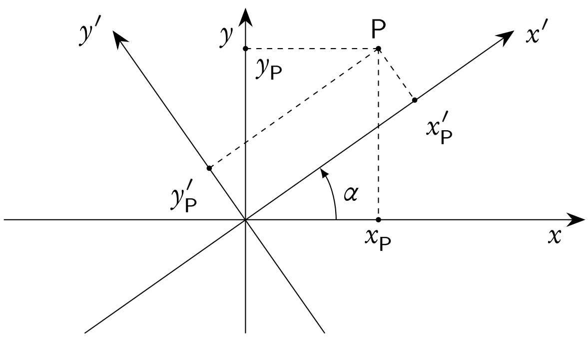


Figura 8.9.: *Rotazione degli assi*

In analogia con il caso della traslazione, segnaliamo che le formule (8.8) servono per ottenere le vecchie coordinate di un punto P, note che siano le nuove e, soprattutto, per passare dall'equazione di una curva data nelle vecchie coordinate all'equazione della stessa curva espressa nelle nuove coordinate. Le formule

[2] Basta moltiplicare la prima per $\cos\alpha$, la seconda per $\sin\alpha$ e sommare: si ricava x'; se invece si moltiplica la prima per $\sin\alpha$ e la seconda per $\cos\alpha$ da una successiva sottrazione si ricava y'.

(8.9) servono per ottenere le nuove coordinate di un punto P, note che siano le vecchie, oppure per passare dall'equazione di una curva data nelle nuove coordinate all'equazione della stessa curva espressa nelle vecchie coordinate.

Esempio 8.5. Data la curva di equazione cartesiana

$$x^2 - y^2 - 1 = 0,$$

trovare la sua equazione in un nuovo sistema ruotato di $\pi/4$ e con la stessa origine.

Le (8.8) si scrivono, in questo caso,

$$\begin{cases} x = x' \dfrac{\sqrt{2}}{2} - y' \dfrac{\sqrt{2}}{2} \\ x = x' \dfrac{\sqrt{2}}{2} + y' \dfrac{\sqrt{2}}{2} \end{cases}.$$

Sostituendo questi valori nell'equazione della curva si ottiene

$$\left(x' \frac{\sqrt{2}}{2} - y' \frac{\sqrt{2}}{2}\right)^2 - \left(x' \frac{\sqrt{2}}{2} + y' \frac{\sqrt{2}}{2}\right)^2 - 1 = 0 \quad \Rightarrow \quad 2x'y' + 1 = 0.$$

8.5. La retta nel piano cartesiano

8.5.1. Generalità

Vale il seguente teorema riguardante l'equazione di una retta nel piano.

Teorema 8.2. *Ogni retta del piano è rappresentata da un'equazione di primo grado in due incognite, ovvero da un'equazione del tipo*

$$ax + by + c = 0, \quad (a, b) \neq (0, 0). \tag{8.10}$$

Viceversa ogni equazione del tipo (8.10) *rappresenta una retta del piano.*

Si tenga ben presente che i valori di a, b, c non sono univocamente individuati dalla retta. Se infatti $ax + by + c = 0$ è l'equazione di una retta, $kax + kby + kc = 0$, per ogni $k \neq 0$, è ancora l'equazione della stessa retta. In realtà, dovendo essere $(a, b) \neq (0, 0)$, dividendo ambo i membri della (8.10) per a, oppure per b, l'equazione stessa può essere scritta in una delle forme

$$x + \frac{b}{a} y + \frac{c}{a} = 0, \quad \frac{a}{b} x + y + \frac{c}{b} = 0,$$

ovvero

$$x + b'y + c' = 0, \quad a''x + y + c'' = 0,$$

e ora i valori a', b', oppure a'', b'', sono univocamente individuati dalla retta. Questa osservazione è di grande importanza: se si vuol determinare l'equazione di una retta occorre determinare 2 parametri e quindi saranno necessarie 2 condizioni tra di loro indipendenti. Esamineremo in seguito le situazioni più comuni.

Sono di facile verifica le seguenti proprietà.

- Se $a = 0$ e $b \neq 0$, la (8.10) può essere scritta nella forma $y = h$, con, $h = -c/b$, e rappresenta una retta parallela all'asse delle ascisse; se poi $c = 0$ si ha $y = 0$, che è l'equazione dell'asse delle ascisse.
- Se $b = 0$ e $a \neq 0$, la (8.10) può essere scritta nella forma $x = k$, con, $k = -c/a$, e rappresenta una retta parallela all'asse delle ordinate; se poi $c = 0$ si ha $x = 0$, che è l'equazione dell'asse delle ordinate.
- Se $c = 0$, la (8.10) rappresenta una retta passante per l'origine.

Il caso $b \neq 0$ (che dunque si riferisce a una retta *non* parallela all'asse delle ordinate o, come si usa dire, *non verticale*) è particolarmente importante. In questo caso la (8.10) si può porre nella forma

$$y = -\frac{a}{b}x - \frac{c}{a},$$

di solito scritta come

$$y = mx + q, \tag{8.11}$$

e la retta può essere pensata come il grafico di una funzione polinomiale di primo grado. Continuano a valere tutte le considerazioni già fatte a questo proposito nel paragrafo 4.8.1 del capitolo 4.

La generica equazione (8.10) è anche detta *equazione implicita*, mentre la (8.11), valida, lo ripetiamo, per rette non verticali, è detta *equazione esplicita*.

Ricordiamo che il numero m nella (8.11) si chiama *coefficiente angolare* o *pendenza* della retta r, mentre il numero q si chiama ordinata all'origine. Data una retta r non parallela agli assi, se α è l'angolo avente vertice nel punto A di intersezione tra r e l'asse x e avente per lati la semiretta positiva di origine A dell'asse x e la semiretta superiore di origine A della retta r, si ha

$$m = \tan\alpha. \tag{8.12}$$

La (8.12) continua a valere anche per rette "orizzontali", purché per esse si assuma $\alpha = 0$ oppure $\alpha = \pi$.

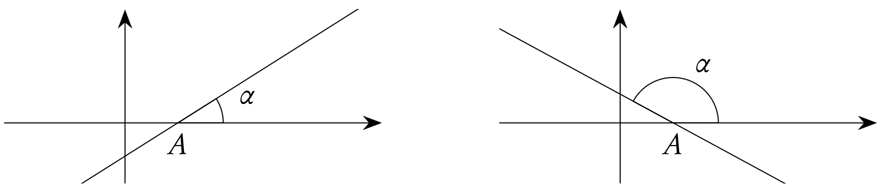


Figura 8.10.: *Coefficiente angolare di una retta:* $m = \tan\alpha$

Atteso il significato del coefficiente angolare, è chiaro che due rette r ed r', di equazioni esplicite $y = mx + q$ ed $y = m'x + q'$ rispettivamente, sono *parallele* se e solo se $m = m'$. È facile provare che se le rette sono invece date in forma implicita, $ax + by + c = 0$ e $a'x + b'y + c'$ rispettivamente, esse sono parallele se e solo se

$$\exists k \neq 0 \text{ tale che } (a' = ka) \wedge (b' = kb). \tag{8.13}$$

Se inoltre $c' = kc$ le due rette sono coincidenti. Se infatti $a = 0$ r è parallela all'asse x e quindi deve essere anche $a' = 0$: basterà prendere $k = b'/b$. Analogo se $b = 0$. Se i quattro coefficienti a, b, a', b' sono tutti diversi da zero, le rette si possono mettere in forma implicita e la condizione di parallelismo diventa

$$-\frac{a}{b} = -\frac{a'}{b'} \Rightarrow \frac{a'}{a} = \frac{b'}{b} = k.$$

Per quanto riguarda la condizione di perpendicolarità, vale il seguente teorema.

Teorema 8.3. *Due rette* r *ed* r'*, di equazioni esplicite* $y = mx + q$ *ed* $y = m'x + q'$ *rispettivamente, sono* perpendicolari *se e solo se* $mm' = -1$*. Due rette* r *ed* r'*, di equazioni implicite* $ax + by + c = 0$ *e* $a'x + b'y + c'$ *rispettivamente, sono perpendicolari se e solo se* $aa' + bb' = 0$*.*

Dimostrazione. Proponiamo la dimostrazione della prima parte del teorema, perché utile anche come esercizio. Osserviamo preliminarmente che nessuna delle due rette deve essere parallela all'asse x, altrimenti l'altra sarebbe verticale e dunque non si potrebbe mettere in forma esplicita. Basterà limitarsi a considerare le parallele $y = mx$ e $y = m'x$ alle due rette date per l'origine. Se le due rette sono ortogonali una giacerà nel primo e terzo quadrante, l'altra nel secondo e quarto. Con riferimento alla figura 8.11, consideriamo i punti $P(1,m)$, $Q(1,m')$ e $H(1,0)$.

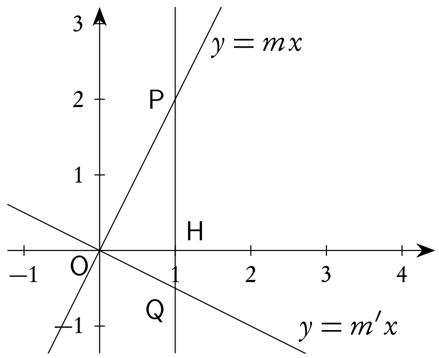


Figura 8.11.: *Perpendicolarità di due rette*

Le due rette sono ortogonali se e solo se l'angolo $\mathsf{P\hat{O}Q}$ è retto. Per il secondo teorema di Euclide si deve avere $|\overline{\mathsf{OH}}|^2 = |\overline{\mathsf{PH}}| \cdot |\overline{\mathsf{HQ}}|$, ovvero $|mm'| = 1$. Poiché m ed m' devono essere discordi, si conclude che $mm' = -1$. Se, viceversa, $|mm'| = 1$ vale il secondo teorema di Euclide e l'angolo $\mathsf{P\hat{O}Q}$ è retto. □

Se una retta non è parallela a nessuno dei due assi e non passa per l'origine, i coefficienti a, b, c sono tutti non nulli. Le intersezioni di una retta siffatta con gli assi sono

$$\left(-\frac{c}{a}, 0\right) = (p, 0) \quad \text{e} \quad \left(0, -\frac{c}{b}\right) = (0, q).$$

I numeri p e q si chiamano *intercette* della retta sugli assi. Si ha

$$ax + by + c = 0 \quad \Rightarrow \quad \frac{x}{-{}^{c}\!/_{a}} + \frac{y}{-{}^{c}\!/_{b}} = 1$$

ovvero

$$\frac{x}{p} + \frac{y}{q} = 1, \tag{8.14}$$

detta *equazione segmentaria* della retta.

Vale anche il seguente teorema concernente la determinazione della distanza di un punto da una retta.

Teorema 8.4 (Distanza di un punto da una retta)**.** *Siano dati un punto* $\mathsf{P}(x_0, y_0)$ *e una retta* r *di equazione implicita* $ax + by + c = 0$*. La distanza di* P_0 *da* r *è data da*

$$d(\mathsf{P}_0, \mathsf{r}) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}. \tag{8.15}$$

Si noti che, se $\mathsf{P}_0 \in \mathsf{r}$, la distanza fornita dalla formula (8.15) è nulla, come deve essere.

8.5.2. Come determinare l'equazione di una retta

Le situazioni più comuni in cui è richiesta la determinazione dell'equazione di una retta sono:

1. retta per un punto e di direzione nota;
2. retta per due punti.

Per quanto riguarda il primo caso sia $\mathsf{P}(x_0, y_0)$ il punto per cui deve passare la retta. Se è nota la direzione, e se la retta non è verticale, significa che ne è noto il coefficiente angolare m. L'equazione è allora

$$y - y_0 = m(x - x_0). \tag{8.16}$$

Se invece la retta è verticale essa avrà equazione $x = x_0$. Se la retta cercata deve essere parallela alla retta di equazione implicita $ax + by + c = 0$ si potrà direttamente scriverne l'equazione come

$$a(x - x_0) + b(y - y_0) = 0. \tag{8.17}$$

Se invece la retta cercata deve essere perpendicolare alla retta di equazione implicita $ax + by + c = 0$ si potrà direttamente scriverne l'equazione come

$$b(x - x_0) - a(y - y_0) = 0. \tag{8.18}$$

In ogni caso è sempre possibile fare riferimento alla forma (8.16), determinando prima, se del caso, il coefficiente angolare corretto.

Per quanto riguarda invece il secondo caso siano $\mathsf{P}(x_1, y_1)$ e $\mathsf{Q}(x_2, y_2)$ i due punti per cui deve passare la retta. L'equazione sarà

$$(y - y_1)(x_2 - x_1) = (x - x_1)(y_2 - y_1). \tag{8.19}$$

La forma (8.19) della retta per due punti è preferibile rispetto a quella tradizionale sotto forma di frazione, mostrata nell'equazione (8.20), in quanto comprende anche i casi di rette parallele ad uno degli assi.

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}. \tag{8.20}$$

8.5.3. Un esempio conclusivo

Proponiamo un esempio, abbastanza complesso, che utilizza la gran parte delle proprietà della retta nel piano cartesiano.

Esempio 8.6. Dato il triangolo di vertici $\mathsf{A}(0,0)$, $\mathsf{B}(0,8)$ e $\mathsf{C}(8,6)$ determinarne il baricentro, il circocentro, l'ortocentro, l'incentro e i tre ex-centri, verificando che i primi tre punti sono allineati. Verificare inoltre le proprietà più rimarchevoli di questi punti.

Per il baricentro (incontro delle mediane) e il circocentro (incontro degli assi) ci servono i punti medi di due dei tre lati, per esempio M di $\overline{\mathsf{AB}}$ e N di $\overline{\mathsf{AC}}$. Si ha

$$x_{\mathsf{M}} = \frac{0+0}{2} = 0,\ y_{\mathsf{M}} = \frac{0+8}{2} = 4; \quad x_{\mathsf{N}} = \frac{0+8}{2} = 4,\ y_{\mathsf{N}} = \frac{0+6}{2} = 3.$$

In vista di calcoli futuri determiniamo anche il punto medio S di $\overline{\mathsf{BC}}$.

$$x_{\mathsf{S}} = \frac{0+8}{2} = 4,\ y_{\mathsf{S}} = \frac{8+6}{2} = 7.$$

Le mediane BN e CM avranno equazione, rispettivamente,

$$(y-8)(4-0) = (x-0)(3-8) \Rightarrow 5x+4y-32 = 0; \quad x-4y+16 = 0.$$

Il baricentro G si potrà trovare come intersezione delle due mediane:

$$\begin{cases} 5x+4y-32 = 0 \\ x-4y+16 = 0 \end{cases} \Rightarrow \quad \mathsf{G} = \left(\frac{8}{3}, \frac{14}{3}\right).$$

Naturalmente questo punto si poteva anche trovare usando la formula (8.2), con lo stesso risultato.

Per trovare il circocentro ci servono due dei tre assi (perpendicolari nei punti medi) dei lati del triangolo. La retta per M e ortogonale ad $\overline{\mathsf{AB}}$ è chiaramente parallela all'asse x e ha equazione $y = 4$. Per determinare la retta per N e ortogonale ad $\overline{\mathsf{AC}}$ usiamo il coefficiente angolare già trovato:

$$y-3 = -\frac{4}{3}(x-4) \Rightarrow 4x+3y-25 = 0.$$

Il circocentro Q si trova sull'intersezione di queste due rette:

$$\begin{cases} y = 4 \\ 4x+3y-25 = 0 \end{cases} \Rightarrow \quad \mathsf{Q} = \left(\frac{13}{4}, 4\right).$$

Per trovare gli assi dei lati si poteva anche usare la proprietà che l'asse di un segmento è il luogo dei punti equidistanti dagli estremi. Ritroviamo, per esercizio, l'equazione dell'asse di $\overline{\mathsf{AC}}$ in questo modo. Detto $\mathsf{P}(s,t)$ un generico punto dell'asse si deve avere

$$|\overline{\mathsf{PA}}| = |\overline{\mathsf{PC}}| \Rightarrow \sqrt{(s-0)^2+(t-0)^2} = \sqrt{(s-8)^2+(t-6)^2}.$$

Quadrando e semplificando si trova $4s+3t-25 = 0$, ovvero la stessa equazione di prima, previa sostituzione(3) di (s,t) con (x,y).

[3] In molti casi è opportuno, nella ricerca di un luogo, indicare inizialmente con (s,t) la coppia di coordinate del generico punto del luogo, sostituendo alla fine (s,t) con (x,y).

Per trovare l'ortocentro ci servono due delle tre altezze. La retta per C e ortogonale ad $\overline{\mathsf{AB}}$ è chiaramente parallela all'asse x e ha equazione $y = 6$. Per determinare la retta per B e ortogonale ad $\overline{\mathsf{AC}}$ troviamo prima il coefficiente angolare di AC, usando la formula (4.15):

$$m_{\mathsf{AC}} = \frac{6-0}{8-0} = \frac{3}{4}.$$

Il coefficiente angolare della perpendicolare sarà $-4/3$ e l'equazione della perpendicolare

$$y - 8 = -\frac{4}{3}(x - 0) \Rightarrow 4x + 3y - 24 = 0.$$

L'ortocentro H si trova sull'intersezione di queste due rette:

$$\begin{cases} y = 6 \\ 4x + 3y - 24 = 0 \end{cases} \quad \Rightarrow \quad \mathsf{H} = \left(\frac{3}{2}, 6\right).$$

Per verificare che questi tre punti sono allineati[4] possiamo trovare la retta per due di essi e verificare che il terzo punto vi appartiene.

$$\mathsf{HQ}:\ (y-6)\left(\frac{13}{4} - \frac{3}{2}\right) = \left(x - \frac{3}{2}\right)(4-6) \Rightarrow 8x + 7y - 54 = 0.$$

È immediato provare che il punto G appartiene a questa retta. La figura 8.12 illustra la situazione descritta.

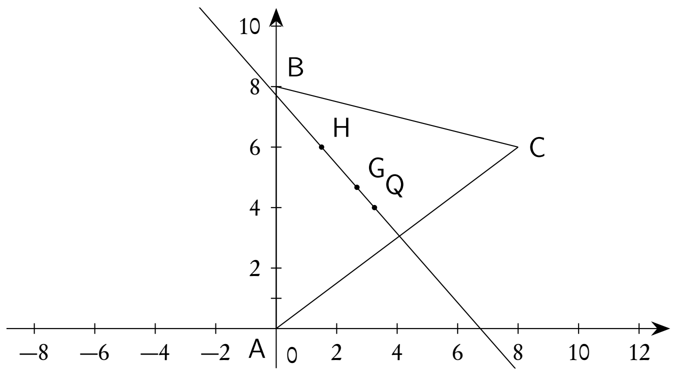


Figura 8.12.: *Baricentro, circocentro, ortocentro e retta di Eulero*

Passiamo ora alla determinazione dell'incentro, punto di intersezione delle bisettrici degli angoli interni. Ne basterebbero due, naturalmente, ma in vista della determinazione anche degli ex-centri le determineremo tutte. Possiamo usare la proprietà delle bisettrici di essere luogo di punti equidistanti dai lati dell'angolo. Le equazioni dei tre lati del triangolo, determinate con metodo ormai noto, sono:

$$\mathsf{AB}: x = 0; \quad \mathsf{AC}: 3x - 4y = 0; \quad \mathsf{BC}: x + 4y - 32 = 0.$$

[4] La retta che li contiene si chiama *retta di Eulero*.

Per le bisettrici degli angoli individuati dalle rette AB e AC, detto $\mathsf{P}(s,t)$ un loro generico punto, si deve avere:

$$\frac{|s|}{\sqrt{1^2+0^2}} = \frac{|3s-4t|}{\sqrt{3^2+4^2}} \Rightarrow 5|s| = |3s-4t| \Rightarrow 5s = \pm(3s-4t).$$

Si ottengono le due rette $2x-y=0$ e $x+2y=0$, tra loro ortogonali, come deve essere. La prima è bisettrice dell'angolo interno in $\widehat{A}$ (perché ha coefficiente angolare positivo), la seconda dei due angoli esterni in $\widehat{A}$ (perché ha coefficiente angolare negativo).

In maniera analoga per le bisettrici degli angoli individuati dalle rette AB e BC si deve avere:

$$\frac{|s|}{\sqrt{1^2+0^2}} = \frac{|s+4t-32|}{\sqrt{1^2+4^2}} \Rightarrow \sqrt{17}s = \pm(s+4t-32),$$

da cui si ottengono le due rette

$$\left(\sqrt{17}+1\right)x+4y-32=0 \quad \text{e} \quad \left(\sqrt{17}-1\right)x-4y+32=0.$$

Anche queste due sono tra loro perpendicolari e la prima è bisettrice dell'angolo interno in $\widehat{B}$ (perché ha coefficiente angolare negativo), la seconda dei due angoli esterni in $\widehat{B}$ (perché ha coefficiente angolare positivo).

Infine per le bisettrici degli angoli individuati dalle rette AC e BC si deve avere:

$$\frac{|3s-4t|}{\sqrt{3^2+4^2}} = \frac{|s+4t-32|}{\sqrt{1^2+4^2}} \Rightarrow \sqrt{17}(3s-4t) = \pm 5(s+4t-32),$$

da cui si ottengono le due rette

$$\left(3\sqrt{17}-5\right)x-\left(4\sqrt{17}+20\right)y+160=0 \quad \text{e} \quad \left(3\sqrt{17}+5\right)x-\left(4\sqrt{17}-20\right)y-160=0.$$

Anche queste due sono tra loro perpendicolari e la prima è bisettrice dell'angolo interno in $\widehat{C}$ (perché ha coefficiente angolare positivo), la seconda dei due angoli esterni in $\widehat{C}$ (perché ha coefficiente angolare negativo).

L'incentro I si trova sull'intersezione di due delle tre bisettrici interne.

$$\begin{cases} 2x-y=0 \\ \left(\sqrt{17}+1\right)x+4y-32=0 \end{cases} \quad \Rightarrow \quad \mathsf{I} = \left(\frac{32}{\sqrt{17}+9}, \frac{64}{\sqrt{17}+9}\right).$$

Per trovare i tre ex-centri basta intersecare due bisettrici di angoli esterni, oppure una di angolo esterno e una di angolo interno, opportunamente scelto. Cominciamo da E_1, ex-centro "dalla parte del lato $\overline{\mathsf{AC}}$".

$$\begin{cases} x+2y=0 \\ \left(\sqrt{17}+1\right)x+4y-32=0 \end{cases} \quad \Rightarrow \quad \mathsf{E}_1 = \left(\frac{32}{\sqrt{17}-1}, \frac{-16}{\sqrt{17}-1}\right).$$

Passiamo poi a E_2, ex-centro "dalla parte del lato $\overline{\mathsf{AB}}$".

$$\begin{cases} x+2y=0 \\ \left(3\sqrt{17}-5\right)x-\left(4\sqrt{17}+20\right)y+160=0 \end{cases} \quad \Rightarrow \quad \mathsf{E}_2 = \left(\frac{-32}{\sqrt{17}+1}, \frac{16}{\sqrt{17}+1}\right).$$

Troviamo infine E_3, ex-centro “dalla parte del lato $\overline{\mathsf{BC}}$”.

$$\begin{cases} 2x-y=0 \\ \left(3\sqrt{17}+5\right)x-\left(4\sqrt{17}-20\right)y-160=0 \end{cases} \quad\Rightarrow\quad \mathsf{E}_3=\left(\frac{32}{9-\sqrt{17}},\frac{64}{9-\sqrt{17}}\right).$$

Ci resta da verificare che, per i vari punti trovati, valgono le note proprietà geometriche. Cominciamo dal baricentro G: esso divide le tre mediane in parti tali che quella che contiene il vertice è doppia dell'altra. Si ha

$$|\overline{\mathsf{AG}}|=\sqrt{\left(0-\frac{8}{3}\right)^2+\left(0-\frac{14}{3}\right)^2}=\frac{2}{3}\sqrt{65};\qquad |\overline{\mathsf{SG}}|=\sqrt{\left(4-\frac{8}{3}\right)^2+\left(7-\frac{14}{3}\right)^2}=\frac{1}{3}\sqrt{65};$$

$$|\overline{\mathsf{BG}}|=\sqrt{\left(0-\frac{8}{3}\right)^2+\left(8-\frac{14}{3}\right)^2}=\frac{2}{3}\sqrt{41};\qquad |\overline{\mathsf{NG}}|=\sqrt{\left(4-\frac{8}{3}\right)^2+\left(3-\frac{14}{3}\right)^2}=\frac{1}{3}\sqrt{41};$$

$$|\overline{\mathsf{CG}}|=\sqrt{\left(8-\frac{8}{3}\right)^2+\left(6-\frac{14}{3}\right)^2}=\frac{4}{3}\sqrt{17};\qquad |\overline{\mathsf{SM}}|=\sqrt{\left(0-\frac{8}{3}\right)^2+\left(4-\frac{14}{3}\right)^2}=\frac{2}{3}\sqrt{17}.$$

Il circocentro Q, punto di intersezione degli assi, è centro della circonferenza circoscritta al triangolo: per provare ciò basterà provare che Q è equidistante dai tre vertici del triangolo. Si ha

$$|\overline{\mathsf{QA}}|=\sqrt{\left(\frac{13}{4}-0\right)^2+(4-0)^2}=\frac{5}{4}\sqrt{17};$$

$$|\overline{\mathsf{QB}}|=\sqrt{\left(\frac{13}{4}-0\right)^2+(4-8)^2}=\frac{5}{4}\sqrt{17};$$

$$|\overline{\mathsf{QC}}|=\sqrt{\left(\frac{13}{4}-8\right)^2+(4-6)^2}=\frac{5}{4}\sqrt{17}.$$

Per quanto riguarda l'ortocentro non si studiano, a livello elementare, speciali proprietà. Tuttavia, tra le molte proprietà di questo punto, ne segnaliamo una, invitando a farne la verifica per il triangolo in esame: il triangolo che ha per vertici i piedi delle tre altezze si chiama *triangolo ortico*. Ebbene, l'ortocentro di un triangolo è l'incentro del suo triangolo ortico.

L'incentro è centro della circonferenza inscritta nel triangolo: per verificare questo basterà provare che l'incentro I è equidistante dai tre lati del triangolo. Si ha:

$$d(\mathsf{I},\mathsf{AB})=|x_{\mathsf{I}}|=\frac{32}{\sqrt{17}+9};$$

$$d(\mathsf{I},\mathsf{AC})=\frac{\left|3\dfrac{32}{\sqrt{17}+9}-4\dfrac{64}{\sqrt{17}+9}\right|}{\sqrt{3^2+4^2}}=\frac{32}{\sqrt{17}+9};$$

$$d(\mathsf{I},\mathsf{BC})=\frac{\left|\dfrac{32}{\sqrt{17}+9}+4\dfrac{64}{\sqrt{17}+9}-32\right|}{\sqrt{1^2+4^2}}=\frac{32}{\sqrt{17}+9}.$$

La proprietà principale di ciascuno degli ex-centri è di essere centro di una delle circonferenze ex-iscritte, ovvero tangenti a un lato e al prolungamento degli altri due. Per verificare questa proprietà basterà provare che ciascuno degli ex-centri è equidistante dalle tre rette di interesse. Proviamolo, per esempio, per E_2. Si ha

$$\begin{aligned} d(\mathsf{E}_2,\mathsf{AB}) &= |x_{\mathsf{E}_2}| = \frac{32}{\sqrt{17}+1}\,; \\ d(\mathsf{E}_2,\mathsf{AC}) &= \frac{\left|3\,\dfrac{-32}{\sqrt{17}+1} - 4\,\dfrac{16}{\sqrt{17}+1}\right|}{\sqrt{3^2+4^2}} = \frac{32}{\sqrt{17}+1}\,; \\ d(\mathsf{E}_2,\mathsf{BC}) &= \frac{\left|\dfrac{-32}{\sqrt{17}+1} + 4\,\dfrac{16}{\sqrt{17}+1} - 32\right|}{\sqrt{1^2+4^2}} = \frac{32}{\sqrt{17}+1}\,. \end{aligned}$$

La figura 8.13 illustra la situazione descritta.

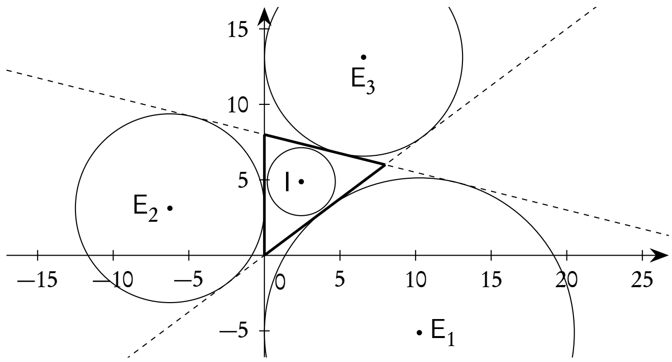


Figura 8.13.: *Incentro ed ex-centri di un triangolo*

Una proprietà interessante che lega l'incentro e i tre ex-centri è che ciascuno di questi quattro punti è ortocentro del triangolo che ha per vertici gli altri tre: la dimostrazione geometrica è particolarmente semplice; se ne faccia la verifica analitica per il triangolo in esame.

8.5.4. Famiglie di curve

Se $f(x,y)=0$ e $g(x,y)=0$ sono le equazioni di due curve, Φ e Γ del piano, allora al variare di λ e μ in $\mathbb{R}$ le equazioni

$$\lambda f(x,y)+\mu g(x,y)=0, \quad (\lambda,\mu)\neq(0,0), \tag{8.21}$$

rappresentano una famiglia $\mathcal{F}$ di curve cui appartengono sia Φ che Γ: la prima per $\lambda=1$ e $\mu=0$, la seconda per $\lambda=0$ e $\mu=1$. Inoltre se un punto P appartiene a entrambe le due curve Φ e Γ, allora appartiene a tutte le curve della famiglia. Le due curve Φ e Γ si possono chiamare le curve base della famiglia. Un'equazione del tipo (8.21) si chiama una *combinazione lineare* delle equazioni delle due curve Φ e Γ.

Se nella (8.21) supponiamo $\lambda \neq 0$, e poniamo $k = \mu/\lambda$, otteniamo

$$f(x,y) + k g(x,y) = 0, \quad k \in \mathbb{R}. \tag{8.22}$$

Al variare di k le equazioni (8.22) forniscono un insieme $\mathcal{G}$ di curve, tutte appartenenti all'insieme $\mathcal{F}$, ma è chiaro che $\mathcal{F}$ e $\mathcal{G}$ non coincidono, almeno se Φ e Γ sono distinte: la curva Γ non appartiene a $\mathcal{G}$, in quanto l'equazione $g(x,y) = 0$ non si ottiene dalle (8.22) per nessun valore di k. Tuttavia questa è l'unica differenza tra $\mathcal{F}$ e $\mathcal{G}$. Se poi si tiene conto che, nella (8.21) la curva Γ si ottiene per $\lambda = 0$ (e quindi $\mu \neq 0$, per esempio $\mu = 1$), si può pensare che il k che compare nella (8.22), al tendere di λ a 0 tenda all'infinito: per questo motivo si usa dire che che, nella (8.22), la curva Γ corrisponde al valore $k = \infty$. Nelle applicazioni che considereremo, la rappresentazione di una famiglia di curve con una equazione "ad un solo parametro", come la (8.22) è molto più comoda che non la rappresentazione con una equazione "a due parametri" come la (8.21): bisogna solo tenere conto che, generalmente, in una rappresentazione ad un solo parametro una delle due curve base risulta esclusa dalla famiglia.

Visti i limiti di questo testo, tratteremo, nel seguito, solo insiemi di rette e di circonferenze, importanti anche dal punto di vista della geometria sintetica.

8.5.5. Fasci di rette

Definizione 8.5 (Fascio di rette)**.** *Si chiama* fascio proprio *di rette l'insieme di tutte le rette di un piano aventi un punto in comune, detto* centro del fascio*; si chiama* fascio improprio *l'insieme di tutte le rette di un piano tra di loro parallele, cioè aventi la stessa direzione.*

Per trattare analiticamente i fasci di rette presenteremo sia la trattazione con un'equazione a due parametri, come la (8.21), che la trattazione con un'equazione ad un solo parametro, come la (8.22), segnalando che quest'ultima è quella normalmente proposta a livello di scuola media superiore.

Equazione a due parametri

Consideriamo due rette di equazione cartesiana

$$\mathrm{r}_1 : a_1 x + b_1 y + c_1 = 0 \quad \text{e} \quad \mathrm{r}_2 : a_2 x + b_2 y + c_2 = 0. \tag{8.23}$$

Un'equazione del tipo

$$\lambda(a_1 x + b_1 y + c_1) + \mu(a_2 x + b_2 y + c_2) = 0, \quad (\lambda, \mu) \in \mathbb{R}^2, \quad (\lambda, \mu) \neq (0,0) \tag{8.24}$$

è un'equazione di grado ≤ 1: dunque rappresenta una retta (se è di primo grado), l'insieme vuoto (se è di grado 0, ma non è un'identità), oppure l'intero piano (se è un'identità). Gli esempi che seguono mostrano ciascuna di queste possibilità.

Esempio 8.7. $\lambda(x + 2y - 1) + \mu(x - y + 2) = 0 \Rightarrow (\lambda + \mu)x + (2\lambda - \mu)y - \lambda + 2\mu = 0$ è sempre di primo grado, qualunque siano λ e μ, non contemporaneamente nulli: i coefficienti di x e y non possono essere contemporaneamente nulli. Dunque si tratta sempre di una retta.

Esempio 8.8. $\lambda(x+y+1)+\mu(x+y+2)=0 \Rightarrow (\lambda+\mu)x+(\lambda+\mu)y+\lambda+2\mu=0$ è di primo grado se $\lambda+\mu\neq 0$, e in questi casi rappresenta una retta; si riduce a $-1=0$, cioè un'equazione di grado 0 priva di soluzioni, se $\lambda+\mu=0$, e in questo caso rappresenta l'insieme vuoto.

Esempio 8.9. $\lambda(x+y+1)+\mu(x+y+1)=0 \Rightarrow (\lambda+\mu)x+(\lambda+\mu)y+\lambda+\mu=0$ è di primo grado se $\lambda+\mu\neq 0$, e in questi casi rappresenta una sola retta, precisamente la $x+y+1=0$; si riduce a $0=0$, cioè un'identità, se $\lambda+\mu=0$, e in questo caso rappresenta l'intero piano.

Si noti come nel primo esempio le due rette base non sono parallele e anzi si incontrano nel punto $\mathsf{P}(-1,1)$: tutte le altre rette della famiglia passeranno ancora per questo punto. Nel secondo esempio le due rette base sono parallele ma distinte: tutte le rette della famiglia sono ancora parallele alle due rette base purché si escludano dalla combinazione le coppie (λ,μ) tali che $\lambda+\mu=0$ ovvero $\mu/\lambda=-1$: chiameremo queste coppie *coppie non ammesse per la combinazione lineare*. Nel terzo esempio le due rette sono la stessa retta e la combinazione lineare fornisce ancora la stessa retta purché, anche qui, si escludano dalla combinazione le coppie (λ,μ) tali che $\lambda+\mu=0$ ovvero $\mu/\lambda=-1$.

Si può provare che le situazioni descritte negli esempi proposti sono di carattere generale e anzi che, date due rette $\mathsf{r}_1: a_1x+b_1y+c_1=0$ e $\mathsf{r}_2: a_2x+b_2y+c_2=0$,

— se esse non sono parallele, la (8.24) fornisce l'equazione dell'intero fascio proprio di centro il punto P comune a r_1 e r_2;
— se esse sono parallele e distinte, la (8.24) fornisce l'equazione dell'intero fascio improprio di rette aventi la direzione comune a r_1 e r_2, purché si escludano le coppie di parametri che annullano i coefficienti di x e y nella combinazione;
— se esse sono la stessa retta, la (8.24) fornisce ancora l'equazione della stessa retta, purché si escludano le coppie di parametri che annullano i coefficienti di x e y nella combinazione.

Se riprendiamo in esame la (8.24) e supponiamo $\lambda\neq 0$, dividendo per λ otteniamo

$$(a_1x+b_1y+c_1)+\frac{\mu}{\lambda}(a_2x+b_2y+c_2)=0,$$

che, posto $k=\mu/\lambda$, si può scrivere nella forma

$$(a_1x+b_1y+c_1)+k(a_2x+b_2y+c_2)=0,\quad k\in\mathbb{R}, \tag{8.25}$$

oppure

$$(a_1+a_2k)x+(b_1+b_2k)y+(c_1+c_2k)=0,\quad k\in\mathbb{R}. \tag{8.26}$$

Come già osservato, la (8.26) differisce dalla (8.24) solo per il fatto che, se $\mathsf{r}_2\neq\mathsf{r}_1$, la r_2 stessa non si ottiene per nessun valore di k, oppure, come si usa dire, corrisponde al valore $k=\infty$. Dunque, con questa limitazione, l'equazione a un parametro (8.26) ha le stesse caratteristiche dell'equazione a due parametri (8.26). Chiameremo *retta esclusa* la retta corrispondente al valore $k=\infty$ nell'equazione (8.26). Vedremo, nel paragrafo successivo, che cosa succede se si parte direttamente da un'equazione del tipo (8.26).

Equazione ad un solo parametro

Consideriamo un'equazione del tipo (8.26), ovvero un'equazione di primo grado nelle incognite x e y, i cui coefficienti siano funzioni lineari di grado ≤ 1 di un parametro reale k:

$$(a_1 + a_2 k)x + (b_1 + b_2 k)y + (c_1 + c_2 k) = 0, \quad k \in \mathbb{R}. \tag{8.27}$$

Riscrivendo questa equazione nella forma

$$(a_1 x + b_1 y + c_1) + k(a_2 x + b_2 y + c_2) = 0, \quad k \in \mathbb{R}, \tag{8.28}$$

non è difficile provare, tenendo anche conto di quanto già detto a proposito dell'equazione a due parametri, che al variare di k in $\mathbb{R}$ gli insiemi descritti dall'equazione (8.27) sono costituiti da un fascio proprio sempre privato di una retta, da un fascio improprio eventualmente privato di una retta, da una singola retta, dall'intero piano, dall'insieme vuoto. Si hanno precisamente le situazioni di seguito descritte.

1. Se $a_1 x + b_1 y + c_1 = 0$ e $a_2 x + b_2 y + c_2 = 0$ sono le equazioni di due rette non parallele, diciamole r_1 e r_2, e P è il loro punto di intersezione, la (8.27) rappresenta il fascio proprio di rette di centro P, privato della retta r_2, che appartiene al fascio, ma non si ottiene dalla (8.27) per nessun valore di k: essa prende il nome di *retta esclusa*. Si parla usualmente ancora di fascio proprio, anche se con un certo abuso di linguaggio. La retta r_1 corrisponde al valore 0 del parametro. In questo caso le rette ottenute dalla retta esclusa per rotazione, in senso orario o antiorario, attorno al centro corrispondono a valori di k crescenti (decrescenti) oppure decrescenti (crescenti). Avvicinandosi alla retta esclusa nel senso dell'aumento, i valori di k "tendono a $+\infty$", nel senso della diminuzione "tendono a $-\infty$".
2. Se $a_1 x + b_1 y + c_1 = 0$ e $a_2 x + b_2 y + c_2 = 0$ sono le equazioni di due rette parallele, diciamole r_1 e r_2, la (8.27) rappresenta il fascio improprio di rette di centro P, privato della retta r_2, che appartiene al fascio, ma non si ottiene dalla (8.27) per nessun valore di k: essa prende il nome di *retta esclusa*. Si parla usualmente ancora di fascio improprio, anche se con un certo abuso di linguaggio. La retta r_1 corrisponde al valore 0 del parametro. Inoltre per il valore di k che, sempre nella(8.27), annulla sia il coefficiente di x che quello di y non si ottiene alcun punto del piano: lo possiamo chiamare valore *non ammesso* di k. In questo caso le rette ottenute dalla retta esclusa per traslazione perpendicolare in uno dei due versi possibili corrispondono a valori di k crescenti (decrescenti) oppure decrescenti (crescenti). Avvicinandosi alla retta esclusa nel senso dell'aumento, i valori di k "tendono a $+\infty$", nel senso della diminuzione "tendono a $-\infty$". Allontanandosi invece dalla retta esclusa i valori di k tendono al valore non ammesso, in un senso per difetto, nell'altro per eccesso.
3. Se una delle due equazioni $a_1 x + b_1 y + c_1 = 0$ e $a_2 x + b_2 y + c_2 = 0$ non è l'equazione di una retta (cioè non è di primo grado), allora si possono avere i casi seguenti.
   a) Se $(a_1 = b_1 = 0) \wedge (c_1 \neq 0)$, la (8.27) rappresenta il fascio improprio di rette parallele a r_2, con esclusione di r_2 stessa, e $k = 0$ è il valore non ammesso (l'equazione non ha nessuna soluzione).
   b) Se $a_1 = b_1 = c_1 = 0$, la (8.27) rappresenta, per $k \neq 0$ solo la retta r_2, mentre per $k = 0$ è un'identità.

c) Se $(a_2 = b_2 = 0) \wedge (c_2 \neq 0)$, la (8.27) rappresenta l'intero fascio improprio di rette parallele a r_1, senza retta esclusa, e con tutti i valori ammessi per il parametro k. Si tratta di una situazione particolarmente importante, in quanto un intero fascio può essere rappresentato da un'equazione con un solo parametro. In generale, data una retta r di equazione $ax+by+c=0$, l'intero fascio di rette ad essa parallele può essere scritto nella forma $ax+by+k=0$, con $k \in \mathbb{R}$.

d) Se $a_2 = b_2 = c_2 = 0$, la (8.27) rappresenta solo la retta r_1, ma questo è un caso banale perché nella (8.27) non compare alcun parametro.

Si vedano gli esempi che seguono.

Esempio 8.10. Rappresentare graficamente le caratteristiche dell'insieme di rette di equazione

$$(1+k)x + (1-2k)y - 3 = 0 \quad \text{ovvero} \quad (x+y-3) + k(x-2y) = 0.$$

Le rette $r_1: x+y-3=0$ e $r_2: x-2y=0$ si intersecano in $P(2,1)$. L'equazione data rappresenta dunque il fascio proprio di centro P, con r_2 come retta esclusa. I valori del parametro aumentano ruotando in senso orario: per provarlo basta tracciare due rette corrispondenti a due diversi valori del parametro e esaminare la situazione restando in uno dei due semipiani individuati dalla retta esclusa; nella figura 8.14 abbiamo scelto $k=0$ (che fornisce la retta base r_1) e $k = 1/2$ (che fornisce la retta verticale per P).

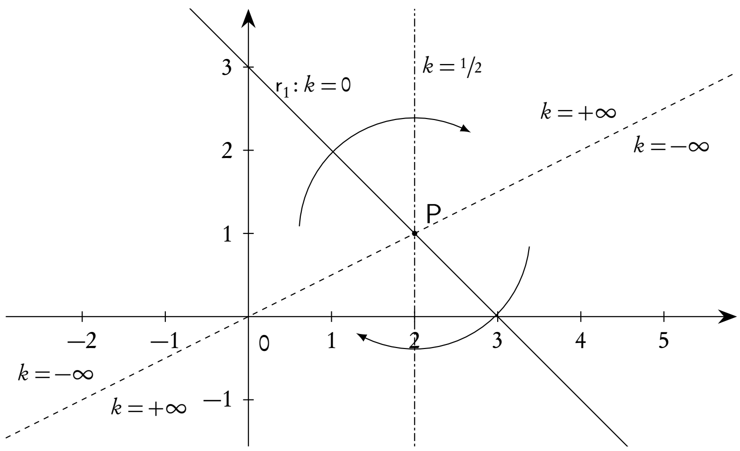


Figura 8.14.: *Il "fascio" proprio di equazione* $(1+k)x + (1-2k)y - 3 = 0$

Esempio 8.11. Rappresentare graficamente le caratteristiche dell'insieme di rette di equazione

$$(x-2y) + k(x-2y+4) = 0 \quad \text{ovvero} \quad (1+k)x - 2(1+k)y + 4k = 0.$$

Le rette $r_1: x-2y=0$ e $r_2: x-2y+4=0$ sono parallele. L'equazione data rappresenta dunque il fascio improprio di direzione coincidente con quella di r_1 ed r_2, con r_2 come retta esclusa. Inoltre il valore $k=-1$ annulla sia il coefficiente di x che quello di y, quindi è il valore non ammesso per il parametro. Il parametro k aumenta come indicato nella figura 8.15: anche qui basta tracciare due rette corrispondenti a due diversi valori del parametro k, rimanendo in uno dei due semipiani individuati dalla retta esclusa.

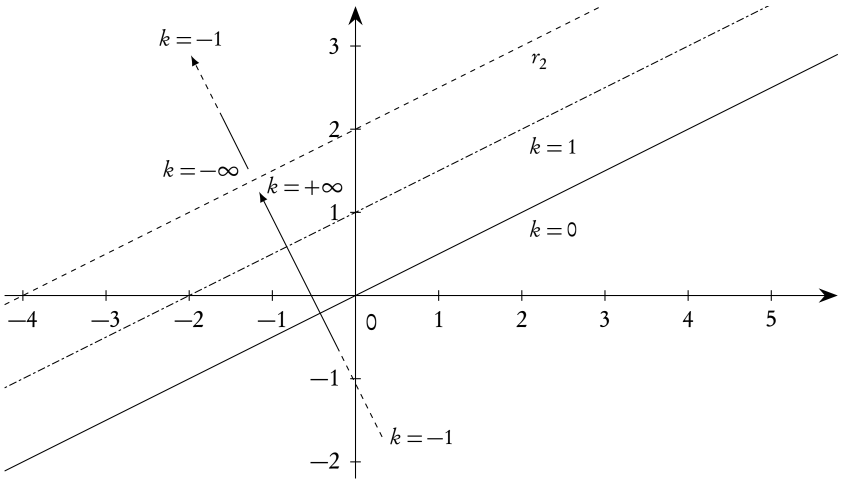


Figura 8.15.: *Il "fascio" improprio* $(x-2y)+k(x-2y+4)=0$

Esempio 8.12. Individuare le caratteristiche dell'insieme di rette di equazione

$$x-2y+1+3k=0 \quad \text{ovvero} \quad (x-2y+1)+k(3)=0.$$

L'equazione $x-2y+1=0$ rappresenta una retta, r_1, mentre l'equazione $3=0$ non ha soluzioni. L'insieme rappresenta dunque l'intero fascio di rette parallele alla retta r_1, senza alcuna retta esclusa e senza alcun valore di k non ammesso.

Un caso particolarmente importante di fascio proprio di rette è quello dell'equazione (8.16), che riportiamo qui di seguito:

$$y-y_0=m(x-x_0). \tag{8.29}$$

Essa, come è noto, serve a determinare l'equazione di una retta non verticale passante per il punto $\mathrm{P}(x_0,y_0)$ e con coefficiente angolare m. Vista come un'equazione di primo grado in due incognite, con parametro m, rappresenta anche il fascio di rette di centro $\mathrm{P}(x_0,y_0)$ e con retta esclusa la retta $x=x_0$, ovvero la retta verticale per P: lo chiameremo il "fascio di rette non verticali per P".

Osservazione 8.6. Si presti particolare attenzione al fatto che un'equazione di primo grado in due incognite con coefficienti dipendenti da un parametro rappresenta un fascio di rette (con le dovute eccezioni) *solo* se la dipendenza dal parametro è lineare per ciascun coefficiente.

Per esempio la famiglia di rette

$$x-2y+k^2=0$$

non rappresenta un fascio di rette. Le rette della famiglia sono tutte parallele alla retta $x-2y=0$, ma hanno ordinata all'origine ≥ 0.

8.6. Le coniche

8.6.1. Generalità. Equazione. Tangenti

Nello studio della geometria euclidea, dopo la retta si incontra solitamente una sola linea curva, la circonferenza, per le sue notevoli proprietà. Le coniche sono, subito dopo la circonferenza, le linee piane più semplici e tra le prime che sono state storicamente studiate: la sistemazione pressoché definitiva della teoria delle coniche come sezioni è opera di Apollonio di Perga, o Perge, in un famoso trattato della fine del terzo secondo a.C. Ne faremo un breve cenno nel paragrafo 8.6.2.

Dal punto di vista della geometria analitica è però più interessante la trattazione delle coniche come particolari luoghi geometrici: la dimostrazione che le sezioni coniche possono essere definite anche come luoghi geometrici è contenuta in un famoso teorema, pubblicato dal matematico belga G.P.Dandelin nel 1822, e detto anche *Teorema del cono gelato*. Tratteremo esplicitamente nel seguito le situazioni più comuni.

Il motivo per cui, in un corso di geometria analitica, le coniche sono trattate subito dopo la retta è legato al fatto che ogni equazione di secondo grado in due incognite rappresenta, nel piano cartesiano, una conica, eventualmente "degenere" e, viceversa, ogni conica è rappresentata nel suo piano in cui si sia introdotto un sistema cartesiano Oxy da un'equazione di secondo grado in due incognite: dal punto di vista analitico il passaggio dalla retta alle coniche equivale al passaggio da equazioni di primo a equazioni di secondo grado in due incognite. In sostanza si dimostra che ogni conica, eventualmente degenere, ha un'equazione del tipo

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0, \text{ con } A, B, C \text{ non contemporaneamente nulli.} \tag{8.30}$$

Viceversa ogni equazione del tipo (8.30) rappresenta una conica, eventualmente degenere.

L'equazione (8.30) contiene 6 parametri; essendo però almeno uno dei primi tre non nullo, l'equazione stessa si può riscrivere in modo da contenere solo 5 parametri: occorreranno dunque 5 condizioni indipendenti (e compatibili) per trovare una conica, per esempio il passaggio per 5 punti. Vedremo nei vari casi altri tipi di situazioni comuni.

Un problema di grande importanza applicativa nella teoria delle coniche è quella della determinazione delle *tangenti* condotte da un punto o aventi direzione assegnata: chiaramente il problema non ha sempre soluzioni, come si evince se si pensa alla circonferenza e si cercano le tangenti condotte da un punto interno alla stessa. Nelle varie situazioni che tratteremo individueremo delle strategie speciali per determinare queste tangenti. Tuttavia è importante considerare fin da subito una tecnica valida per tutti i tipi di coniche e dovuta ad una proprietà delle tangenti ampiamente nota per la circonferenza ed estendibile a tutte le coniche: una retta è tangente a una conica se e solo se ha con la conica stessa due intersezioni coincidenti. Detto in termini analitici una conica e una retta sono tra di loro tangenti se il sistema tra la conica e la retta

$$\begin{cases} Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \\ ax + by + c = 0 \end{cases} \tag{8.31}$$

ha due soluzioni coincidenti. In pratica per risolvere il sistema (8.31) si ricava la x o la y dalla seconda equazione e la si sostituisce nella prima, ottenendo un'equazione in un'incognita di grado ≤ 2, detta *equazione risolvente*: la retta e la conica sono tangenti se e solo se l'equazione risolvente è di secondo grado e ha il discriminante nullo. Si presti particolare attenzione al fatto che l'equazione risolvente deve essere di secondo grado (oltre ad avere discriminante nullo). Se per esempio consideriamo la conica $x^2+3x-2y-2=0$ (che, come vedremo, è una parabola) e la retta $x-1=0$, il sistema (8.31) diventa

$$\begin{cases} x^2+3x-2y-2=0 \\ x-1=0 \end{cases}$$

ed ha la sola soluzione $(1,1)$, ma equazione risolvente $y-1=0$, che non è di secondo grado: la retta e la conica non sono tangenti, pur avendo un solo punto in comune.

In generale dunque per trovare le tangenti ad una conica da un punto $\mathsf{P}(x_0,y_0)$, si considera il fascio proprio di rette di centro P e si impone la condizione che il discriminante dell'equazione risolvente del sistema tra il fascio e la conica sia nullo. Per trovare le tangenti di direzione assegnata basterà considerare un fascio improprio di rette anziché un fascio proprio. Consideriamo un semplice esempio.

Esempio 8.13. Trovare le eventuali tangenti alla conica di equazione $x^2+2y^2+2x=0$ condotte dal punto $\mathsf{P}(0,1)$.

Consideriamo il fascio di rette per P nella forma $y-1=mx$, da cui resta esclusa la retta $x=0$, che considereremo a parte. Il sistema tra la conica e il fascio diventa

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

con equazione risolvente

$$(1+2m^2)x^2+2(2m+1)x+2=0.$$

L'equazione risolvente, di secondo grado, ha $\Delta=0$ se $m={}^1/_4$. Dunque la retta

$$y=\frac{1}{4}x+1$$

è tangente alla conica. Esaminiamo ora la retta $x=0$. Il sistema con la conica è immediato e ha equazione risolvente $y^2=0$ che è di secondo grado e ha discriminante nullo. Dunque anche la retta $x=0$ è tangente alla conica.

Si noti che, per risolvere questo problema, non è stato necessario scoprire il tipo di conica in esame: la strategia indicata non dipende dal tipo di conica. Solo per completezza segnaliamo che, come vedremo, la conica in questione è un'ellisse che abbiamo rappresentato, assieme alle tangenti, nella figura 8.16.

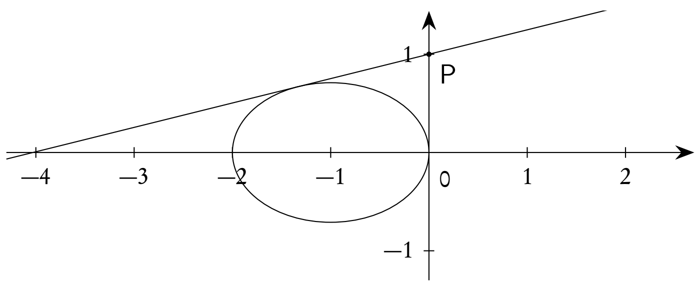


Figura 8.16.: *L'ellisse $x^2+2y^2+2x=0$ e le tangenti condotte dal punto* $\mathsf{P}(0,1)$

Se poi si deve determinare l'unica tangente alla conica in un suo punto $\mathsf{P}(x_0, y_0)$ si possono usare le cosiddette *formule di sdoppiamento*. Precisamente data una conica di equazione $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ e un suo punto $\mathsf{P}(x_0, y_0)$, l'equazione della tangente in P_0 si ottiene direttamente dall'equazione della conica con le seguenti sostituzioni.

$$
\begin{array}{ccccccc}
x^2 & \to & x_0 x & ; & y^2 & \to & y_0 y \\
x & \to & \dfrac{x + x_0}{2} & ; & y & \to & \dfrac{y + y_0}{2} \\
 & & xy & \to & \dfrac{xy_0 + x_0 y}{2} & &
\end{array}
\tag{8.32}
$$

Esempio 8.14. Data la conica

$$4x^2 + y^2 + 5xy + x - y = 0,$$

l'equazione della tangente nel suo punto $\mathsf{P}_0(0, 1)$ è

$$y + 5\frac{x}{2} + \frac{x}{2} - \frac{y+1}{2} = 0 \quad \Rightarrow \quad 6x + y - 1 = 0.$$

Anche in questo caso non è necessario conoscere il tipo di conica per scrivere l'equazione della tangente. Sempre per completezza riportiamo, nella figura 8.17, il grafico della curva (che risulta essere un'iperbole) e della sua tangente nel punto P_0.

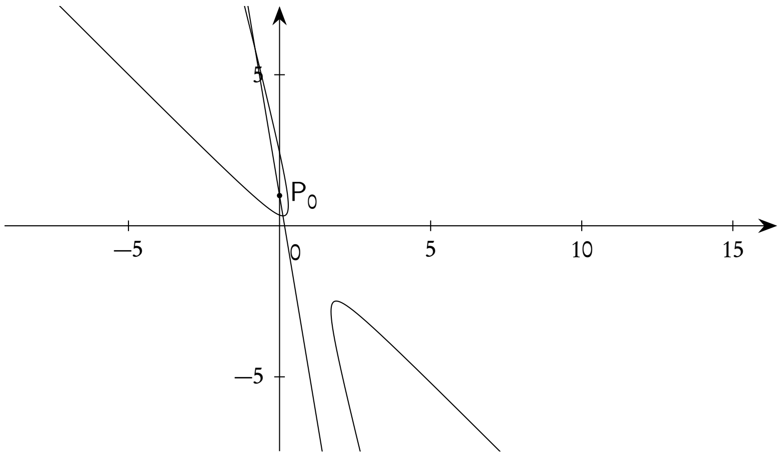


Figura 8.17.: *L'iperbole* $4x^2 + y^2 + 5xy + x - y = 0$ *e la sua tangente nel punto* $\mathsf{P}_0(0, 1)$

8.6.2. Le coniche come sezioni

In un piano π consideriamo due rette distinte a e g, non perpendicolari, e immaginiamo di far compiere, nello spazio, una rotazione completa di g attorno ad a. Se a e g si intersecano in un punto V, individuando un angolo acuto α $(0 < \alpha < \pi/2)$, otterremo una superficie detta *cono circolare a due falde indefinito* di *semiapertura* α, se a e g sono parallele otterremo un *cilindro circolare indefinito*, che si può anche chiamare un *cono degenere*. In entrambi i casi la retta a si chiama *asse* (del cono o del cilindro),

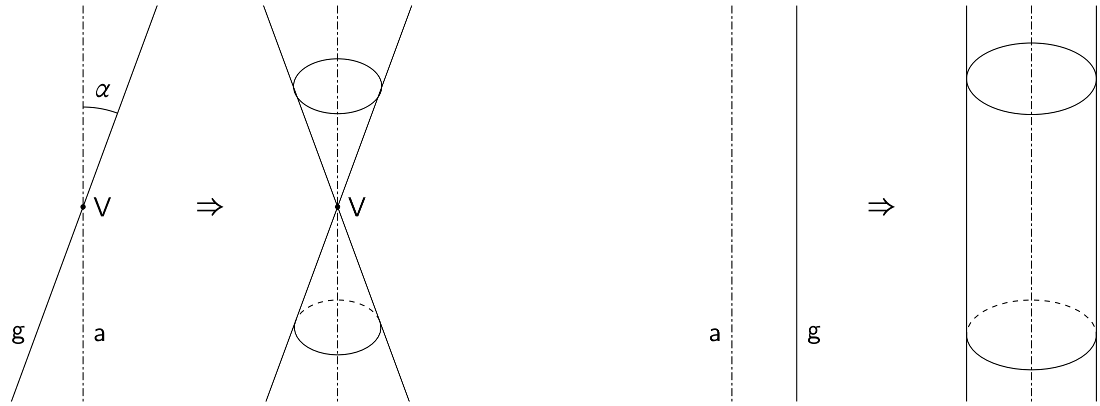


Figura 8.18.: *Cono e cilindro circolari indefiniti*

mentre la retta g e quelle che si ottengono dalla rotazione di g si chiamano *generatrici*. Il punto V si chiama *vertice* del cono. Si veda la figura 8.18.

Costruiti in questo modo il cono a due falde o il cilindro indefiniti, consideriamo un piano δ nello spazio. Se δ non passa per il vertice del cono o non è parallelo all'asse del cilindro, intersecherà il cono stesso o il cilindro lungo una curva che prende il nome di *conica non degenere*. Nel caso del cilindro si otterrà sempre una *ellisse* (e come caso particolare una circonferenza se δ risulta perpendicolare all'asse a): si veda la figura 8.25. Nel caso del cono, detto β l'angolo non ottuso tra δ e a, $(0 \le \beta \le \pi/2)$ si ottiene:

- un'*ellisse* se $\beta > \alpha$ (una circonferenza se $\beta = \pi/2$): si veda la figura 8.19;
- una *parabola* se $\beta = \alpha$: si veda la figura 8.21;
- un'*iperbole* se $\beta < \alpha$: si veda la figura 8.20. In quest'ultimo caso il piano δ interseca entrambe le falde del cono, per cui la curva avrà due rami.

Se invece il piano δ passa per il vertice del cono o è parallelo all'asse del cilindro le curve sezione si chiamano *coniche degeneri*. Nel caso del cilindro si ottiene:

- l'insieme vuoto se il piano è esterno al cilindro;
- una coppia di rette coincidenti se il piano è tangente al cilindro: si veda la figura 8.27;
- una coppia di rette parallele nel restante caso: si veda la figura 8.26.

Nel caso del cono si ottiene:

- un punto se $\beta > \alpha$: si veda la figura 8.24;
- una coppia di rette coincidenti se $\beta = \alpha$ (piano tangente al cono): si veda la figura 8.23;
- una coppia di rette incidenti se $\beta < \alpha$: si veda la figura 8.22.

Le figure dalla 8.19 alla 8.27 illustrano le situazioni considerate. In ognuna delle figure, tranne la 8.27, sono tracciati il cono o il cilindro con il piano sezionante, la parte di cono o cilindro "soprastante" il piano sezionante, la parte di cono o cilindro "sottostante" il piano sezionante. È inoltre ovviamente tracciata la curva sezione.

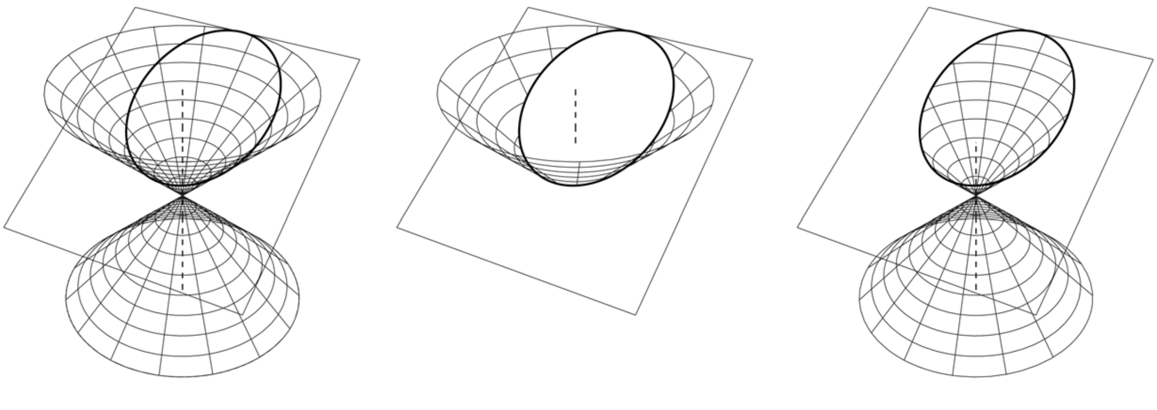


Figura 8.19.: *Ellisse come sezione di un piano con un cono a due falde*

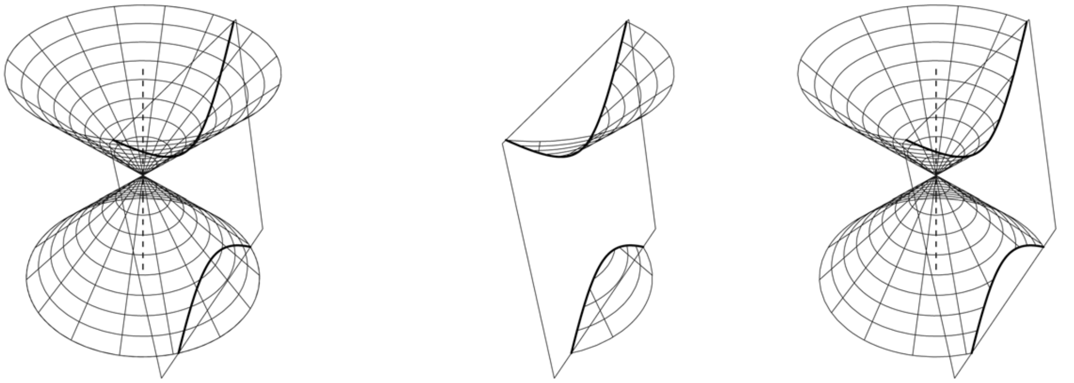


Figura 8.20.: *Iperbole come sezione di un piano con un cono a due falde*

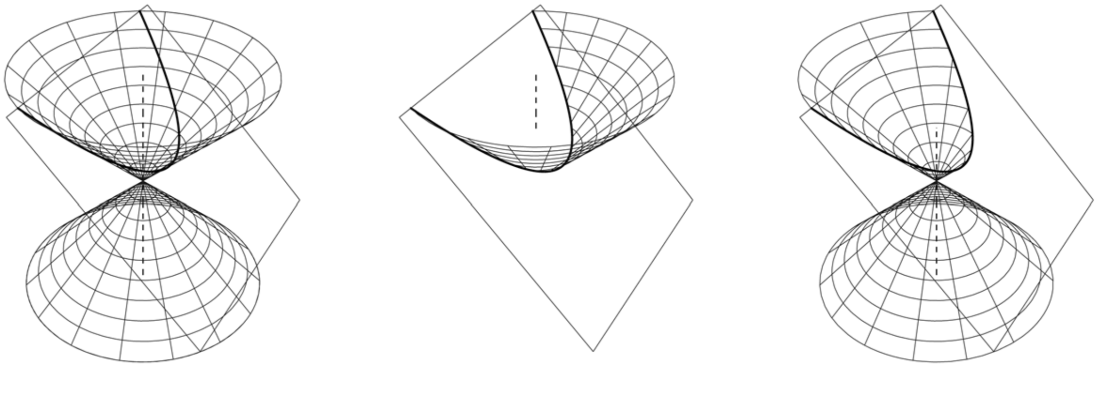


Figura 8.21.: *Parabola come sezione di un piano con un cono a due falde*

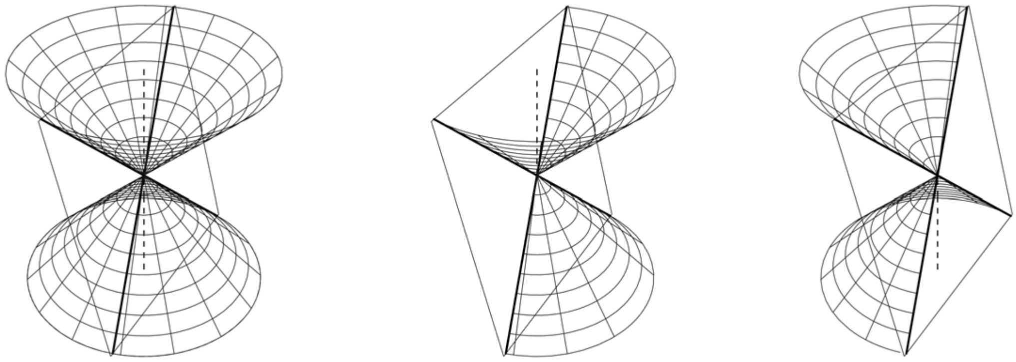


Figura 8.22.: *Coppia di rette incidenti come sezione di un piano con un cono a due falde*

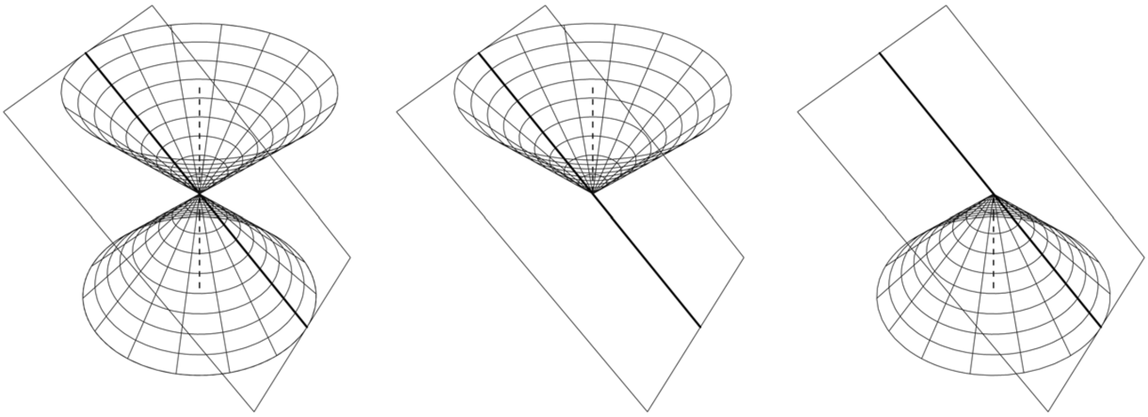


Figura 8.23.: *Una retta "doppia" come sezione di un piano tangente a un cono a due falde*

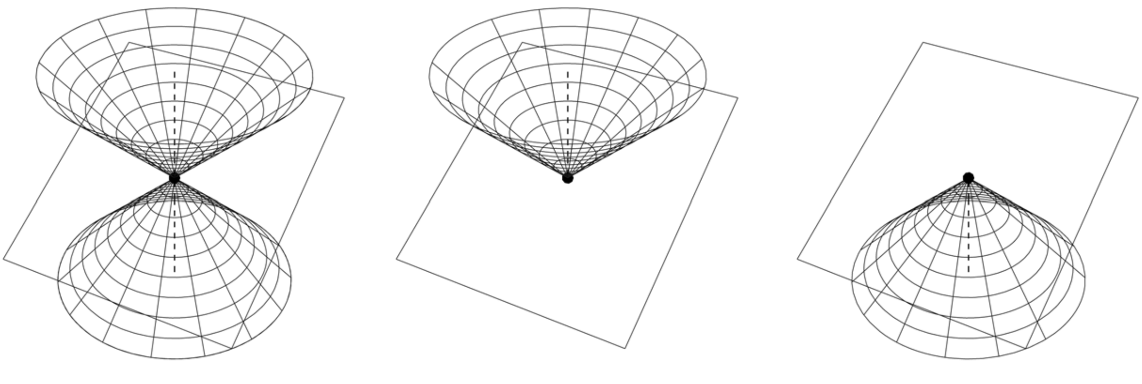


Figura 8.24.: *Un punto come sezione di un piano con un cono a due falde*

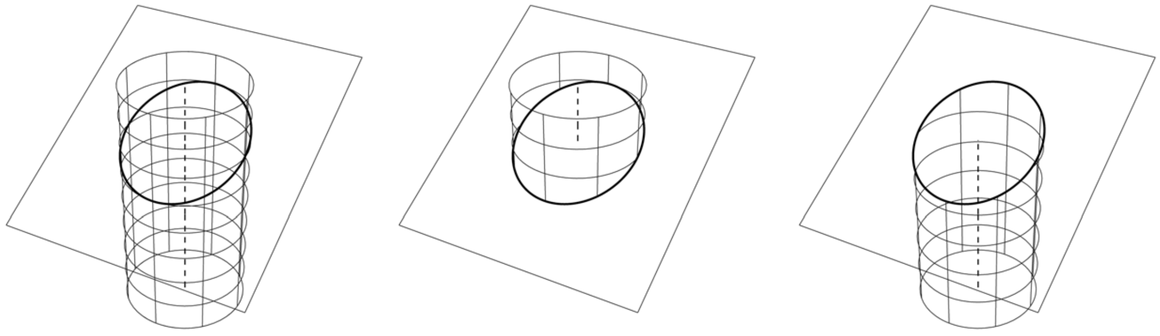


Figura 8.25.: *Ellisse come sezione di un piano con un cilindro*

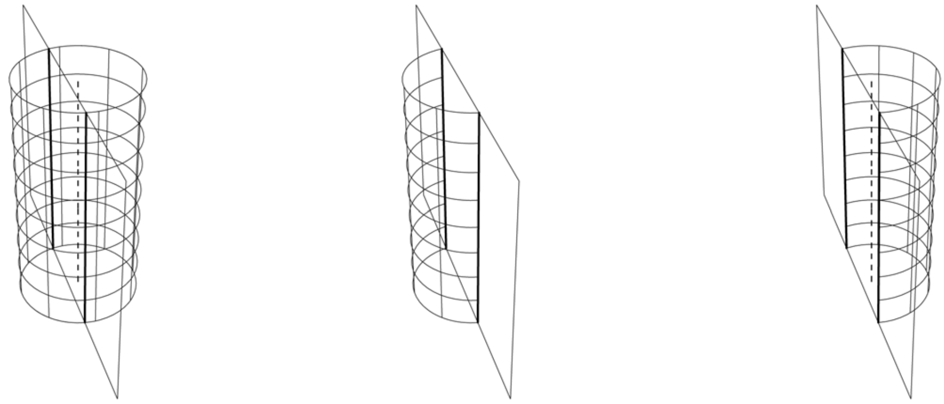


Figura 8.26.: *Coppia di rette parallele come sezione di un piano con un cilindro*

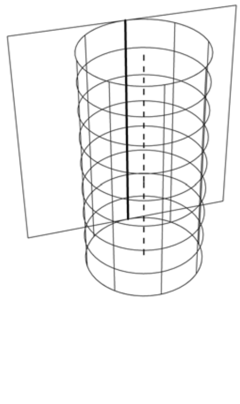


Figura 8.27.: *Una retta "doppia" come sezione di un piano tangente a un cilindro*

8.6.3. Le coniche come luoghi geometrici

Come già accennato, sulla base del teorema di Dandelin, le coniche possono essere definite anche come opportuni luoghi geometrici del piano e ci sono diverse possibilità. Vista la natura di questo testo, sceglieremo la via più semplice e considereremo solo le situazioni più comuni, in particolare facendo riferimento solo alle coniche non degeneri. Valgono i seguenti teoremi.

Teorema 8.7. *Dati, in un piano π, due punti F_1 e F_2 detti* fuochi, *l'*ellisse *è il luogo geometrico dei punti del piano la somma delle cui distanze dai fuochi è costante (e maggiore della distanza tra i fuochi stessi). Se i fuochi coincidono l'ellisse si riduce a una circonferenza.*

Teorema 8.8. *Dati, in un piano π, due punti F_1 e F_2 detti* fuochi, *l'*iperbole *è il luogo geometrico dei punti del piano per i quali è costante, e non nullo, il modulo della differenza delle distanze dai fuochi.*

Teorema 8.9. *Dati, in un piano π, un punto F detto* fuoco *e una retta d detta* direttrice, *con $\mathsf{F} \notin \mathsf{d}$, la* parabola *è il luogo geometrico dei punti del piano equidistanti da F e da d.*

Come mostrano i teoremi appena enunciati, la proprietà che caratterizza la parabola è diversa da quelle, sostanzialmente simili, dell'ellisse e dell'iperbole. Tuttavia sarebbe possibile dare una definizione come luogo con una condizione dello stesso tipo per tutte e tre le coniche, ma questo esula dagli scopi di questo testo.

I tre teoremi 8.7, 8.8 e 8.9 consentono, come vedremo, di determinare le equazioni cartesiane di queste tre curve, anche se noi ci limiteremo a particolari scelte degli assi coordinati.

8.7. La circonferenza nel piano cartesiano

Sia sulla base del teorema 8.7 che della ben nota definizione della geometria euclidea, la circonferenza è il luogo geometrico dei punti del piano aventi da un punto fisso detto *centro* una distanza costante r detta raggio. Scelto allora nel piano un sistema cartesiano ortogonale $\mathsf{O}xy$, indichiamo con $(x_\mathsf{C}, y_\mathsf{C})$ le coordinate del centro. Detto $\mathsf{P}(x,y)$ un generico punto del luogo si deve avere

$$\sqrt{(x-x_\mathsf{C})^2+(y-y_\mathsf{C})^2}=r, \tag{8.33}$$

ovvero, elevando al quadrato ambo i membri senza problemi perché entrambi positivi e perché non ci sono condizioni per il dominio del radicale,

$$(x-x_\mathsf{C})^2+(y-y_\mathsf{C})^2=r^2. \tag{8.34}$$

Eseguendo i calcoli e semplificando si ottiene

$$x^2+y^2-2x_\mathsf{C}x-2y_\mathsf{C}y+x_\mathsf{C}^2+y_\mathsf{C}^2-r^2=0. \tag{8.35}$$

Posto

$$a=-2x_\mathsf{C}x, \quad b=-2y_\mathsf{C}, \quad c=x_\mathsf{C}^2+y_\mathsf{C}^2-r^2, \tag{8.36}$$

l'equazione 8.35 assume la forma

$$x^2+y^2+ax+by+c=0, \tag{8.37}$$

detta *equazione canonica della circonferenza*.

Viceversa una equazione del tipo (8.37) rappresenta una circonferenza se, e solo se,

$$\frac{a^2}{4}+\frac{b^2}{4}-c\geq 0. \tag{8.38}$$

Se la 8.38 è verificata, il centro della circonferenza e il raggio sono dati dalle formule

$$x_C=-\frac{a}{2},\quad y_C=-\frac{b}{2},\quad r=\sqrt{\frac{a^2}{4}+\frac{b^2}{4}-c}. \tag{8.39}$$

Si noti che la (8.37) rientra nella forma generale dell'equazione di una conica

$$Ax^2+Bxy+Cy^2+Dx+Ey+F=0, \tag{8.40}$$

con le condizioni

$$A=C\neq 0,\quad B=0. \tag{8.41}$$

Dunque la più generale equazione di una circonferenza nel piano è del tipo

$$Ax^2+Ay^2+Dx+Ey+F=0,\quad A\neq 0, \tag{8.42}$$

equazione che può essere scritta in forma canonica dividendo ambo i membri per A. Dopo aver ottenuto la forma canonica[5] si può controllare la validità della condizione (8.38) e applicare le (8.39) per trovare il centro e il raggio.

Usando la tecnica del completamento dei quadrati di cui abbiamo già parlato nella pagina 89, si possono determinare eventuale centro e raggio della circonferenza scritta in forma canonica, senza far ricorso ad alcuna formula. Si procede come di seguito indicato.

$$\begin{aligned} x^2+y^2+ax+by+c &= (x^2+ax)+(y^2+by)+c= \\ &= \left(x^2+ax+\frac{a^2}{4}\right)+\left(y^2+by+\frac{b^2}{4}\right)+c-\frac{a^2}{4}-\frac{b^2}{4}= \\ &= \left(x+\frac{a}{2}\right)^2+\left(y+\frac{b}{2}\right)+c-\frac{a^2}{4}-\frac{b^2}{4}. \end{aligned}$$

L'equazione canonica può allora essere scritta nella forma

$$\left(x+\frac{a}{2}\right)^2+\left(y+\frac{b}{2}\right)=\frac{a^2}{4}+\frac{b^2}{4}-c, \tag{8.43}$$

che, per confronto con la (8.34), fornisce subito la condizione (8.38), le eventuali coordinate del centro e il raggio.

[5] Si presti attenzione a questo fatto: le (8.38) e (8.39) valgono solo per la forma canonica dell'equazione.

Esempio 8.15. Trovare eventuale centro e raggio della circonferenza di equazione

$$4x^2+4y^2-4x-16y+13=0, \text{ ovvero, in forma canonica, } x^2+y^2-x-4y+\frac{13}{4}=0$$

Applicando la tecnica del completamento dei quadrati si ottiene:

$$\begin{aligned} x^2+y^2-x-4y+\frac{13}{4} &= (x^2-x)+(y^2-4y)+\frac{13}{4}= \\ &= \left(x^2-x+\frac{1}{4}\right)+(y^2-4y+4)-\frac{1}{4}-4+\frac{13}{4}= \\ &= \left(x-\frac{1}{2}\right)^{+}(y-2)^2-1. \end{aligned}$$

L'equazione data si può scrivere nella forma

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

da cui si deduce che si tratta in effetti di una circonferenza di centro $C(1/2,2)$ e raggio $r=1$.

8.7.1. Tangenti a una circonferenza

Per trovare le tangenti a una circonferenza si può procedere con i metodi generali indicati per le coniche, ma, viste le speciali proprietà di questa curva, si possono adottare anche strategie molto più efficienti.

Se si deve trovare la tangente ad una circonferenza in un *suo punto* $P(x_0,y_0)$ si può trovare la perpendicolare al raggio, passante per P_0, oppure trovare la retta avente distanza dal centro uguale al raggio. Il secondo metodo è adatto anche a trovare le due tangenti condotte da un punto esterno. Non è escluso che si possano applicare anche altre tecniche, basate sulle proprietà geometriche della circonferenza e delle sue tangenti.

Esempio 8.16. Trovare, usando le diverse strategie proposte, le eventuali tangenti alla circonferenza di equazione $x^2+y^2+2x-2y-3=0$ condotte dal punto $P(1,2)$.

La circonferenza si può riscrivere (completamento dei quadrati!) come $(x+1)^2+(y-1)^2=5$, dunque ha centro $C(-1,1)$ e raggio $r=\sqrt{5}$. È immediato che il punto P appartiene alla circonferenza. Vediamo di applicare i vari metodi proposti e adatti a questo caso.

— Considerazioni elementari indicano che una retta verticale per P non può essere tangente alla circonferenza. Consideriamo perciò il fascio di rette non verticali per P, fascio che ha equazione $y-2=m(x-1)$. Il sistema con la circonferenza è

$$\begin{cases} x^2+y^2+2x-2y-3=0 \\ y-2=m(x-1) \end{cases},$$

che ha come equazione risolvente

$$(1+m^2)x^2+2(1+m-m^2)x+m^2-2m-3=0.$$

Uguagliando a zero il discriminante si trova la sola[6] soluzione $m = -2$. l'equazione della tangente è allora $y = -2x + 4$.

— Usando le formule di sdoppiamento si devono fare le seguenti sostituzioni nell'equazione della circonferenza:

$$x^2 \to x; \quad y^2 \to 2y; \quad x \to \frac{x+1}{2}; \quad y \to y \to \frac{y+2}{2}.$$

Si ottiene

$$x + 2y + 2\frac{x+1}{2} - 2\frac{y+2}{2} - 3 = 0 \quad \Rightarrow \quad 2x + y - 4 = 0,$$

in perfetto accordo con il risultato precedente.

— La retta CP ha coefficiente angolare

$$m = \frac{\Delta y}{\Delta x} = \frac{2-1}{1-(-1)} = \frac{1}{2}.$$

La tangente, che è perpendicolare a CP, avrà dunque coefficiente angolare $m = -2$ e equazione

$$y - 2 = -2(x-1) \quad \Rightarrow \quad y = -2x + 4,$$

sempre in accordo con il risultato già trovato.

— Riconsideriamo il fascio di rette non verticali per P, scrivendolo in forma implicita:

$$mx - y - m + 2 = 0,$$

e cerchiamo la retta del fascio che ha distanza $\sqrt{5}$ dal centro. Si deve avere

$$\frac{|-m-1-m+2|}{\sqrt{m^2+1}} = \sqrt{5}.$$

Si può elevare tranquillamente al quadrato (ambo i membri sono positivi!). Semplificando si ottiene

$$m^2 + 4m + 4 = 0 \quad \Rightarrow \quad m = -2,$$

ancora in accordo con il risultato già trovato.

Esempio 8.17. Trovare, usando le diverse strategie proposte, le eventuali tangenti condotte alla circonferenza di equazione $x^2 + y^2 + x + y - 2 = 0$ dal punto $P(3/4, 3/4)$.

L'equazione della circonferenza si può riscrivere nella forma

$$\left(x + \frac{1}{2}\right)^2 + \left(y + \frac{1}{2}\right)^2 = \frac{5}{2},$$

da cui si trovano il centro $C(-1/2, -1/2)$ e il raggio $r = \sqrt{5/2}$.

Poiché

$$d(P, C) = \sqrt{\left(-\frac{1}{2} - \frac{3}{4}\right)^2 + \left(-\frac{1}{2} - \frac{3}{4}\right)^2} = \frac{5}{4}\sqrt{2} > \sqrt{\frac{5}{2}},$$

il punto P è esterno alla circonferenza: ci saranno due tangenti condotte da P. Vediamo di applicare i metodi proposti e adatti a questo caso.

[6] Esiste una sola tangente condotta a una circonferenza da un suo punto!

— Considerazioni elementari indicano che una retta verticale per P non può essere tangente alla circonferenza. Consideriamo perciò il fascio di rette non verticali per P e il sistema tra questo fascio e la circonferenza data,

$$\begin{cases} x^2+y^2+x+y-2=0 \\ y-\dfrac{3}{4}=m\left(x-\dfrac{3}{4}\right) \end{cases},$$

di equazione risolvente

$$(16+16m^2)x^2-(24m^2+40m+16)x+9m^2-30m-11=0.$$

Uguagliando a zero il discriminante di quest'equazione si ottiene

$$3m^2+10m+3=0,$$

che ha le due soluzioni $m_1=-3$ e $m_2=-1/3$. Le due rette tangenti sono dunque, dopo semplificazione

$$3x+y-3=0 \quad \text{e} \quad x+3y-3=0.$$

— Riconsideriamo il fascio di rette non verticali per P, scrivendolo in forma implicita:

$$4mx-4y-3m+3=0,$$

e cerchiamo le rette del fascio che hanno distanza $\sqrt{5/2}$ dal centro. Si deve avere

$$\frac{|-2m+2-3m+3|}{\sqrt{16m^2+16}}=\sqrt{\frac{5}{2}}.$$

Quadrando e semplificando si ottiene nuovamente l'equazione

$$3m^2+10m+3=0$$

in accordo con quanto già trovato.

— Si sarebbe anche potuto procedere con una soluzione "più geometrica": i punti di tangenza delle due rette tangenti condotte da P devono stare sull'intersezione tra la circonferenza data e la circonferenza γ di centro il punto medio M di $\overline{\mathrm{CP}}$ e passante per C (o per P). Per M si ha

$$x_\mathrm{M}=\frac{1}{8}, \quad y_\mathrm{M}=\frac{1}{8}.$$

Il raggio di γ sarà

$$\frac{|\overline{\mathrm{PC}}|}{2}=\frac{5}{8}\sqrt{2}.$$

L'equazione di γ è dunque

$$\left(x-\frac{1}{8}\right)^2+\left(y-\frac{1}{8}\right)^2=\frac{25}{32} \quad \Rightarrow \quad x^2+y^2-\frac{1}{4}x-\frac{1}{4}y-\frac{3}{4}=0.$$

Le intersezioni tra le due circonferenze si trovano mettendo a sistema le due equazioni:

$$\begin{cases} x^2+y^2+x+y-2=0 \\ x^2+y^2-\dfrac{1}{4}x-\dfrac{1}{4}y-\dfrac{3}{4}=0 \end{cases}.$$

Si trovano facilmente i punti $\mathsf{A}(1,0)$ e $\mathsf{B}(0,1)$. Le rette PA e PB sono le due tangenti cercate e si ritrovano le stesse equazioni già note. La figura 8.28 illustra la situazione.

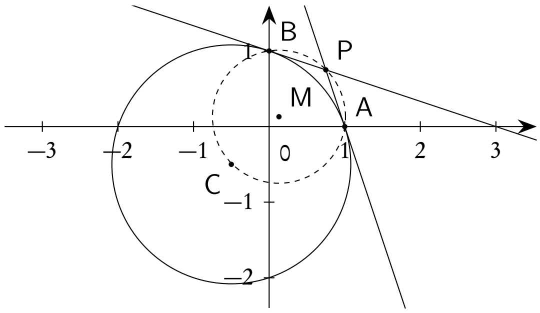


Figura 8.28.: *La circonferenza* $x^2+y^2+x+y-2=0$ *e le tangenti condotte da* $\mathsf{P}(3/4, 3/4)$

8.7.2. Come determinare l'equazione di una circonferenza

L'equazione di una circonferenza, in forma canonica, dipende da tre parametri, dunque per determinarla occorreranno tre condizioni indipendenti, e compatibili.

La condizione più comune è quella del passaggio per un punto: per sfruttare questa condizione la strategia standard(7) è quella di imporre che le coordinate del punto siano soluzioni dell'equazione, ottenendo così una equazione, di primo grado, nei tre parametri che determinano l'equazione della circonferenza. Se è richiesto il passaggio per tre punti distinti si otterrà un sistema di tre equazioni nei tre parametri: se il sistema è risolubile (cosa che succede se i tre punti non sono allineati) si otterrà l'equazione dell'unica circonferenza che passa per i tre punti.

Esempio 8.18. Trovare l'equazione della circonferenza passante per $\mathsf{A}(0,0)$, $\mathsf{B}(1,1)$, $\mathsf{C}(4,0)$.

Applicando la condizione di passaggio per i tre punti alla generica equazione di una circonferenza

$$x^2+y^2+ax+by+c=0,$$

si ottiene il sistema

$$\begin{cases} c=0 \\ 1+1+a+b+c=0 \\ 16+0+4a+0+c=0 \end{cases},$$

di immediata soluzione: $(a,b,c)=(-4,2,0)$. L'equazione richiesta è dunque

$$x^2+y^2-4x+2y=0.$$

[7] Si tratta della strategia comune a tutte le equazioni in due incognite: una curva di equazione $f(x,y)=0$ passa per il punto $\mathsf{P}(x_0,y_0)$ se e solo se le coordinate di P verificano l'equazione, ovvero se e solo se sostituendo (x_0,y_0) ad (x,y) in $f(x,y)=0$ si ottiene un'identità.

La seconda condizione standard è la condizione di tangenza ad una retta: si scrive il sistema costituito dall'equazione della retta e dalla generica equazione di una circonferenza, si ricava poi l'equazione risolvente di questo sistema e infine si impone la condizione che questa equazione risolvente (di secondo grado) abbia discriminante nullo. In questo modo si perviene ad un'equazione, di secondo grado, nei parametri a, b, c.

Esempio 8.19. Si scriva la condizione sui coefficienti a, b, c affinché la circonferenza $x^2+y^2+ax+by+c=0$ sia tangente alla retta $x-2y-1=0$.

Si scrive il sistema

$$\begin{cases} x^2+y^2+ax+by+c=0 \\ x-2y-1=0 \end{cases},$$

la cui equazione risolvente è

$$5y^2+(4+2a+b)y+1+a+c=0.$$

Annullando il discriminante si ottiene la condizione richiesta

$$4a^2+b^2+4ab-4a+8b-20c-4=0.$$

Tuttavia per la circonferenza non sempre le strategie standard appena descritte sono le più efficienti. Per esempio la richiesta di determinare l'equazione delle circonferenze tangenti a tre rette porta a scrivere un sistema di tre equazioni di secondo grado, ovvero un sistema di ottavo grado, anche se è noto a priori che non ci possono essere più di quattro soluzioni. Conviene sempre, se possibile, sfruttare le proprietà geometriche della circonferenza. Gli esempi che seguono mostreranno alcune delle strategie utilizzabili in casi elementari. L'idea di base è quella di determinare centro e raggio della circonferenza cercata, in modo da poter usare poi direttamente la formula (8.34).

Esempio 8.20. Riprendiamo in esame il problema, già trattato nell'esempio 8.18, di trovare l'equazione della circonferenza passante per $A(0,0)$, $B(1,1)$, $C(4,0)$.

In alternativa al procedimento prima seguito si può osservare che $\overline{AB}$, $\overline{BC}$ e $\overline{AC}$ sono corde della circonferenza cercata: il centro O deve trovarsi sull'intersezione degli assi dei tre segmenti $\overline{AB}$, $\overline{BC}$ e $\overline{AC}$, mentre il raggio deve essere uguale alla distanza $|\overline{OA}|$, oppure $|\overline{OB}|$, oppure $|\overline{OC}|$. Basterà trovare O come intersezione di due dei tre assi citati. Per l'asse di $\overline{AC}$ si ha, banalmente, l'equazione $x=2$. L'asse di $\overline{AB}$ deve avere coefficiente angolare -1 (in quanto $\overline{AB}$ ha coefficiente angolare 1) e deve passare per $M(1/2, 1/2)$: si ottengono facilmente l'equazione $x+y-1=0$ e, di conseguenza, le coordinate del centro O, $O(2,-1)$. Si ha poi $|\overline{OA}|=\sqrt{5}$ da cui l'equazione

$$(x-2)^2+(y+1)^2=5.$$

Dal punto di vista tecnico non ci sono grosse differenze rispetto alla strategia seguita precedentemente, ma questo procedimento risulta sicuramente più elegante. La figura 8.29 evidenzia la strategia seguita.

Esempio 8.21. Riprendiamo in esame il problema, cui abbiamo già fatto cenno, di trovare le circonferenze tangenti a tre rette date. Anziché seguire la strategia del "delta$=0$", che porta a un sistema di tre equazioni in tre incognite, di ottavo grado, possiamo osservare che, se le tre rette individuano un triangolo, il problema può essere ricondotto a quello di trovare le circonferenze inscritta ed ex-inscritte nel triangolo, problema che abbiamo già trattato nell'esempio 8.6 a pagina 219, dove abbiamo propio trovato l'incentro, i tre ex-centri e i quattro raggi. Se solo due delle tre rette sono parallele la tecnica può essere facilmente

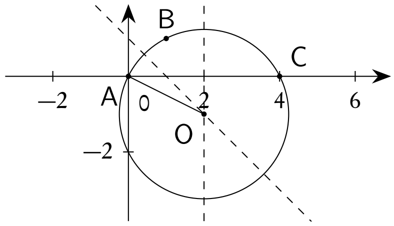


Figura 8.29.: *Circonferenza per tre punti*

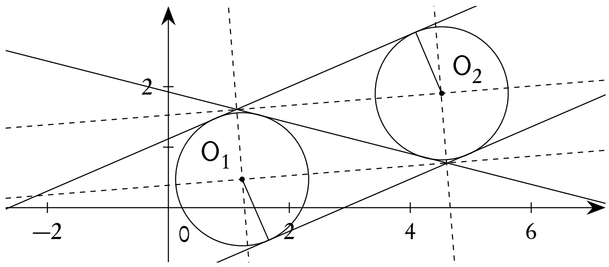


Figura 8.30.: *Circonferenze tangenti a tre rette*

adattata: si veda la figura 8.30, nella quale le rette tratteggiate sono le bisettrici degli angoli individuati dalle tre rette date.

È poi chiaro che, se le tre rette date sono distinte e appartengono a un fascio, proprio o improprio, il problema non ha soluzioni.

Esempio 8.22. Trovare le circonferenze tangenti a una retta r e ad una retta s in un suo punto P, nell'ipotesi che le due rette siano incidenti in un punto A, diverso da P.

Le circonferenze cercate devono avere il centro sulla perpendicolare per P ad s e sulle bisettrici degli angoli individuati dalle due rette r ed s, quindi sulle loro intersezioni. Per trovare i raggi basta trovare poi le distanze dei due centri dal punto P. Si veda la figura 8.31. Se le due rette date fossero parallele, basterebbe trovare la bisettrice della striscia da loro individuata, anziché le bisettrici degli angoli da loro individuati.

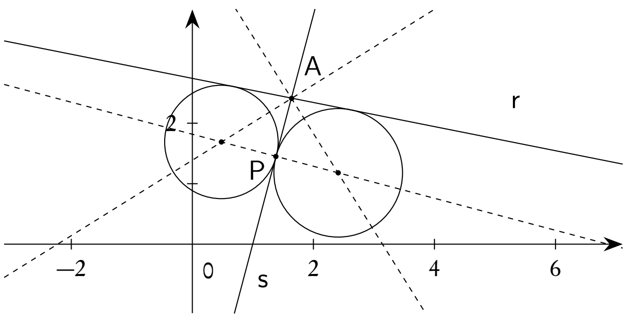


Figura 8.31.: *Circonferenze tangenti a due rette*

Esempio 8.23. Trovare la circonferenza passante per un punto A e tangente ad una retta r in un suo punto P.

La circonferenza cercata deve avere il centro sull'intersezione tra la perpendicolare in P ad r e l'asse del segmento $\overline{AP}$; il raggio è poi la distanza di questo centro da P, o da A. Si veda la figura 8.32.

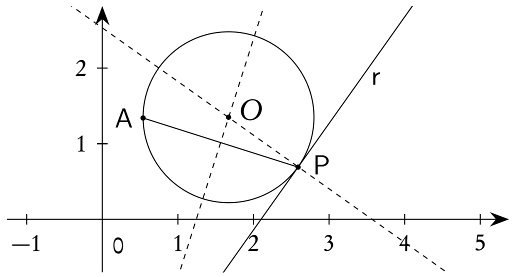


Figura 8.32.: *Circonferenza per un punto e tangente a una retta*

Esempio 8.24. Trovare le circonferenze passanti per due punti A e B e tangenti ad una retta r, nell'ipotesi che AB non sia parallela ad r, che i punti A e B non appartengano ad r e stiano entrambi nello stesso semipiano individuato da r.

Questo problema potrebbe essere agevolmente risolto con le tecniche standard: il passaggio per A e B fornisce due condizioni di primo grado, la tangenza ad r una condizione di secondo grado, per cui si ottiene un sistema di secondo grado, di non difficile risoluzione.

Tuttavia vogliamo proporre un soluzione più "geometrica", con lo scopo di ripassare una proprietà importante delle tangenti e secanti ad una circonferenza. Se le circonferenze devono passare per A e B, significa che la retta AB è una loro secante; indichiamo con S il punto di intersezione tra AB ed r e con T l'incognito punto di tangenza tra r e le circonferenze cercate. Il teorema della secante e della tangente implica che $\overline{ST}$ è medio proporzionale tra $\overline{SA}$ ed $\overline{SB}$: il problema si riduce dunque a trovare i punti di r che siano medi proporzionali tra $\overline{SA}$ ed $\overline{SB}$. Vediamo come procedere su un esempio numerico. Siano $A(2,1)$, $B(4,2)$ ed r di equazione $2x-y=0$. La retta AB ha equazione $x-2y=0$ e interseca la retta r nell'origine O. Si ha poi

$$|\overline{OA}| = \sqrt{5}, \quad |\overline{OB}| = \sqrt{20}.$$

Si deve dunque avere

$$|\overline{OT}|^2 = |\overline{OA}| \cdot |\overline{OB}| = 10 \quad \Rightarrow \quad |\overline{OT}| = \sqrt{10}.$$

Per trovare i punti T su r che soddisfano questa condizione basta fare l'intersezione tra la circonferenza di centro O e raggio $|\overline{OT}|$ e la retta r stessa:

$$\begin{cases} x^2+y^2=10 \\ 2x-y=0 \end{cases}.$$

Si trovano due punti

$$T_1 = \left(-\sqrt{2}, -2\sqrt{2}\right), \quad T_2 = \left(\sqrt{2}, 2\sqrt{2}\right).$$

Ci saranno dunque due circonferenze che soddisfano le condizioni richieste: esse avranno centro sull'intersezione tra l'asse di $\overline{AB}$ e le perpendicolari ad r rispettivamente per T_1 e per T_2. L'asse di $\overline{AB}$ ha equazione $4x+2y-15=0$. Le perpendicolari ad r per T_1 e per T_2 hanno equazione, rispettivamente,

$$x+2y+5\sqrt{2}=0, \quad x+2y-5\sqrt{2}=0.$$

I centri delle due circonferenze saranno

$$O_1=\left(\frac{15+5\sqrt{2}}{3},-\frac{15+20\sqrt{2}}{6}\right), \quad O_2=\left(\frac{15-5\sqrt{2}}{3},-\frac{15-20\sqrt{2}}{6}\right).$$

Non resta che trovare i due raggi $|\overline{O_1T_1}|$ e $|\overline{O_2T_2}|$. Si veda la figura 8.33.

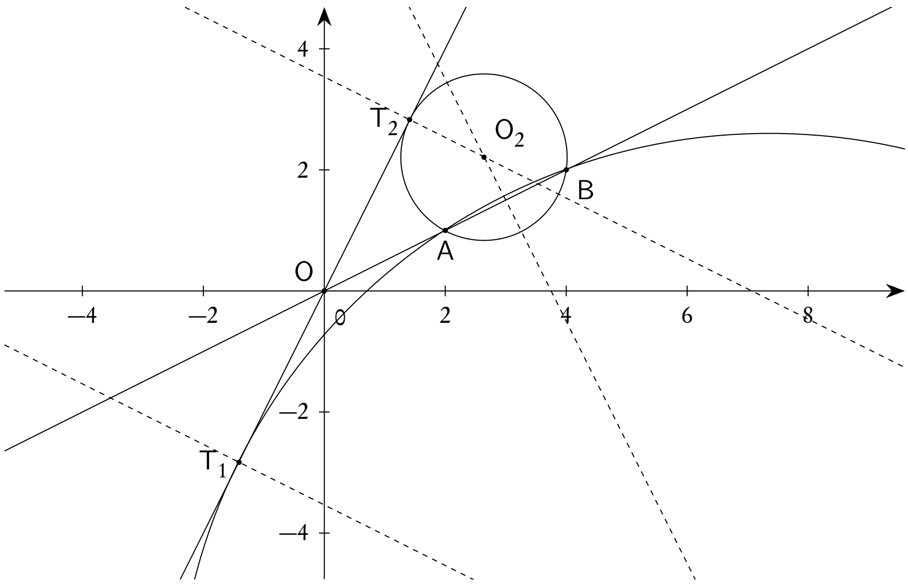


Figura 8.33.: *Circonferenze per due punti e tangenti a una retta*

8.7.3. Fasci di circonferenze

Con riferimento al paragrafo 8.5.4 sulle famiglie di curve, siano ora $\gamma_1\colon x^2+y^2+a_1x+b_1y+c_1=0$ e $\gamma_2\colon x^2+y^2+a_2x+b_2y+c_2=0$ le equazioni di due circonferenze. Una loro combinazione lineare

$$\lambda(x^2+y^2+a_1x+b_1y+c_1)+\mu(x^2+y^2+a_2x+b_2y+c_2)=0, \tag{8.44}$$

ovvero

$$(\lambda+\mu)(x^2+y^2)+(a_1\lambda+a_2\mu)x+(b_1\lambda+b_2\mu)y+(c_1\lambda+c_2\mu)=0, \tag{8.45}$$

se non è l'insieme vuoto, rappresenta una circonferenza se $\lambda+\mu\neq 0$, altrimenti rappresenta una retta.

Supposto $\lambda\neq 0$ la (8.45) si può riscrivere, dopo aver posto $k={}^{\mu}/_{\lambda}$,

$$(1+k)x^2+(1+k)y^2+(a_1+a_2k)x+(b_1+b_2k)x+(c_1+c_2k)=0 \tag{8.46}$$

che rappresenta la stessa famiglia di circonferenze per $k \neq -1$, con l'esclusione della circonferenza γ_1, che comunque si pensa appartenente al fascio per $k = \infty$.

Questa famiglia di circonferenze si chiama un *fascio di circonferenze*. Nel caso $\lambda + \mu = 0$, oppure $k = -1$, le (8.45), oppure (8.46), se non sono l'insieme vuoto, sono le equazioni di una retta, detta *asse radicale del fascio*. L'asse radicale può essere pensata come una circonferenza del fascio con "raggio infinito" e, per questo motivo, è chiamata circonferenza degenere del fascio.

Viceversa ogni equazione del tipo (8.46) rappresenta, se non è l'insieme vuoto o un'identità, una circonferenza ($k \neq -1$) o una retta ($k = -1$), l'asse radicale del fascio. I due circoli, se non vuoti, ottenuti dalla (8.46) per $k = 0$ e per $k = \infty$ si chiamano circoli base del fascio.

Valgono le seguenti proprietà dei fasci di circonferenze.

— Sostituendo, nella (8.46), le due circonferenze base con altre due circonferenze qualsiasi dello stesso fascio, si ottiene ancora lo stesso fascio. Se si combina una circonferenza del fascio con l'asse radicale, quando esistente, si ottiene un'equazione del tipo

$$x^2 + y^2 + (a_1 + a_2 k)x + (b_1 + b_2 k)x + (c_1 + c_2 k) = 0$$

che rappresenta ancora lo stesso fascio di circonferenze.

— I centri delle circonferenze (se non concentriche) del fascio appartengono a una retta perpendicolare all'asse radicale, detta *asse centrale del fascio*.

— Se le due circonferenze base del fascio si intersecano in due punti A e B, tutte le altre circonferenze del fascio passano per gli stessi punti e, viceversa, ogni circonferenza che passa per quei punti appartiene al fascio; la retta AB è l'asse radicale del fascio. In questo caso il fascio è detto *fascio ellittico*.

— Se le due circonferenze base del fascio sono tangenti in un punto A, tutte le circonferenze sono tra di loro tangenti in A e l'asse radicale è la tangente comune; viceversa ogni circonferenza tangente in A all'asse radicale fa parte del fascio. In questo caso il fascio è detto *fascio parabolico*. La circonferenza degenere di raggio nullo coincidente con l'intersezione tra l'asse radicale e l'asse centrale fa parte del fascio: essa si chiama anche un *punto limite* o *punto di Poncelet.*

— Se le due circonferenze base del fascio non hanno punti in comune tutte le circonferenze del fascio saranno prive di punti comuni. Se le due circonferenze base sono concentriche, tutte le circonferenze del fascio saranno concentriche: non esiste né l'asse radicale né l'asse centrale. Se le due circonferenze base non sono concentriche esistono due circonferenze di raggio nullo che appartengono al fascio, una da una parte e una dall'altra dell'asse radicale: esse si chiamano *punti limite* o *punti di Poncelet*. In questo caso il fascio si chiama *fascio iperbolico*.

Per trovare gli eventuali punti limite basta, dalla (8.46), trovare il raggio e uguagliarlo a 0:

$$r_k = \sqrt{\frac{(a_1 + a_2 k)^2 + (b_1 + b_2 k)^{-} 4(c_1 + c_2 k)(1 + k)}{(1 + k)^2}} = 0 :$$

si ottiene un'equazione di secondo grado in k che ha, dunque, al massimo due soluzioni.

Esempio 8.25. Trovare le caratteristiche del fascio di circonferenze di equazione

$$(1 + k)x^2 + (1 + k)y^2 - 6kx - 1 + 8k = 0, \quad \text{ovvero} \quad (x^2 + y^2 - 1) + k(x^2 + y^2 - 6x + 8) = 0.$$

Le due circonferenze base del fascio, $x^2+y^2-1=0$ e $x^2+y^2-6x+8=0$, non hanno punti in comune e non sono concentriche: si tratta dunque di un fascio iperbolico. L'asse radicale ha equazione $x={}^3/_2$. Il centro e il raggio della generica circonferenza sono

$$C_k\left(\frac{3k}{1+k},0\right), \quad r_k=\sqrt{\frac{9k^2-(8k-1)(1+k)}{(1+k)^2}}.$$

Uguagliando r_k a zero si trovano i due valori

$$k=\frac{7\pm3\sqrt{5}}{2},$$

che forniscono, sostituendoli nelle coordinate del centro, i due punti limite

$$L_{1,2}=\left(\frac{7\pm3\sqrt{5}}{3\pm\sqrt{5}},0\right).$$

La figura 8.34 illustra la situazione. In particolare:

- se $k\to-\infty$ le circonferenze del fascio tendono, dall'esterno, alla circonferenza base destra;
- se $-\infty<k<-1$ le circonferenze del fascio sono comprese tra l'asse radicale e la circonferenza base destra;
- se $k=-1$ si ottiene l'asse radicale;
- se $-1<k<0$ le circonferenze del fascio sono comprese tra la circonferenza base sinistra e l'asse radicale;
- se $k=0$ si ottiene la circonferenza base sinistra;
- se $0<k<{}^{7-3\sqrt{5}}/_2$ le circonferenze del fascio sono interne alla circonferenza base sinistra;
- se $k={}^{7-3\sqrt{5}}/_2$ si ottiene il punto limite L_2;
- se ${}^{7-3\sqrt{5}}/_2<k<{}^{7+3\sqrt{5}}/_2$ non si ha alcuna circonferenza;
- se $k={}^{7+3\sqrt{5}}/_2$ si ottiene il punto limite L_1;
- se ${}^{7+3\sqrt{5}}/_2<k<+\infty$ le circonferenze del fascio sono interne alla circonferenza base destra;
- se $k\to+\infty$ le circonferenze del fascio tendono, dall'interno, alla circonferenza base destra.

Le circonferenze tracciate nella figura si riferiscono ai valori di k seguenti, cominciando da L_2 e finendo con L_1: ${}^{(7-3\sqrt{5})}/_2$, 0.1, 0 (circonferenza base sinistra), −0.2, −0.4, −0.6, −1 (asse radicale), −2, −4, −10, $\mp\infty$ (circonferenza base destra), 10, ${}^{(7+3\sqrt{5})}/_2$.

Esempio 8.26. Trovare le caratteristiche del fascio di circonferenze di equazione

$$(1+k)x^2+(1+k)y^2-4x-1-k=0, \quad \text{ovvero} \quad (x^2+y^2-4x-1)+k(x^2+y^2-1)=0.$$

Le due circonferenze base del fascio $x^2+y^2-4x-1=0$, di centro C_2, e $x^2+y^2-1=0$, di centro C_1, hanno i punti $(0,\pm1)$ in comune: tutte le altre circonferenze del fascio passeranno per questi due punti. L'asse radicale, che corrisponde al valore $k=-1$, ha equazione $x=0$, cioè è l'asse delle ordinate.

La figura 8.35 illustra la situazione. In particolare:

- se $k\to-\infty$ le circonferenze del fascio hanno centro che tende a C_1 da sinistra;

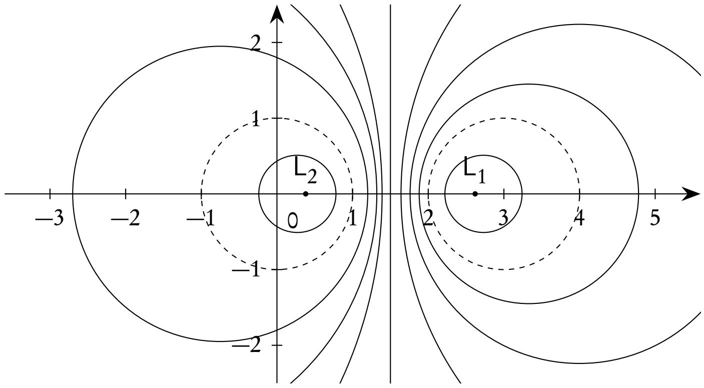


Figura 8.34.: *Il fascio di circonferenze di equazione* $(1+k)x^2+(1+k)y^2-6kx-1+8k=0$

— se $-\infty < k < -1$ le circonferenze del fascio hanno centro a sinistra di C_1;
— se $k=-1$ si ottiene l'asse radicale (circonferenza con centro all'infinito);
— se $-1 < k < 0$ le circonferenze del fascio hanno centro oltre C_2;
— se $k=0$ si ottiene la circonferenza base destra;
— se $0 < k < +\infty$ le circonferenze del fascio hanno centro tra C_1 e C_2;
— se $k \to +\infty$ le circonferenze del fascio hanno centro che tende a C_1 da destra.

Le circonferenze rappresentate in figura si riferiscono ai seguenti valori di k, cominciando da quella con centro più a sinistra e finendo con quella con centro più a destra: $k=-2$, $k=-4$, $k=-20$, $k=\mp\infty$ (circonferenza base sinistra), $k=10$, $k=2$, $k=0.5$, $k=0$ (circonferenza base destra), $k=-0.4$, $k=-0.8$.

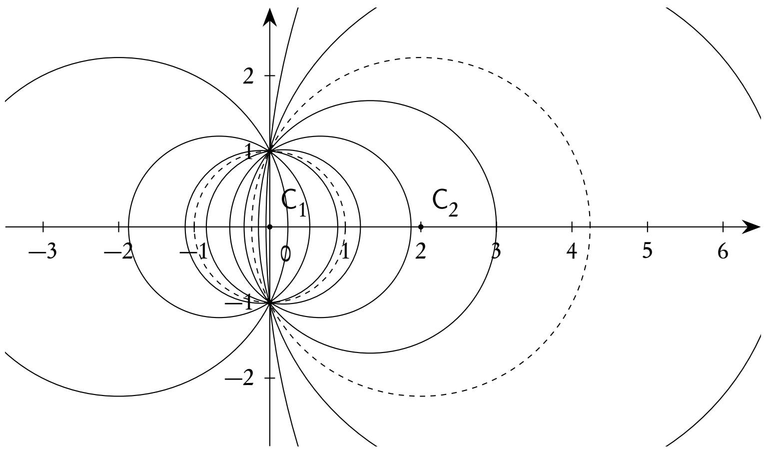


Figura 8.35.: *Il fascio di circonferenze di equazione* $(1+k)x^2+(1+k)y^2-4x-1-k=0$

Esempio 8.27. Trovare le caratteristiche del fascio di circonferenze di equazione

$$(1+k)x^2+(1+k)y^2-6x+51-k=0, \quad \text{ovvero} \quad (x^2+y^2-6x+5)+k(x^2+y^2-1)=0.$$

Le due circonferenze base del fascio $x^2+y^2-6x+5=0$, di centro C_2, e $x^2+y^2-1=0$, di centro C_1, sono tangenti nel punto $T(1,0)$: tutte le altre circonferenze del fascio saranno tra di loro tangenti in questo punto. L'asse radicale, che corrisponde a3 valore $k=-1$, ha equazione $x=1$. Le circonferenze hanno raggio

$$r_k=\sqrt{\left(\frac{3}{1+k}\right)^2-\frac{5-k}{1+k}}=\left|\frac{k-2}{1+k}\right|.$$

Dunque per $k=2$ si ottiene la circonferenza di raggio nullo, coincidente con il punto di tangenza comune alle circonferenze del fascio, mentre per ogni altro valore di $k\neq -1$ si ottiene una circonferenza non degenere.

La figura 8.36 illustra la situazione. In particolare:

- se $k\to-\infty$ le circonferenze del fascio hanno centro che tende a C_1 da sinistra;
- se $-\infty<k<-1$ le circonferenze del fascio hanno centro a sinistra di C_1;
- se $k=-1$ si ottiene l'asse radicale (centro all'infinito);
- se $-1<k<0$ le circonferenze del fascio hanno centro a destra di C_2;
- se $k=0$ si ottiene la circonferenza base destra;
- se $0<k<2$ le circonferenze del fascio hanno centro tra T e C_2;
- se $k=2$ si ottiene la circonferenza degenere coincidente con T (punto limite o di Poncelet);
- se $k>2$ le circonferenze del fascio hanno centro tra C_1 e T;
- se $k\to+\infty$ le circonferenze del fascio hanno centro che tende a C_1 da destra.

Le circonferenze tracciate nella figura si riferiscono ai valori di k seguenti, cominciando da quella con centro più a sinistra e finendo con quella con centro più a destra: $k=-2$, $k=-4$, $k=-10$, $k=\mp\infty$ (circonferenza base sinistra), $k=10$, $k=2$ (punto limite T), $k=1$, $k=0.2$, $k=0$ (circonferenza base destra), $k=-0.2$, $k=-0.5$. È inoltre tracciata l'asse radicale ($k=-1$).

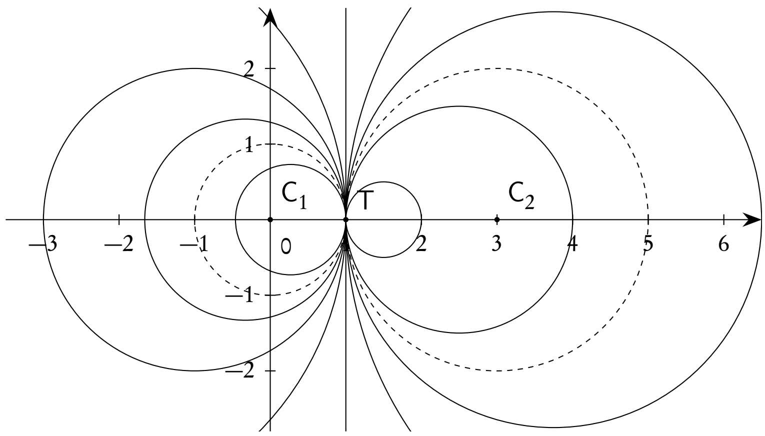


Figura 8.36.: *Il fascio di circonferenze di equazione* $(1+k)x^2+(1+k)y^2-6x+51-k=0$

8.8. La parabola in forma canonica

Sappiamo, vedi il teorema 8.9 nella pagina 236, che la parabola è il luogo geometrico dei punti di un piano equidistanti da un punto F detto fuoco e da una retta d detta direttrice, nell'ipotesi che[8] $\mathsf{F} \notin \mathsf{d}$.

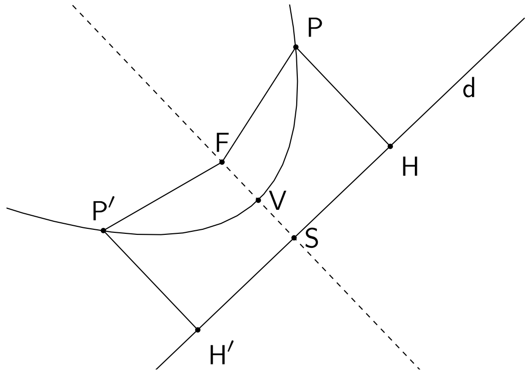


Figura 8.37.: *Parabola come luogo*

La retta per F ortogonale a d è chiaramente di simmetria per la curva in questione e si chiama *asse della parabola*. Detto S il punto di intersezione tra l'asse e la direttrice, il punto medio V di $\overline{\mathsf{FS}}$ appartiene alla parabola e si chiama *vertice*.

Noi ci limiteremo a trovare l'equazione della parabola nei casi in cui la direttrice sia parallela ad uno dei due assi coordinati, e ci riferiremo a queste situazioni parlando di *equazione canonica della parabola*.

Cominciamo a considerare il caso in cui la direttrice è parallela all'asse x e indichiamo con (p,q) le coordinate di F e con $y=k$ l'equazione della direttrice, tenendo conto che deve essere $q-k\neq 0$ perché il fuoco non deve stare sulla direttrice. Il vertice V avrà allora coordinate $(p, {}^{q+k}/_{2}$. Detto $\mathsf{P}(s,t)$ un generico punto del luogo cercato e $\mathsf{H}(s,k)$ la proiezione di P sulla retta d, si deve avere

$$|\overline{\mathsf{PF}}| = |\overline{\mathsf{PH}}| \quad \Rightarrow \quad \sqrt{(s-p)^2+(t-q)^2} = |t-k|. \tag{8.47}$$

Quadrando e semplificando l'ultima uguaglianza (senza problemi perché si tratta di uguaglianza tra numeri positivi e con un radicale sempre definito) si ottiene

$$2y(q-k) = x^2 - 2px + p^2 + q^2 - k^2 .$$

Si possono dividere ambo i membri per $2(q-k)$ che, come già osservato, è diverso da zero:

$$y = \frac{1}{2(q-k)}x^2 - \frac{p}{q-k}x + \frac{p^2+q^2-k^2}{2(q-k)}. \tag{8.48}$$

Posto

$$a = \frac{1}{2(q-k)}, \quad b = -\frac{p}{q-k}, \quad c = \frac{p^2+q^2-k^2}{2(q-k)} = \frac{p^2}{2(q-k)} + \frac{q+k}{2}, \tag{8.49}$$

[8] Se $\mathsf{F} \in \mathsf{d}$ il luogo degenera nella retta per F perpendicolare a d.

la (8.49) si riscrive nella forma

$$y = ax^2 + bx + c, \quad a \neq 0. \tag{8.50}$$

Il coefficiente a nella (8.50) è positivo se $q > k$, ovvero se F sta sopra la direttrice, negativo se $q < k$, ovvero se F sta sotto la direttrice. Nel primo caso la parabola è una curva piana convessa (oppure rivolge la concavità verso l'alto), nel secondo caso è una curva piana concava (oppure rivolge la concavità verso il basso).

Dunque ogni parabola con direttrice orizzontale (e quindi asse verticale) ha un'equazione del tipo (8.50), con i coefficienti dati dalle (8.49) e con a positivo o negativo a seconda che la concavità sia verso l'alto o verso il basso.

Vale anche il viceversa: ogni equazione del tipo (8.50) rappresenta una parabola il cui fuoco e la cui direttrice si trovano risolvendo il sistema

$$\begin{cases} a = \dfrac{1}{2(q-k)} \\ b = -\dfrac{p}{q-k} \\ c = \dfrac{p^2}{2(q-k)} + \dfrac{q+k}{2} \end{cases}$$

nelle incognite p, q, k. Si può procedere come di seguito indicato.

$$\begin{cases} q - k = \dfrac{1}{2a} \\ b = -2ap \Rightarrow p = -\dfrac{b}{2a} \\ c = ap^2 + \dfrac{q+k}{2} \Rightarrow c = \dfrac{b^2}{4a} + \dfrac{q+k}{2} \end{cases} .$$

Dalla prima e terza equazione si ottiene il seguente sistema

$$\begin{cases} q - k = \dfrac{1}{2a} \\ q + k = \dfrac{4ac - b^2}{2a} \end{cases} ,$$

da cui è facile ricavare q e k. Posto, com'è tradizione, $\Delta = b^2 - 4ac$, si trova

$$\mathsf{F}\left(-\frac{b}{2a}, \frac{1-\Delta}{4a}\right); \quad \mathsf{d}: y = \frac{-1-\Delta}{4a}. \tag{8.51}$$

Per quanto riguarda il vertice si avrà, tenendo conto delle proprietà già richiamate,

$$\mathsf{V}\left(-\frac{b}{2a}, -\frac{\Delta}{4a}\right). \tag{8.52}$$

Per quanto riguarda il vertice si tenga comunque conto che, essendo V un punto della parabola, se ne può ricavare l'ordinata semplicemente sostituendo l'ascissa nell'equazione (8.50).

Dalla (8.50) si deduce anche che la parabola con asse verticale è il grafico di una funzione polinomiale di secondo grado: si veda il paragrafo 4.8.2 nella pagina 120.

Nel caso in cui la direttrice sia verticale, e quindi l'asse orizzontale, si otterranno risultati simili, previo lo scambio della x con la y. Riportiamo, per completezza, le formule relative.

Equazione di una parabola con asse orizzontale:

$$x = ay^2 + by + c \tag{8.53}$$

In questo caso la parabola rivolgerà la concavità verso destra se $a > 0$, verso sinistra se $a < 0$.

Coordinate del fuoco ed equazione della direttrice per una parabola con asse orizzontale:

$$\mathsf{F}\left(\frac{1-\Delta}{4a}, -\frac{b}{2a}\right); \quad \mathsf{d}\colon x = \frac{-1-\Delta}{4a}. \tag{8.54}$$

Coordinate del vertice per una parabola con asse orizzontale:

$$\mathsf{V}\left(-\frac{\Delta}{4a}, -\frac{b}{2a}\right). \tag{8.55}$$

Per quanto riguarda il vertice si tenga comunque ora conto che, essendo V un punto della parabola, se ne può ricavare l'ascissa semplicemente sostituendo l'ordinata nell'equazione (8.53).

La parabola con asse orizzontale non può essere pensata come grafico di una funzione, se si vuole mantenere la convenzione che il dominio si rappresenta sull'asse delle x, mentre il codominio su quello delle y: a un punto del dominio possono corrispondere anche due punti del codominio, cosa vietata nel caso delle funzioni.

Occupiamoci con qualche maggiore dettaglio del caso di parabola con asse verticale, perché più frequente nelle applicazioni e perché, come già osservato, in questo caso la parabola può essere pensata come grafico di una funzione, mantenendo le convenzioni standard.

Se nella equazione generica (8.50) $c = 0$, la parabola passa per l'origine (cosa che succede per tutte le curve razionali intere la cui equazione è prima di termine noto); se anche $b = 0$ la parabola ha vertice nell'origine. In questo caso è facile constatare che il parametro a caratterizza la maggiore o minore "apertura" della parabola: si veda la figura 8.38

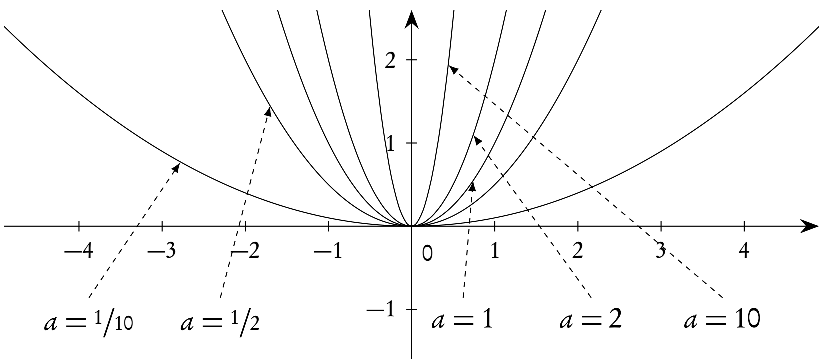


Figura 8.38.: *Parabola con vertice nell'origine e diverse "aperture"*

Per i casi in cui b, oppure c, oppure entrambi, siano diversi da zero, possiamo utilmente applicare la tecnica del completamento dei quadrati che ci consentirà di tracciare la parabola, senza far ricorso a formule particolari, semplicemente come traslata di una parabola per l'origine. Tenendo conto della (3.11), l'equazione di una parabola con asse verticale si può scrivere nella forma

$$y = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}, \quad \text{ovvero} \quad y + \frac{b^2 - 4ac}{4a} = a\left(x + \frac{b}{2a}\right)^2. \tag{8.56}$$

Da qui si capisce subito che la generica parabola $y = ax^2 + bx + c$ è semplicemente ottenuta traslando la parabola $y = ax^2$ in modo che il vertice anziché nell'origine si trovi nel punto

$$\mathsf{V}\left(-\frac{\Delta}{4a}, -\frac{b}{2a}\right),$$

formula che fornisce appunto le coordinate del vertice riportate nella (8.55).

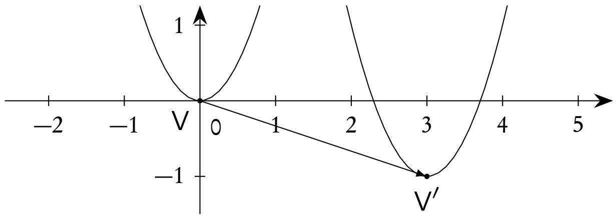


Figura 8.39.: *La parabola $y = 2x^2$ e la sua traslata $y = 2x^2 - 12x + 17$*

Per esempio la parabola $y = 2x^2 - 12x + 17$ si può riscrivere nella forma $y + 1 = 2(x - 3)^2$, che evidenzia il fatto che si tratta della traslazione della parabola $y = 2x^2$ in modo che il vertice trasli nel punto $(3, -1)$. Si veda al figura 8.39.

8.8.1. Tangenti a una parabola

Per trovare le tangenti a una parabola si può procedere con i metodi generali indicati per le coniche; nel caso però si debba trovare la tangente ad una parabola con asse verticale in un *suo punto* $\mathsf{P}(x_0, y_0)$, conviene utilizzare le considerazioni[9] che faremo qui di seguito.

Sia data una parabola con asse verticale: $y = ax^2 + bx + c$ e sia $\mathsf{P}(x_0, y_0)$ un suo punto. Cerchiamo ora la tangente con le formule di sdoppiamento.

$$y = ax^2 + bx + c \quad \Rightarrow \quad \frac{y + y_0}{2} = ax_0x + b\frac{x + x_0}{2} + c \quad \Rightarrow \quad y = (2ax_0 + b)x + bx_0 + 2c - y_0.$$

La cosa più importante nella formula appena ottenuta è il valore del coefficiente angolare della tangente: si può concludere che se $\mathsf{P}(x_0, y_0)$ è un punto di una parabola con asse verticale, il coefficiente angolare della tangente in P alla parabola è

$$m = 2ax_0 + b. \tag{8.57}$$

[9] La formula che otterremo è legata al fatto che la parabola con asse verticale è pensabile come grafico di una funzione, e il coefficiente angolare della tangente è la derivata della funzione stessa: tutto questo sarà argomento dei futuri corsi di analisi.

Esempio 8.28. Trovare la tangente alla parabola $y = 2x^2 + 3x - 1$ nel suo punto di ascissa 1.

Se il punto di ascissa 1 appartiene alla parabola, la sua ordinata deve essere 4. Applicando la (8.57) si ottiene

$$m = 2 \times 2 \times 1 + 3 = 7.$$

La tangente richiesta sarà allora

$$y - 4 = 7(x - 1) \quad \Rightarrow \quad 7x - y - 3 = 0.$$

Applicando le formule di sdoppiamento i calcoli sarebbero stati altrettanto rapidi:

$$y = 2x^2 + 3x - 1 \quad \Rightarrow \quad \frac{y+4}{2} = 2x + 3\frac{x+1}{2} - 1 \quad \Rightarrow \quad 7x - y - 3 = 0.$$

Applicando invece il metodo del "$\Delta = 0$" si sarebbe ottenuto il sistema

$$\begin{cases} y = 2x^2 + 3x - 1 \\ y - 4 = m(x - 1) \end{cases},$$

di equazione risolvente

$$2x^2 + (3 - m)x + m - 5 = 0.$$

Annullando il discriminante si ottiene

$$m^2 - 14m + 49 = 0 \quad \Rightarrow \quad m = 7,$$

in perfetto accordo con quanto già trovato.

Esempio 8.29. Trovare le tangenti alla parabola di equazione $y = x^2 - 4x + 3$ condotte dal punto $P(2, -5)$.

Poiché una retta verticale non può essere tangente ad una parabola con asse verticale, possiamo considerare il fascio di rette non verticali per P e metterlo a sistema con la parabola.

$$\begin{cases} y = x^2 - 4x + 3 \\ y + 5 = m(x - 2) \end{cases}.$$

L'equazione risolvente il sistema è

$$x^2 - (4 + m)x + 3 + 2m + 5 = 0,$$

il cui discriminante si annulla se $m^2 - 16 = 0$, ovvero $m = \pm 4$. Si trovano le due tangenti

$$4x - y - 13 = 0 \quad \text{e} \quad 4x + y - 3 = 0.$$

8.8.2. Come trovare l'equazione di una parabola

L'equazione di una parabola con asse verticale o orizzontale dipende da tre parametri, a, b, c: saranno dunque necessarie tre condizioni indipendenti e compatibili per determinarla. Riepiloghiamo di seguito quelle che sono le situazioni più comuni e successivamente proporremo alcuni esempi significativi, tenendo conto che deve essere precisato se la parabola cercata ha asse verticale o orizzontale.

— La conoscenza del fuoco e della direttrice consente di scrivere subito l'equazione applicando la definizione.
— Il passaggio per un punto fornisce una condizione di primo grado sui parametri a, b, c.
— La conoscenza del vertice fornisce due condizioni: una è il passaggio della parabola per il vertice, l'altra si ottiene tenendo conto che deve essere $x_V = -b/2a$. Non conviene utilizzare l'ordinata del vertice che fornisce una condizione di secondo grado.
— La conoscenza del fuoco fornisce due condizioni, ottenute uguagliando le coordinate del fuoco alle due formule generali per il fuoco stesso.
— La conoscenza della direttrice fornisce una condizione ottenuta utilizzando la formula generale per la sua equazione.
— La conoscenza della tangente e del punto P di tangenza fornisce due condizioni: una è il passaggio per il punto, l'altra si ottiene dalla formula $m = 2ax_0 + b$.
— La conoscenza di una tangente senza il punto di tangenza fornisce una condizione ottenuta con il metodo del "$\Delta = 0$".

Esempio 8.30. Trovare l'equazione della parabola avente

$$F\left(\frac{5}{4}, 0\right) \quad \text{e} \quad V\left(\frac{5}{4}, -\frac{1}{8}\right).$$

La parabola ha asse verticale, in quanto l'ascissa del fuoco e del vertice coincidono. Si ha il sistema

$$\begin{cases} -\dfrac{b}{2a} = \dfrac{5}{4} & \text{(ascissa del vertice e del fuoco)} \\ \dfrac{1-\Delta}{4a} = 0 & \text{(ordinata del fuoco)} \\ -\dfrac{1}{8} = a\left(\dfrac{5}{4}\right)^2 + b\left(\dfrac{5}{4}\right) + c & \text{(passaggio per il vertice)} \end{cases}.$$

Si trova $y = 2x^2 - 5x + 3$.

Esempio 8.31. Trovare la parabola che ha direttrice $y = 21/4$ e vertice $V(0,5)$.

La parabola ha asse verticale, avendo direttrice orizzontale. Si ha il sistema

$$\begin{cases} -\dfrac{b}{2a} = 0 & \text{(ascissa del vertice)} \\ 5 = c & \text{(passaggio per il vertice)} \\ \dfrac{-1-\Delta}{4a} = \dfrac{21}{4} & \text{(per l'equazione delle direttrice)} \end{cases}.$$

Si trova $y = -x^2 + 5$.

Esempio 8.32. Trovare l'equazione della parabola con asse orizzontale passante per $(-3,0)$, $(7,-2)$, $(-2,1)$.

Si ha il sistema (passaggio per i tre punti)

$$\begin{cases} -3 = c \\ 7 = 4a - 2b + c \\ -2 = a + b + c \end{cases}.$$

Si trova $x = 2y^2 - y - 3$.

Esempio 8.33. Trovare l'equazione della parabola con asse verticale, passante per $(-1,1)$, $(2,1)$ e tangente alla retta $x - y + 3 = 0$.

Per la tangenza si considera il sistema

$$\begin{cases} y = ax^2 + bx + c \\ x - y + 3 = 0 \end{cases},$$

la cui equazione risolvente è:

$$ax^2 + (b-1)x + c - 3 = 0.$$

Annullando il discriminante si ottiene un'equazione in a, b, c che, messa a sistema con le condizioni di passaggio per i due punti dati, fornisce il seguente sistema.

$$\begin{cases} (b-1)^2 - 4a(c-3) = 0 \\ 1 = a - b + c \\ 1 = 4a + 2b + c \end{cases}.$$

Si trovano le due soluzioni

$$y = -x^2 + x + 3 \quad \text{e} \quad y = -\frac{1}{9}x^2 + \frac{1}{9}x + \frac{11}{9}.$$

Esempio 8.34. Trovare l'equazione della parabola con asse verticale, passante per $\mathsf{A}(-1,2)$, B di ascissa 1 e tangente in B alla retta $3x - y - 3 = 0$.

Se B sta sulla retta, la sua ordinata deve essere 0. Imponendo il passaggio per A e B e la condizione $m = 2ax_0 + b$, con $m = 3$ e $x_0 = 1$, si ottiene il sistema

$$\begin{cases} 2 = a - b + c \\ 0 = a + b + c \\ 3 = 2a + b \end{cases}.$$

Si trova $y = 2x^2 - x - 1$.

8.8.3. Una notevole proprietà della parabola

Vogliamo provare, analiticamente, la seguente proprietà delle parabole: se la parabola è una superficie perfettamente riflettente, ogni raggio uscente dal fuoco si riflette parallelamente all'asse e, viceversa, ogni raggio parallelo all'asse si riflette passando per il fuoco. Si tratta di una proprietà notevole[10] e costituisce inoltre un ottimo esercizio conclusivo si rette e parabole.

Con riferimento alla figura 8.40, consideriamo una parabola con asse verticale, di equazione $y = ax^2 + bx + c$, e un suo generico punto $\mathsf{P}(p,q)$. Siano poi

— t la retta tangente in P alla parabola;

[10] Per l'ellisse vale una proprietà simile: i raggi uscenti da un fuoco si riflettono sull'ellisse passando per l'altro fuoco. La proprietà valida per la parabola si può pensare come la situazione limite dell'ellisse quando uno dei due fuochi si allontana indefinitamente.

— n la retta normale (perpendicolare) in P alla parabola;
— f la retta FP;
— r la retta per P parallela all'asse della parabola.

Per provare che gli angoli α e β sono uguali basterà provare che n e t sono le bisettrici degli angoli individuati da f e r.

Indichiamo con $\mathsf{F}(s,u)$ le coordinate del fuoco, ove

$$s=-\frac{b}{2a}, \quad u=\frac{1-b^2+4ac}{4a}.$$

Essendo P un punto della parabola si ha $q=ap^2+bp+c$. Inoltre il coefficiente angolare della tangente t e della normale n sono dati, rispettivamente, da

$$m_{\mathsf{t}}=2ap+b, \quad m_{\mathsf{n}}=\frac{-1}{2ap+b}.$$

L'equazione della retta r è $x-p=0$; per la retta f si ha

$$(x-p)(u-q)=(y-q)(s-p) \quad \text{ovvero} \quad x(u-q)-y(s-p)-pu+qs=0.$$

Per trovare le bisettrici utilizziamo il fatto che esse sono il luogo dei punti equidistanti dalle due rette r e f. Dette (x,y) le coordinate di un generico punto su queste bisettrici si ha

$$|x-s|=\frac{|x(u-q)-y(s-p)-pu+qs|}{\sqrt{(u-q)^2+(s-p)^2}}, \quad \text{ovvero} \quad x-s=\pm\frac{x(u-q)-y(s-p)-pu+qs}{\sqrt{(u-q)^2+(s-p)^2}}.$$

Consideriamo per esempio la prima delle due, cioè quella con il segno "+" al secondo membro. Il suo coefficiente angolare è:

$$m_1=\frac{u-q-\sqrt{(u-q)^2+(s-p)^2}}{s-p}.$$

Dovremo dunque provare che questo coefficiente angolare coincide con quello di una delle due rette t o n, essendo chiaro che si verificherà una sola delle due possibilità. Proveremo che $m_1=m_{\mathsf{n}}$ cioè che

$$\frac{-1}{2ap+b}=\frac{u-q-\sqrt{(u-q)^2+(s-p)^2}}{s-p}.$$

Sostituendo, al denominatore del secondo membro, s con la sua espressione e poi semplificando e riordinando si ottiene

$$\sqrt{(u-q)^2+(s-p)^2}=(u-q)-\frac{1}{2a}.$$

Quadrando, semplificando e sostituendo i valori di s, u e q in funzione di a, b e c si conclude.

Questa notevole proprietà delle parabole è sfruttata, per esempio, nelle parabole dei fari delle automobili: se la lampada è accesa nel fuoco, i raggi vengono riflessi uscendo parallelamente all'asse della parabola (posizione di "fari abbaglianti"); un'altra applicazione si ha nelle parabole usate come antenne riceventi: raggi provenienti da una sorgente distante possono essere considerati tra di loro paralleli e essere concentrati nel fuoco della antenna parabolica.

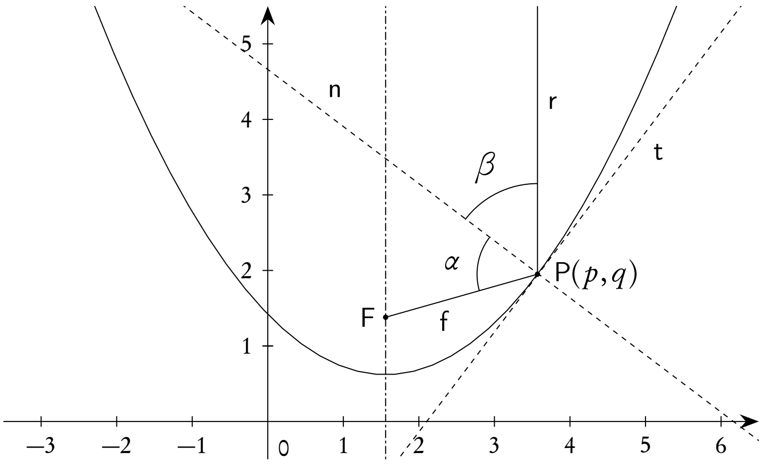


Figura 8.40.: *Parabola e riflessione dei raggi luminosi*

8.9. L'ellisse e l'iperbole in forma canonica

8.9.1. L'ellisse

Sappiamo, vedi il teorema 8.7, che l'ellisse è il luogo geometrico dei punti del piano tali che la somma delle distanze da due punti fissi detti fuochi sia costante (e maggiore della distanza tra i fuochi stessi).

La retta passante per F_1 ed F_2, che è di simmetria, si chiama *asse focale* o *asse maggiore*; il punto medio C del segmento $\overline{F_1F_2}$ si chiama *centro*; la retta per il centro e perpendicolare all'asse focale, che è di simmetria, si chiama *asse minore*; i punti di intersezione dell'ellisse con l'asse focale e con l'asse trasverso si chiamano *vertici*. Il segmento $\overline{CV_2}$, oppure $\overline{CV_1}$, si chiama *semiasse maggiore*, il segmento $\overline{CV_3}$, oppure $\overline{CV_4}$, si chiama *semiasse minore*.

La circonferenza si può ritenere caso limite dell'ellisse quando i due fuochi coincidono.

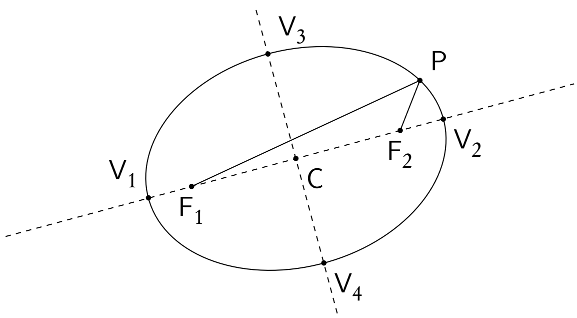


Figura 8.41.: *Ellisse come luogo*

Per trovare l'equazione dell'ellisse ci metteremo in una situazione di comodo, introducendo un sistema Oxy in cui l'asse delle ascisse coincida con l'asse focale e il centro con l'origine degli assi. Indicheremo con $-c$ e c rispettivamente le ascisse dei due fuochi ($c > 0$): c è dunque la *semidistanza focale*. Indicheremo inoltre con $2a$ la somma costante delle distanze di un generico punto dell'ellisse dai due fuochi. Con riferimento alla figura 8.41, è chiaro che

$$2a = |\overline{V_1V_2}|.$$

Detto $P(x,y)$ un generico punto dell'ellisse, si deve avere

$$\sqrt{(x-c)^2+y^2}+\sqrt{(x+c)^2+y^2}=2a, \quad \text{ovvero} \quad \sqrt{(x+c)^2+y^2}=2a-\sqrt{(x-c)^2+y^2}. \tag{8.58}$$

Quadrando e semplificando l'ultima equazione si ottiene

$$a\sqrt{(x-c)^2+y^2}=a^2-cx.$$

Una nuova elevazione al quadrato porta, dopo semplificazione, a

$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2). \tag{8.59}$$

Poiché abbiamo elevato due volte al quadrato potremmo aver introdotto soluzioni estranee; in realtà si può mostrare che la (8.59) è perfettamente equivalente alla (8.58), ovvero che l'elevazione al quadrato era sempre lecita. Poiché $a>c>0$, si ha anche $a^2>c^2$, e dunque esiste $b>0$, e minore di a, tale che $b^2=a^2-c^2$. La (8.59) assume la forma

$$b^2x^2+a^2y^2=a^2b^2, \tag{8.60}$$

che viene abitualmente riscritta nella forma, detta canonica,

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{8.61}$$

I punti $(\pm a,0)$ e $(0,\pm b)$ sono i quattro vertici dell'ellisse e dunque a è il semiasse maggiore, b il semiasse minore. Noti a e b si ricava facilmente c:

$$c=\sqrt{a^2-b^2} \quad \text{con} \quad a>b. \tag{8.62}$$

Se invece, pur mantenendo l'asse focale parallelo all'asse delle ascisse, spostiamo il centro dell'ellisse nel punto $C(x_C,y_C)$, l'equazione diventa

$$\frac{(x-x_C)^2}{a^2}+\frac{(y-y_C)^2}{b^2}=1. \tag{8.63}$$

Scegliendo invece l'asse focale sull'asse y, con i fuochi nei punti $(0,\pm c)$, e indicando questa volta con $2b$ la somma costante delle distanze di un generico punto dell'ellisse dai due fuochi, troveremmo ancora una equazione dello stesso tipo della (8.61), ma questa volta avendo posto con $a^2=b^2-c^2$ e quindi con $a<b$. Noti a e b il valore di c è dato questa volta da

$$c=\sqrt{b^2-a^2} \quad \text{con} \quad a<b. \tag{8.64}$$

Le equazioni ottenute per l'ellisse con asse focale l'asse x oppure l'asse y si dicono anche *equazioni dell'ellisse riferita agli assi*, in quanto gli assi dell'ellisse coincidono con gli assi del sistema di coordinate.

Le formule ottenute si possono compendiare vantaggiosamente come segue.

L'equazione di un'ellisse, con centro nel punto $\mathsf{C}(x_\mathsf{C}, y_\mathsf{C})$ e con i fuochi sono su una parallela agli assi coordinati, è un'equazione del tipo

$$\frac{(x-x_\mathsf{C})^2}{a^2}+\frac{(y-y_\mathsf{C})^2}{b^2}=1, \tag{8.65}$$

dove a e b sono i semiassi. Se l'asse focale è orizzontale si ha $a>b$, se è verticale si ha $a<b$. In ogni caso la semidistanza focale è data da

$$c=\sqrt{|a^2-b^2|}. \tag{8.66}$$

Ci si potrà anche utilmente riferire al parametro a come alla lunghezza del semiasse orizzontale, al parametro b come alla lunghezza del semiasse verticale: a seconda che $a>b$ oppure $a<b$, i fuochi staranno sull'asse orizzontale oppure su quello verticale.

8.9.2. L'iperbole

Sappiamo, vedi il teorema 8.8, che l'iperbole è il luogo geometrico dei punti del piano tali che il modulo della differenza delle distanze da due punti fissi detti fuochi sia costante e non nullo.

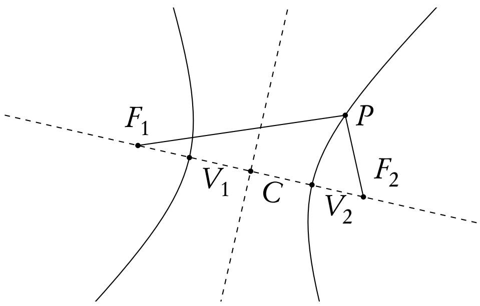


Figura 8.42.: *L'iperbole come luogo*

Si mantengono per l'iperbole le stesse denominazioni usate per l'ellisse, quando applicabili. In particolare la retta passante per F_1 ed F_2, che è di simmetria, si chiama *asse focale* o *asse principale* o *asse trasverso*; il punto medio C del segmento $\overline{\mathsf{F}_1\mathsf{F}_2}$ si chiama *centro*; la retta per il centro e perpendicolare all'asse focale, che è di simmetria, si chiama *asse secondario* o *asse non trasverso*; i punti di intersezione dell'iperbole con l'asse focale *vertici*. Il segmento $\overline{\mathsf{CV}_2}$, oppure $\overline{\mathsf{CV}_1}$, si chiama *semiasse trasverso*.

Per trovare l'equazione dell'iperbole ci si mette sempre nella comoda situazione in cui l'asse focale sia orizzontale o verticale. Si usano le stesse convenzioni già usate per l'ellisse: se l'asse focale è orizzontale $2a$ è il valore costante del modulo della differenza delle distanze di un punto dell'iperbole dai due fuochi, se invece l'asse focale è verticale tale costante si indica con $2b$. La distanza focale si indica sempre con $2c$. Ripetendo calcoli e considerazioni simili a quelle fatte per l'ellisse si perviene a un'equazione del tipo

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad \text{(se l'asse focale è l'asse } x\text{)}, \tag{8.67}$$

oppure del tipo

$$-\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad \text{(se l'asse focale è l'asse } y\text{)}, \tag{8.68}$$

ove, in entrambi i casi, si ha

$$c=\sqrt{a^2+b^2}. \tag{8.69}$$

Le rette di equazione

$$\frac{x}{a}=\pm\frac{y}{b}, \quad \text{ovvero} \quad y=\frac{b}{a}x \quad \text{e} \quad y=-\frac{b}{a}x \tag{8.70}$$

si chiamano *asintoti dell'iperbole*. Esse hanno la caratteristica di avvicinarsi indefinitamente ai due rami dell'iperbole, man mano che ci si allontana dall'origine. Con una locuzione parzialmente impropria si possono definire "tangenti all'infinito". Si veda la figura 8.43, che evidenzia anche il significato dei parametri b ed a dell'equazione, tenendo conto che nel primo caso i vertici hanno ascissa $\pm a$, nel secondo hanno ordinata $\pm b$.

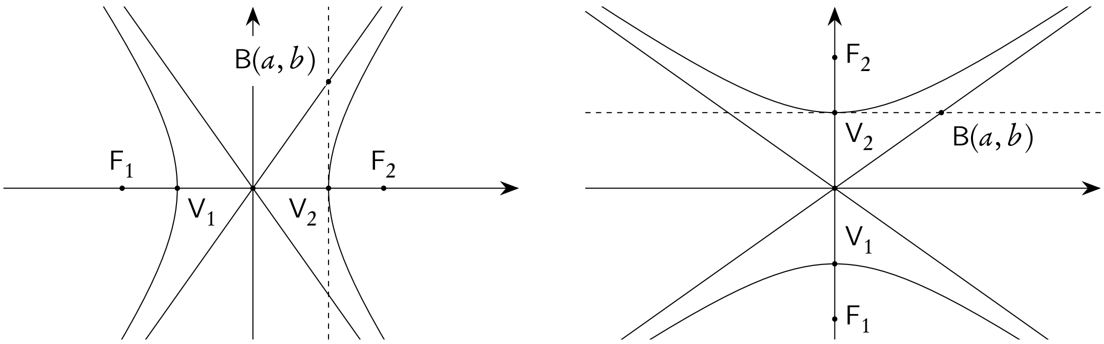


Figura 8.43.: *Iperbole e suoi asintoti*

Le equazioni ottenute per l'iperbole con asse focale l'asse x oppure l'asse y si dicono anche, in analogia al caso dell'ellisse, *equazioni dell'iperbole riferita agli assi*, in quanto gli assi dell'iperbole coincidono con gli assi del sistema di coordinate.

Se poi il centro non coincide con l'origine degli assi basterà eseguire una traslazione, simile a quella già fatta per l'ellisse, ottenendo l'equazione

$$\frac{(x-x_C)^2}{a^2}-\frac{(y-y_C)^2}{b^2}=1 \quad \text{(se l'asse focale è orizzontale)}, \tag{8.71}$$

oppure la

$$-\frac{(x-x_C)^2}{a^2}+\frac{(y-y_C)^2}{b^2}=1 \quad \text{(se l'asse focale è verticale)}. \tag{8.72}$$

Analogamente al caso dell'ellisse ci si potrà anche utilmente riferire al parametro a come alla lunghezza del semiasse orizzontale, al parametro b come alla lunghezza del semiasse verticale: a seconda della successione dei segni del primo membro dell'equazione, i fuochi staranno sull'asse orizzontale oppure su quello verticale.

8.9.3. Riepilogo su ellisse ed iperbole

Come si può constatare dalle considerazioni svolte nei paragrafi precedenti, le equazioni canoniche dell'ellisse e dell'iperbole sono molto simili: è vantaggioso compendiare le formule come segue.

Le equazioni canoniche di un'ellisse o di una iperbole con assi paralleli agli assi coordinati, aventi semidistanza focale c, semiassi orizzontale a e verticale b e centro di coordinate (x_C, y_C) sono del tipo

$$\pm\frac{(x-x_C)^2}{a^2}\pm\frac{(y-y_C)^2}{b^2}=1. \tag{8.73}$$

Si possono presentare 4 casi, a seconda della successione dei segni del primo membro.

1. "$-,-$": l'equazione non ha nessuna soluzione, in quanto il primo membro è minore o uguale a zero, il secondo strettamente positivo.
2. "$+,+$": si tratta di un'ellisse, con centro in $\mathrm{C}(x_C, y_C)$, con asse orizzontale di lunghezza $2a$ e verticale di lunghezza $2b$; i fuochi stanno sull'asse maggiore.
3. "$+,-$": si tratta di un'iperbole, con centro in $\mathrm{C}(x_C, y_C)$, con asse orizzontale di lunghezza $2a$ e verticale di lunghezza $2b$; i fuochi stanno sull'asse orizzontale.
4. "$-,+$": si tratta di un'iperbole, con centro in $\mathrm{C}(x_C, y_C)$, con asse orizzontale di lunghezza $2a$ e verticale di lunghezza $2b$; i fuochi stanno sull'asse verticale.

Nel caso dell'ellisse la semidistanza focale è data da

$$c=\sqrt{|a^2-b^2|}. \tag{8.74}$$

Nel caso dell'iperbole la semidistanza focale è data da

$$c=\sqrt{a^2+b^2}. \tag{8.75}$$

Sempre nel caso dell'iperbole gli asintoti hanno equazione

$$\frac{x-x_C}{a}=\pm\frac{y-y_C}{b}, \tag{8.76}$$

che si possono ricavare semplicemente uguagliano a zero il primo membro della (8.73).

Per tracciare le curve si può procedere nel seguente modo.

1. Si individua il centro $\mathrm{C}(x_C, y_C)$ e i quattro vertici di coordinate $(x_C \pm a, y_C)$ e $(x_C, y_C \pm b)$.
2. Si traccia il rettangolo con assi paralleli agli assi coordinati e passante per i quattro vertici.
3. Se si tratta di un'ellisse essa è tangente internamente a questo rettangolo.
4. Se si tratta di un'iperbole si tracciano anche le rette che contengono le diagonali del rettangolo: esse sono gli asintoti dell'iperbole. Dopodiché l'iperbole è tangente esternamente al rettangolo in due dei quattro vertici e asintoticamente tangente agli asintoti e avrà l'asse focale orizzontale oppure verticale a seconda della successione dei segni del primo membro.

Si veda la figura 8.44, dove abbiamo considerato il caso in cui il centro è situato nell'origine.

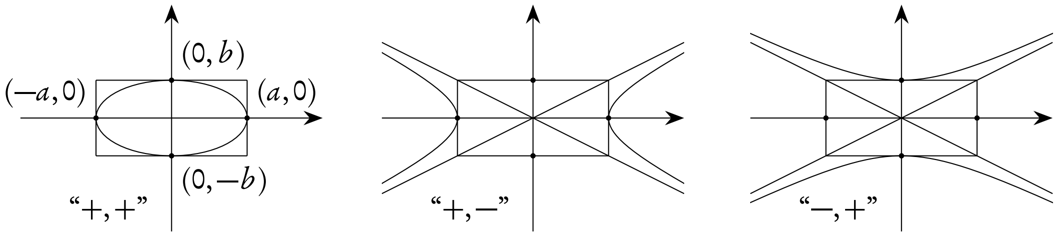


Figura 8.44.: *Tracciamento di ellissi e iperboli (centro nell'origine)*

Eseguendo i calcoli e le semplificazioni, l'equazione (8.73) si può porre nella seguente forma[11]

$$Ax^2 + Cy^2 + Dx + Ey + F = 0, \tag{8.77}$$

con opportuni valori dei coefficienti: si tratta dunque di un'equazione di secondo grado in due incognite, priva del termine misto e inoltre con i due coefficienti di x^2 e y^2 non nulli. Discuteremo nel paragrafo 8.10 il problema inverso, cioè quello di determinare le caratteristiche e tracciare il grafico delle soluzioni di un'equazione generale di secondo grado in due incognite, priva del termine misto e con i coefficienti di x^2 e y^2 non nulli.

8.9.4. Eccentricità

Sia per l'ellisse che per l'iperbole il rapporto fra la semidistanza focale c e il semiasse focale (che nel caso dell'ellisse è sempre il semiasse maggiore) si chiama *eccentricità* e si indica con e. Per l'ellisse si ha sempre $e < 1$, per l'iperbole $e > 1$. Si potrebbe definire l'eccentricità anche per la parabola, ma questo esula dagli scopi di questo corso: in ogni caso segnaliamo che l'eccentricità della parabola sarebbe 1.

L'eccentricità è una misura di quanto "schiacciate" siano l'ellisse e l'iperbole. Per esempio nel caso dell'ellisse con fuochi sull'asse orizzontale l'eccentricità, che vale c/a, è tanto più vicina a 1 quanto più i fuochi sono vicini ai vertici sull'asse focale, come si capisce subito tenendo conto del valore di c

$$e = \frac{c}{a} = \frac{\sqrt{a^2 - b^2}}{a} = \sqrt{1 - \frac{b^2}{a^2}}.$$

Si veda la figura 8.45 dove sono rappresentate tre diverse situazioni.

Si noti che la circonferenza ha eccentricità nulla, in quanto $c = 0$ (oltreché $a = b$). Nel caso dell'iperbole, se $a = b$, l'eccentricità vale $\sqrt{2}$, in quanto per l'iperbole con asse focale orizzontale si ha

$$e = \frac{c}{a} = \frac{\sqrt{a^2 + b^2}}{a} = \sqrt{1 + \frac{a^2}{b^2}},$$

da cui, appunto, $e = \sqrt{2}$ se $a = b$. Tratteremo questa situazione nel paragrafo 8.11 relativo all'iperbole equilatera.

[11] Abbiamo utilizzato le lettere A, C, D, E, F per i coefficienti, "saltando" la lettera B, che riserviamo al coefficiente di del termine misto xy, quando presente.

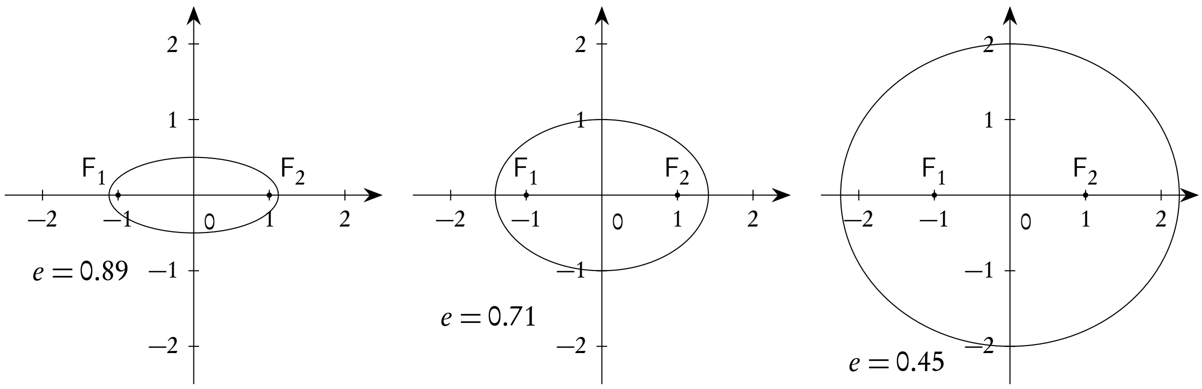


Figura 8.45.: *Ellissi con stessi fuochi e diverse eccentricità*

8.10. L'equazione di secondo grado in due incognite senza termine misto

Consideriamo un'equazione di secondo grado in due incognite, priva del termine misto, ovvero del tipo

$$Ax^2+Cy^2+Dx+Ey+F=0, \tag{8.78}$$

Per decidere di che tipo di curva si tratta e rappresentarla graficamente, distinguiamo i vari casi che si possono presentare, tenendo conto che A e C non possono essere contemporaneamente nulli.

1. $A=0 \wedge D=0$
   L'equazione si riduce alla

$$Cy^2+Ey+F=0, \tag{8.79}$$

cioè ad una equazione di secondo grado nella sola incognita y: se essa ha due soluzioni, diciamole y_1 e y_2, rappresenta due rette parallele all'asse delle x, $y=y_1$ e $y=y_2$; se ha una sola soluzione rappresenta una sola retta parallela all'asse delle x, oppure, come si usa dire, una coppia di rette coincidenti, $y=y_1$; se non ha soluzioni rappresenta l'insieme vuoto. In ogni caso si tratta di una conica degenere.

2. $A=0 \wedge D\neq 0$
   L'equazione si può porre nella forma

$$x=-\frac{C}{D}y^2-\frac{E}{D}y-\frac{F}{D}, \tag{8.80}$$

e rappresenta una parabola con asse parallelo all'asse delle x.

3. $C=0 \wedge E=0$
   L'equazione si riduce alla forma

$$Ax^2+Dx+F=0, \tag{8.81}$$

cioè ad una equazione di secondo grado nella sola incognita x: se essa ha due soluzioni, diciamole x_1 e x_2, rappresenta due rette parallele all'asse delle y, $x=x_1$ e $x=x_2$; se ha una sola soluzione rappresenta una sola retta parallela all'asse delle y, oppure, come si usa dire, una coppia di rette coincidenti, $x=x_1$; se non ha soluzioni rappresenta l'insieme vuoto. In ogni caso si tratta di una conica degenere.

4. $C=0 \wedge E \neq 0$
L'equazione si può porre nella forma

$$y=-\frac{A}{E}x^2-\frac{D}{E}x-\frac{F}{E}, \tag{8.82}$$

e rappresenta una parabola con asse parallelo all'asse delle y.

5. $A \neq 0 \wedge C \neq 0$
Si procede con la tecnica del completamento dei quadrati.

$$\begin{aligned} Ax^2+Cy^2+Dx+Ey+F &= A\left(x^2+\frac{D}{A}x\right)+C\left(y^2+\frac{E}{C}y\right)+F= \\ &= A\left(x^2+\frac{D}{A}x+\frac{D^2}{4A^2}\right)+C\left(y^2+\frac{E}{C}y+\frac{E^2}{4C^2}\right)+F-\frac{D^2}{4A^2}-\frac{E^2}{4C^2}= \\ &= A\left(x+\frac{D}{2A}\right)^2+C\left(y+\frac{E}{2C}\right)^2+F-\frac{D^2}{4A^2}-\frac{E^2}{4C^2}. \end{aligned}$$

Posto

$$x_C=-\frac{D}{2A}, \quad y_C=-\frac{E}{2C}, \quad H=-F+\frac{D^2}{4A^2}+\frac{E^2}{4C^2}, \tag{8.83}$$

la (8.78) si può scrivere nella forma

$$A(x-x_C)^2+C(y-y_C)^2=H. \tag{8.84}$$

Occorrerà distinguere due sottocasi.

5a. $H=0$
Se A e C sono concordi l'equazione avrà come unica soluzione il punto (x_C, y_C), ovvero una conica degenere. Se A e C sono discordi, portando una delle due quantità a secondo membro ed estraendo le radici quadrate si otterrà

$$\sqrt{|A|}(x-x_C)=\pm\sqrt{|C|}(y-y_C), \tag{8.85}$$

cioè una coppia di rette incidenti nel punto (x_C, y_C): ancora una conica degenere.

5b. $H \neq 0$
Si possono dividere ambo i membri per H, portando anche A e C al denominatore. Si ottiene

$$\frac{(x-x_C)^2}{H/A}+\frac{(y-y_C)^2}{H/C}=1.$$

Posto

$$\sqrt{|H/A|}=a^2, \quad \sqrt{|H/C|}=b^2,$$

si ottiene infine

$$\pm\frac{(x-x_C)^2}{a^2}\pm\frac{(y-y_C)^2}{b^2}=1. \tag{8.86}$$

L'equazione (8.78) rappresenta allora in questo caso un'ellisse, un'iperbole oppure l'insieme vuoto come discusso nel paragrafo 8.9.3.

Esempio 8.35. Tracciare il grafico della seguente conica:

$$x^2+2y^2-2x-8y+7=0.$$

Si ha

$$\begin{aligned} x^2+2y^2-2x-8y+7 &= (x^2-2x)+2(y^2-4y)+7= \\ &= (x^2-2x+1)+2(y^2-4y+4)+7-1-8= \\ &= (x-1)^2+2(y-2)^2-2. \end{aligned}$$

L'equazione data si può porre nella forma

$$\frac{(x-1)^2}{2}+(y-2)^2=1.$$

Se ne deduce che si tratta di un'ellisse con centro in $(1,2)$ e semiassi $\sqrt{2}$ e 1: dunque l'asse focale è orizzontale. Si ha

$$c=\sqrt{2-1}=1,$$

per cui i fuochi sono $\mathsf{F}_1(0,2)$ e $\mathsf{F}_2(2,2)$. Si veda la figura 8.46.

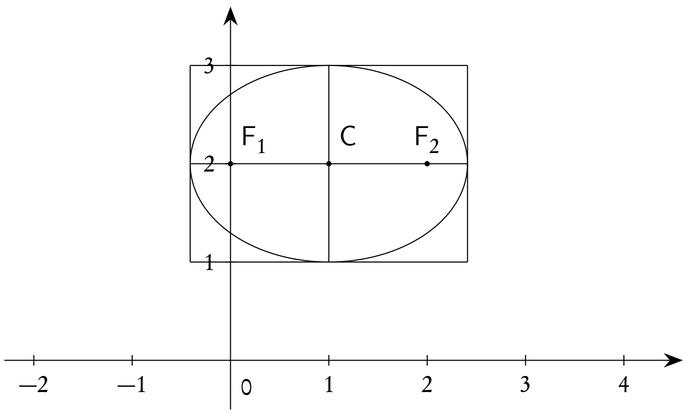


Figura 8.46.: *L'ellisse* $x^2+2y^2-2x-8y+7=0$

Esempio 8.36. Tracciare il grafico della seguente conica:

$$4x^2-y^2-16x+2y+15=0.$$

Si ha

$$\begin{aligned} 4x^2-y^2-16x+2y+15 &= 4(x^2-4x)-(y^2-2y)+15= \\ &= 4(x^2-4x+4)-(y^2-2y+1)+15-16+1= \\ &= 4(x-2)^2-(y-1)^2. \end{aligned}$$

L'equazione si può porre nella forma

$$(x-2)^2 = \frac{(y-1)^2}{4}$$

e rappresenta la coppia di rette

$$x-2 = \pm\frac{y-1}{2}$$

che si intersecano nel punto $\mathsf{C}(2,1)$, centro della conica degenere.

Esempio 8.37. Tracciare il grafico della seguente conica

$$x^2-2y^2-2x+4y-3=0.$$

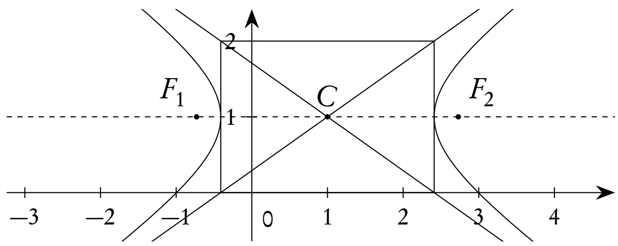


Figura 8.47.: *L'iperbole* $x^2-2y^2-2x+4y-3=0$

Si ha

$$\begin{aligned} x^2-2y^2-2x+4y-3 &= (x^2-2x)-2(y^2-2y)-3= \\ &= (x^2-2x+1)-2(y^2-2y+1)-3-1+2= \\ &= (x-1)^2-2(y-1)^2-2. \end{aligned}$$

L'equazione si può porre nella forma

$$\frac{(x-1)^2}{2}-(y-1)^2=1$$

e rappresenta un'iperbole con asse focale orizzontale, semiassi $\sqrt{2}$ e 1 e centro nel punto $\mathsf{C}(1,1)$. Si ha poi

$$c=\sqrt{2+1}=\sqrt{3},$$

per cui i fuochi sono $\mathsf{F}_1(1-\sqrt{3},1)$ e $\mathsf{F}_2(1-\sqrt{3},1)$. Le equazioni degli asintoti sono

$$\frac{x-1}{\sqrt{2}}=\pm(y-1).$$

8.11. L'iperbole equilatera

Un'iperbole si dice *equilatera*[(12)] se $a = b$. Come già accennato, vedi la pagina 263, un'iperbole equilatera ha eccentricità $e = \sqrt{2}$. In questo caso l'equazione si riduce alla forma

$$(x - x_C)^2 - (y - y_C)^2 = \pm a^2. \tag{8.87}$$

Se poi l'iperbole ha centro nell'origine l'equazione diventa

$$x^2 - y^2 = \pm a^2. \tag{8.88}$$

Essendo gli asintoti tra di loro perpendicolari e, nel caso di centro nell'origine, coincidenti con le bisettrici dei quadranti, è possibile effettuare una rotazione degli assi di $\pm\pi/4$ fino a farli coincidere con gli asintoti stessi. Effettuiamo, per esempio, una rotazione di $\pi/4$. Le equazioni (8.8) della rotazione si scrivono, in questo caso,

$$\begin{cases} x = x'\dfrac{\sqrt{2}}{2} - y'\dfrac{\sqrt{2}}{2} \\ x = x'\dfrac{\sqrt{2}}{2} + y'\dfrac{\sqrt{2}}{2} \end{cases} \tag{8.89}$$

e sostituendo questi valori nella (8.88) si ottiene, dopo facili calcoli,

$$x'y' = \pm\frac{a^2}{2}. \tag{8.90}$$

Possiamo dunque concludere, usando di nuovo x e y per gli assi, che un'iperbole equilatera riferita ad un sistema cartesiano Oxy in cui gli assi coincidano con gli asintoti, o, come si usa dire, un'*iperbole equilatera riferita agli asintoti* ha un'equazione del tipo

$$xy = k, \quad k \neq 0, \quad \text{ove} \quad |k| = \frac{a^2}{2} = \frac{b^2}{2}. \tag{8.91}$$

I due rami dell'iperbole si situano nel primo e terzo quadrante se $k > 0$, nel secondo e quarto se $k < 0$. I vertici dell'iperbole stanno sull'intersezione dell'iperbole stessa con una delle due bisettrici dei quadranti e hanno ascissa e ordinata uguali oppure opposte, con modulo $\sqrt{|k|}$ e quindi con i seguenti valori

$$\left(\sqrt{k}, \sqrt{k}\right) \quad \text{e} \quad \left(-\sqrt{k}, -\sqrt{k}\right), \quad \text{oppure} \quad \left(-\sqrt{-k}, \sqrt{-k}\right) \quad \text{e} \quad \left(\sqrt{-k}, -\sqrt{-k}\right). \tag{8.92}$$

Per quanto riguarda i fuochi, essendo

$$c = \sqrt{a^2 + b^2} = \sqrt{2a^2} = a\sqrt{2}$$

le loro coordinate si otterranno da quelle dei vertici moltiplicandole per $\sqrt{2}$.

[12] L'analogo concetto per l'ellisse porta a considerare la circonferenza.

Essendo, nella (8.91), $x \neq 0$, la (8.91) stessa si potrà scrivere nella forma

$$y = \frac{k}{x}. \tag{8.93}$$

La (8.93) mostra che l'iperbole è, in questo caso, il grafico di una legge di proporzionalità inversa, situazione che abbiamo già trattato nella pagina 123. Anche per questo motivo la situazione in esame è di grande importanza applicativa.

8.11.1. La funzione omografica

Consideriamo ora un'iperbole equilatera con centro fuori dall'origine. Essa avrà un'equazione del tipo

$$(x - x_C)^2 - (y - y_C)^2 = \pm a^2, \tag{8.94}$$

ovvero

$$x^2 - y^2 - 2x_C x + 2y_C y x + x_C^2 - y_C^2 = \pm a^2. \tag{8.95}$$

Sottoponiamo gli assi ad una rotazione di $\pi/4$ e sostituiamo le (8.89) nella (8.95). Otteniamo, dopo semplificazione,

$$x'y' + x'\frac{\sqrt{2}}{2}(x_C - y_C) - y'\frac{\sqrt{2}}{2}(x_C + y_C) = -x_C^2 + y_C^2 \pm a^2. \tag{8.96}$$

Tenendo conto che, in base alle (8.9),

$$x'_C = x_C\frac{\sqrt{2}}{2} + y_C\frac{\sqrt{2}}{2}, \quad y'_C = -x_C\frac{\sqrt{2}}{2} + y_C\frac{\sqrt{2}}{2},$$

la (8.96) si può riscrivere nella forma

$$x'y' - y'_C x' - x'_C y' = h. \tag{8.97}$$

Possiamo dunque concludere, usando di nuovo le lettere x e y per gli assi, che un'iperbole equilatera con centro nel punto (x_C, y_C) e assi paralleli agli asintoti ha equazione

$$xy - y_C x - x_C y = h. \tag{8.98}$$

Esempio 8.38. Consideriamo l'iperbole equilatera di equazione

$$(x-1)^2 - (y-2)^2 = 1,$$

avente centro in $C(1,2)$ e semiassi di lunghezza 1. Se sottoponiamo gli assi ad una rotazione di $\pi/4$, applicando la tecnica sopra descritta otteniamo l'equazione

$$x'y' - \frac{\sqrt{2}}{2}x' - \frac{3\sqrt{2}}{2}y' = -2.$$

Il centro C ha ora le nuove coordinate

$$C\left(\frac{3\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

Si veda la figura 8.48.

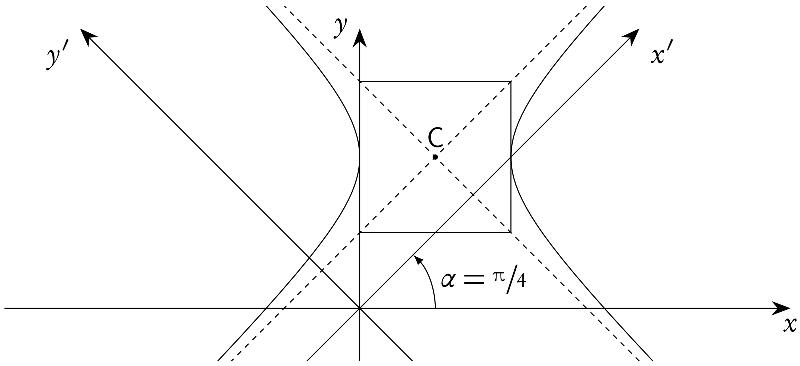


Figura 8.48.: *L'iperbole di equazioni* $x^2 - y^2 - 2x + 4y - 4 = 0$

Tenendo conto che, nella (8.98), $x \neq x_C$ la (8.98) stessa può essere posta nella forma

$$y = \frac{y_C x + b}{x - x_C},$$

che si usa scrivere come

$$y = \frac{ax + b}{cx + d}. \tag{8.99}$$

La (8.99) mostra che l'iperbole equilatera può essere pensata, in questo caso, come il grafico di una funzione reale di variabile reale

$$f(x) = \frac{ax + b}{cx + d} \tag{8.100}$$

di dominio naturale $x \neq -d/c$, detta *funzione omografica*.

Tuttavia occorre tenere presente che *non* ogni funzione del tipo (8.100) ha come grafico una iperbole equilatera. Infatti

1. se $c = 0$ la (8.100) si riduce alla forma $f(x) = a/d\, x + b/d$ che ha come grafico una retta;
2. se $ad - bc = 0$ si ha

   $$b = \frac{ad}{c},$$

   e quindi

   $$f(x) = \frac{ax + b}{cx + d} = \frac{ax + \frac{ad}{c}}{cx + d} = \frac{a}{c}\,\frac{cx + d}{cx + d} = \frac{a}{c} \quad \text{se} \quad cx + d \neq 0,$$

   che rappresenta una retta parallela all'asse delle ascisse, con l'esclusione del punto di ascissa $-d/c$.

8.12. L'equazione di secondo grado in due incognite con termine misto

Abbiamo già incontrato, trattando l'iperbole equilatera, equazioni di secondo grado in due incognite in cui è presente il termine misto xy, Vogliamo ora presentare una strategia generale per trattarle. Consideriamo dunque una generica equazione di secondo grado in due incognite, in cui sia presente il termine misto

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0, \quad B \neq 0, \tag{8.101}$$

e cerchiamo di individuare, se possibile, una rotazione degli assi in modo che, nel nuovo sistema, questo termine sia assente: a questo punto potremo usare la tecnica già proposta nel paragrafo 8.10.

Una generica rotazione degli assi ha un'equazione del tipo (8.8). Sostituendo le espressioni di x e y nella (8.101), eseguendo tutti i calcoli e le semplificazioni, si ottiene una equazione in cui il termine misto $x'y'$ ha coefficiente

$$-2A\sin\alpha\cos\alpha + B\cos^2\alpha - B\sin^2\alpha + 2C\sin\alpha\cos\alpha = (-A + C)\sin 2\alpha + B\cos 2\alpha. \tag{8.102}$$

Se vogliamo che l'equazione sia priva del termine misto, questo coefficiente dovrà essere zero: l'angolo α di rotazione dovrà essere tale che

$$\cot 2\alpha = \frac{A - C}{B} \tag{8.103}$$

e sarà sufficiente scegliere, tra le infinite soluzioni della (8.103), un angolo acuto.

Esempio 8.39. Rappresentare graficamente la conica di equazione

$$x^2 + xy + y^2 - x - 1 = 0.$$

Applicando la (8.103) si trova $\cot 2\alpha = 0$, da cui $\alpha = \pi/4$. Le formule della rotazione sono allora

$$\begin{cases} x = x'\dfrac{\sqrt{2}}{2} - y'\dfrac{\sqrt{2}}{2} \\ x = x'\dfrac{\sqrt{2}}{2} + y'\dfrac{\sqrt{2}}{2} \end{cases}.$$

Sostituendo questi valori nella equazione data si ottiene, dopo semplificazione,

$$3(x')^2 + (y')^2 - x'\sqrt{2} + y'\sqrt{2} - 2 = 0.$$

Quest'ultima equazione, con la tecnica del completamento dei quadrati, può essere riscritta come

$$\frac{\left(x' - \dfrac{\sqrt{2}}{6}\right)^2}{\dfrac{8}{9}} + \frac{\left(y' + \dfrac{\sqrt{2}}{2}\right)^2}{\dfrac{8}{3}} = 1$$

Il grafico, nel sistema ruotato, è dunque una ellisse di centro

$$C\left(\frac{\sqrt{2}}{6}, -\frac{\sqrt{2}}{2}\right)$$

e semiassi

$$a = \sqrt{\frac{8}{9}} \quad \text{e} \quad b = \sqrt{\frac{8}{3}}.$$

Si veda la figura 8.49.

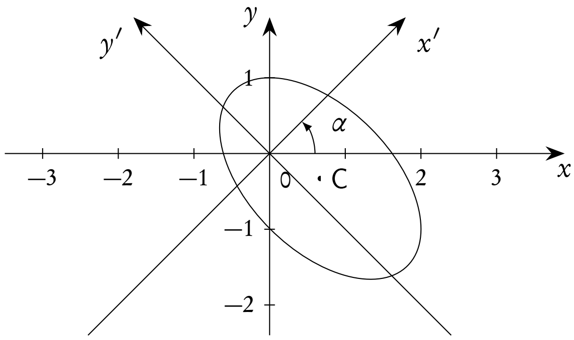


Figura 8.49.: *L'ellisse di equazione* $x^2 + xy + y^2 - x - 1 = 0$

8.13. Altri luoghi geometrici

Utilizzando la geometria analitica è facile scoprire le caratteristiche di molti luoghi geometrici notevoli, la cui descrizione sintetica è spesso più complessa. Ne proponiamo due, a titolo d'esempio e di utile esercizio.

La circonferenza di Apollonio

Dati due punti A e B del piano, il luogo dei punti P dello stesso piano tali che

$$\frac{|\overline{AP}|}{|\overline{PB}|} = k, \quad k > 0 \land k \neq 1,$$

è una circonferenza detta *circonferenza di Apollonio*. Il caso $k = 1$ va escluso perché banale: si ottiene semplicemente l'asse del segmento $\overline{AB}$.

Per trattare analiticamente un problema come questo è conveniente introdurre un sistema di coordinate opportuno, in modo da semplificare al massimo i calcoli. Scegliamo un sistema Oxy in cui i punti dati abbiano coordinate $A(0,0)$ e $B(a,0)$, con $a > 0$. Detto $P(x,y)$ un punto del piano, esso appartiene al luogo se e solo se

$$\sqrt{x^2 + y^2} = k\sqrt{(x-a)^2 + y^2},$$

che, quadrando, semplificando e dividendo per $k^2 - 1$ (cosa possibile perché $k > 0$ e $k \neq 1$) fornisce

$$x^2 + y^2 - \frac{2ak^2}{k^2 - 1} + \frac{k^2}{k^2 - 1}.$$

L'equazione ottenuta mostra che si tratta di una circonferenza di centro e raggio dati da

$$\mathrm{C}\left(\frac{ak^2}{k^2-1},0\right), \qquad r=a\left|\frac{k}{k^2-1}\right|.$$

Poiché

$$\frac{k^2}{k^2-1}<0 \vee \frac{k^2}{k^2-1}>1,$$

il centro si trova sempre all'esterno del segmento $\overline{\mathrm{AB}}$ e precisamente dalla parte di A se $0<k<1$, dalla parte di B se $k>1$. Se $0<k<1$ la distanza tra il centro e il punto A vale

$$\frac{ak^2}{1-k^2},$$

che è minore del raggio: il cerchio contiene il punto A; analogamente se $k>1$ il cerchio contiene il punto B. Del resto queste proprietà sono evidenti perché un punto del segmento $\overline{\mathrm{AB}}$ deve appartenere al luogo.

La figura 8.50 illustra la situazione con $a=2$ e $k=2$.

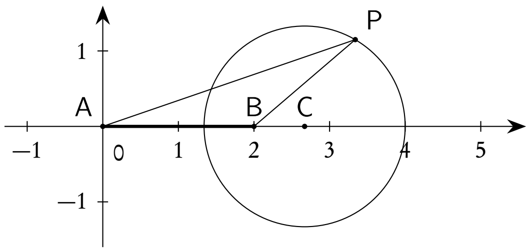


Figura 8.50.: *Ciconferenza di Apollonio*

Ancora ellisse ed iperbole

Dati, in un piano, una retta d e un punto $\mathrm{F}\notin \mathrm{d}$, il luogo geometrico dei punti del piano tali che

$$\frac{|\overline{\mathrm{PF}}|}{d(\mathrm{F},\mathrm{d})}=k, \quad k>o \wedge k\neq 1,$$

è un'ellisse se $0<k<1$, un'iperbole se $k>1$. Il caso $k=1$ è banale perché, come già sappiamo, si tratta di una parabola di fuoco F e direttrice d. I nomi di fuoco e direttrice si mantengono anche per il caso in esame.

Scegliamo un sistema di coordinate in cui d coincide con l'asse y e il punto F ha coordinate $(p,0)$, con $p>0$. Detto $\mathrm{P}(x,y)$ un punto del piano, esso appartiene al luogo se e solo se

$$\sqrt{(x-p)^2+y^2}=k|x|,$$

che, quadrando e semplificando, fornisce

$$(1-k^2)x^2+y^2-2px+p^2=0.$$

L'equazione ottenuta mostra che si tratta di una conica. Procedendo con la tecnica del completamento dei quadrati si ottiene

$$\frac{\left(x-\dfrac{p}{1-k^2}\right)^2}{\dfrac{p^2k^2}{(1-k^2)^2}}+\frac{y^2}{\dfrac{p^2k^2}{1-k^2}}=1.$$

Si tratta dunque di una conica con asse focale l'asse x, centro, semiassi e semidistanza focale dati da

$$\mathrm{C}\left(\frac{p}{1-k^2},0\right),\quad a=\frac{pk}{|1-k^2|},\quad b=\frac{pk}{\sqrt{|1-k^2|}},\quad c=\frac{pk^2}{|1-k^2|}.$$

Precisamente

1. se $0<k<1$ si tratta di un'ellisse;
2. se $k>1$ si tratta di un'iperbole.

In entrambi i casi se si considerano il simmetrico F' del punto F rispetto al centro e la retta d' simmetrica di d rispetto al centro, si ha ancora

$$\frac{|\overline{\mathrm{PF}'}|}{d(\mathrm{F}',\mathrm{d}')}=k,\quad k>o\wedge k\neq 1,$$

ovvero F' e d' sono anch'essi un fuoco e una direttrice. Si può poi verificare che valgono le già note proprietà focali di ellisse e iperbole:

$$|\overline{\mathrm{PF}'}|+|\overline{\mathrm{PF}}|=2a\quad\text{oppure}\quad\left||\overline{\mathrm{PF}'}|-|\overline{\mathrm{PF}}|\right|=2a,$$

cioè che la definizione ora data è perfettamente equivalente alla definizione con le proprietà focali.

È poi immediato che $k={}^c/_a$, cioè che il valore del rapporto

$$\frac{|\overline{\mathrm{PF}}|}{d(\mathrm{F},\mathrm{d})}$$

è esattamente l'eccentricità dell'ellisse o dell'iperbole.

Si noti che questa definizione di ellisse e iperbole mediante fuoco e direttrice "uniforma", in un certo senso, le definizioni delle tre sezioni coniche, che solo con le proprietà focali sembravano diverse: ellisse, parabola e iperbole sono i luoghi geometrici dei punti del piano tali che il rapporto delle distanze da un punto fisso detto fuoco e da una retta fissa detta direttrice (con il fuoco esterno alla direttrice) valga k. Se $0<k<1$ si ha un'ellisse, se $k=1$ si ha una parabola, se $k>1$ si ha un'iperbole.

Come mostra l'equazione

$$(1-k^2)x^2+y^2-2px+p^2=0$$

sopra ottenuta, questo luogo può essere una circonferenza solo se $k=0$. Il valore $k=0$ non si può però mai ottenere direttamente dal rapporto $|\overline{\mathrm{PF}}|/d(\mathrm{F},\mathrm{d})$: si può pensare comunque che esso possa essere ottenuto allontanando la direttrice all'infinito. In questo caso il centro C si sovrappone a F e quindi anche F' si sovrappone a F. La circonferenza è dunque, secondo questa ottica, una conica con "eccentricità limite" nulla, con i due fuochi e il centro coincidenti e con direttrici "all'infinito".

La figura 8.51 illustra la situazione con $p=3$ e $k=2$ (iperbole), la 8.52 con $p=3$ e $k={}^1/_2$ (ellisse).

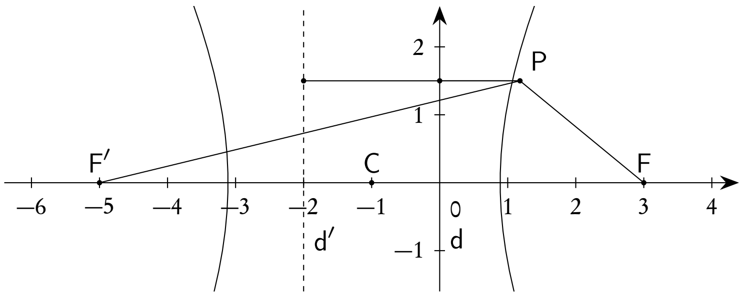


Figura 8.51.: *Iperbole con fuochi e direttrici*

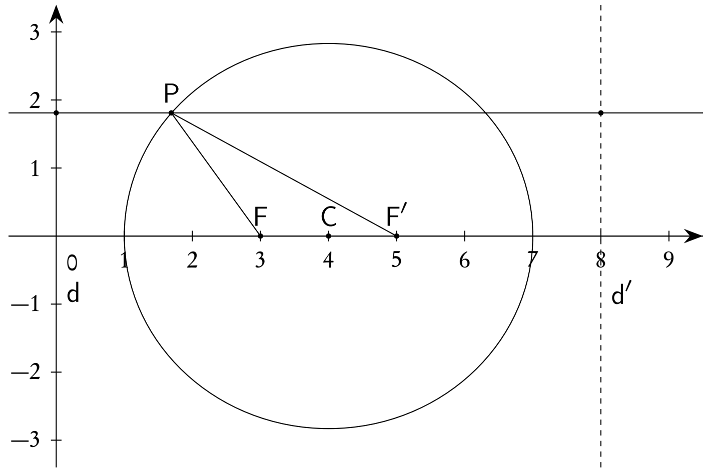


Figura 8.52.: *Ellisse con fuochi e direttrici*

8.14. Altri tipi di equazioni

Le curve algebriche aventi equazioni di grado superiore al secondo richiedono tecniche che sono al di fuori degli scopi di questo testo; tuttavia alcune situazioni possono essere trattate anche a livello elementare e ne faremo qualche cenno.

La situazione più semplice è quella in cui l'equazione $f(x,y) = 0$ della curva può essere riscritta fattorizzando il primo membro in un prodotto di fattori con grado minore o uguale a 2: basterà allora applicare la legge di annullamento del prodotto. Vediamo un esempio.

Esempio 8.40. Rappresentare graficamente la curva di equazione $(x-y^2)(x^2+y^2-1)=0$.

Si tratta dell'unione dei grafici di $x-y^2=0$ e $x^2+y^2-1=0$, che rappresentano, rispettivamente, una parabola e una circonferenza. Si veda la figura 8.53.

Negli esempi che seguono proponiamo qualche altra tecnica per trattare curve di grado superiore al secondo o curve non razionali.

Esempio 8.41. Tracciare il grafico della curva di equazione cartesiana $x^3-y^2=0$.

Riscrivendo l'equazione nella forma $y^2=x^3$ si vede facilmente che deve essere $x \geq 0$. Dopodiché si ha

$$y = \pm\sqrt{x^3} = \pm x^{3/2}:$$

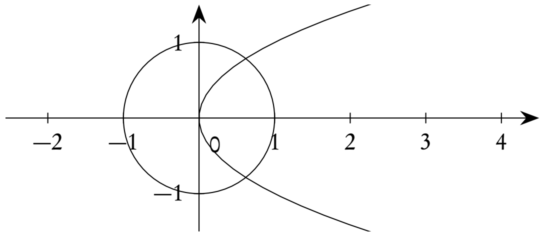


Figura 8.53.: *La curva di equazione* $(x-y^2)(x^2+y^2-1)=0$

la curva sarà semplicemente l'unione di due funzioni potenza. Si veda la figura 8.54.

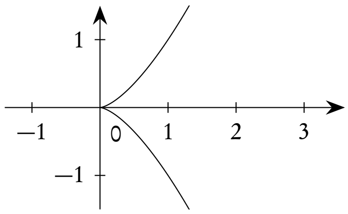


Figura 8.54.: *La curva di equazione* $x^3-y^2=0$

Esempio 8.42. Tracciare il grafico della curva di equazione $y=\sqrt{x^2-2x}$.

Intanto si deve avere $x<0 \vee x>2$, dopodiché l'equazione data è equivalente al sistema

$$\begin{cases} y \geq 0 \\ y^2 = x^2-2x \end{cases}.$$

L'equazione ha come grafico un'iperbole di centro $(1,0)$ e semiassi $a=b=1$. La disequazione $y \geq 0$ implica che si deve considerare solo la parte soprastante l'asse delle x. Si veda la figura 8.55.

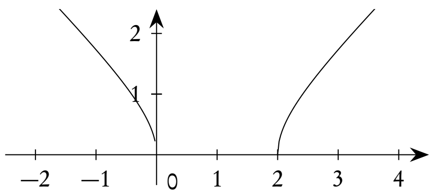


Figura 8.55.: *La curva di equazione* $y=\sqrt{x^2-2x}$

Esempio 8.43. Tracciare il grafico della curva di equazione $y=\sqrt{1-x^2}+1$.

Intanto si deve avere $-1 \leq x \leq 1$. Dopodiché, scrivendo l'equazione nella forma

$$y-1=\sqrt{1-x^2},$$

si vede che l'equazione è equivalente al sistema

$$\begin{cases} y \geq 1 \\ (y-1)^2 = 1-x^2 \end{cases}.$$

L'equazione rappresenta una circonferenza di centro $(0,1)$ e raggio 1; la disequazione implica che si deve prendere solo la semicirconferenza soprastante la retta $y=1$. Si veda la figura 8.56.

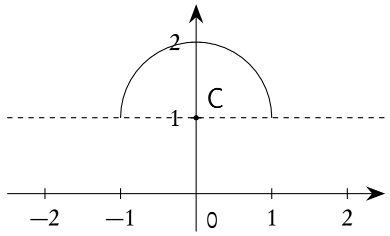


Figura 8.56.: *La curva di equazione* $y=\sqrt{1-x^2}+1$

8.15. Equazioni parametriche

Finora abbiamo considerato rette e curve nel piano mediante una equazione in due incognite, chiamandola rappresentazione cartesiana della retta o curva. Esistono altri metodi per rappresentare gli stessi luoghi e faremo un breve cenno alla cosiddetta *rappresentazione parametrica*, segnalando comunque che si tratta solo di qualche semplice considerazione introduttiva: la trattazione dettagliata sarà compito dei futuri corsi universitari.

Se f e g sono due funzioni reali di una variabile reale, definite in un comune intervallo I, indichiamo con t la variabile indipendente e con x e y rispettivamente le variabili dipendenti. Al variare di t in I, il punto di coordinate $(x,y)=\big(f(t),g(t)\big)$ descrive un insieme del piano: sotto opportune condizioni questo insieme è una curva nel senso intuitivo del termine. Scriveremo

$$\begin{cases} x=f(t) \\ y=g(t) \end{cases}, \tag{8.104}$$

e chiameremo le (8.104) *equazioni parametriche della curva*.

Esempio 8.44. Tracciare, per punti, la curva di equazioni parametriche

$$\begin{cases} x=t^3-2t^2 \\ y=t^3-t \end{cases}.$$

Si può costruire la seguente tabella di valori.

t	-2	$-\frac{3}{2}$	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2
x	-16	$-\frac{63}{8}$	-3	$-\frac{5}{8}$	0	$-\frac{3}{8}$	-1	$-\frac{9}{8}$	0
y	-6	$-\frac{15}{8}$	0	$\frac{3}{8}$	0	$-\frac{3}{8}$	0	$\frac{15}{8}$	6

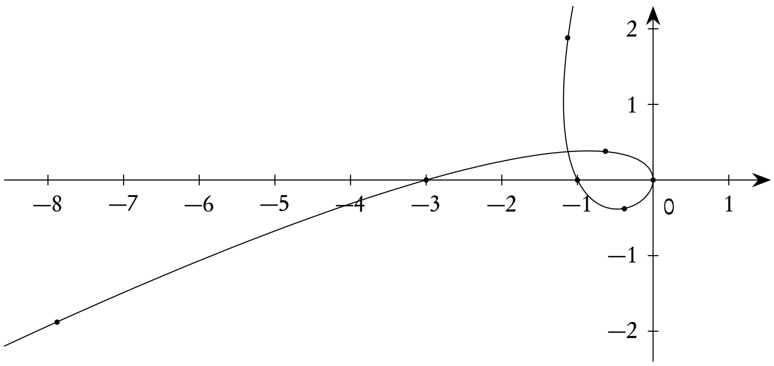


Figura 8.57.: *La curva* (t^3-2t^2, t^3-t)

Naturalmente per un grafico accettabile occorreranno molti più punti: la figura 8.57 mostra la curva ottenuta con un software di calcolo, con evidenziati alcuni dei punti della tabella precedente.

Se dalle (8.104) è possibile eliminare il parametro t (per esempio ricavandolo in una delle due e sostituendone il valore nell'altra), si ottiene un'equazione nelle incognite x ed y, che è l'equazione cartesiana della curva.

Esempio 8.45. Trovare l'equazione cartesiana della curva di equazioni parametriche

$$\begin{cases} x = t-1 \\ y = t^2 \end{cases} .$$

Si ha, dalla prima equazione, $t = x+1$ che sostituito nella seconda porge $y = x^2+2x+1$, che è l'equazione di una parabola.

Esempio 8.46. Trovare l'equazione cartesiana della curva di equazioni parametriche

$$\begin{cases} x = 1+2\cos t \\ y = \sin t \end{cases}, \quad t \in [0, 2\pi].$$

Riscrivendo la prima equazione nella forma $(x-1)/2 = \cos t$, quadrando e sommando con il quadrato della seconda si ottiene l'equazione

$$\frac{(x-1)^2}{4} + y^2 = 1,$$

che è l'equazione di un'ellisse con centro in $(1,0)$ e semiassi di lunghezza 2 e 1.

L'importanza delle equazioni parametriche è anche legata al fatto che equazioni del tipo (8.104) possono essere pensate come le leggi orarie del moto di un punto nel piano: il parametro t ha esattamente il significato di tempo e la curva grafico è la traiettoria descritta dal punto al variare del tempo.

Anche se non possiamo entrare nei dettagli segnaliamo che è molto importante il fatto che una curva può avere diverse rappresentazioni parametriche; se siamo interessati solo alla curva come insieme di punti, le rappresentazioni si equivalgono, ma se siamo interessati per esempio alle leggi orarie di un moto esse possono riferirsi a situazioni radicalmente differenti, come mostra l'esempio che segue.

Esempio 8.47. Tracciare il grafico delle curve di equazioni parametriche

$$\gamma_1\colon \begin{cases} x = t \\ y = 1 \end{cases}, \qquad \gamma_2\colon \begin{cases} x = t^3 \\ y = 1 \end{cases}, \qquad \gamma_3\colon \begin{cases} x = t \sin t \\ y = 1 \end{cases}.$$

In tutti e tre i casi la "curva" è costituita dalla retta $y = 1$ e, al variare di t, in ciascuno dei tre casi si ottengono tutti i punti della retta. Dal punto di vista fisico però γ_1 rappresenta un moto rettilineo ed uniforme, γ_2 un moto accelerato, e con accelerazione variabile, γ_3 un moto oscillante, con oscillazioni di ampiezza sempre più grande. Anche se la dimostrazione accurata di questi tutti fatti richiede l'uso delle derivate, lo si può capire a livello elementare calcolando i valori di x ad intervalli di tempo costanti (la y rimane sempre 1). Si veda la tabella che segue.

t	0	$\pi/2$	π	$3\pi/2$	2π	$5\pi/2$	3π	$7\pi/2$	4π
$x(\gamma_1)$	0	1.57	3.14	4.71	6.28	7.85	9.42	11	12.57
$x(\gamma_2)$	0	3.88	31.01	104.65	248.05	484.47	837.17	1329.39	1984.4
$x(\gamma_3)$	0	1.57	0	−4.71	0	7.85	0	−11	0

8.15.1. Equazioni parametriche della retta

Sia data una retta di equazione cartesiana $ax + bx + c = 0$. Se $a \neq 0$, posto $y = t$ si ottiene

$$\begin{cases} x = -\dfrac{b}{a}\,t - \dfrac{c}{a} \\ y = t \end{cases}; \tag{8.105}$$

se $b \neq 0$, posto $x = t$ si ottiene

$$\begin{cases} x = t \\ y = -\dfrac{a}{b}\,t - \dfrac{c}{c}b \end{cases}. \tag{8.106}$$

Le (8.105) e (8.105) sono del tipo

$$\begin{cases} x = x_0 + l\,t \\ y = y_0 + m\,t \end{cases}. \tag{8.107}$$

Viceversa, ricavando la t da una delle (8.107) e sostituendolo nell'altra, si ottiene un'equazione di primo grado in x e y, cioè l'equazione di una retta.

Le (8.107) si chiamano *equazioni parametriche canoniche della retta*. Il punto di coordinate (x_0, y_0) appartiene alla retta e si ottiene per il valore 0 di t: si può chiamare punto iniziale o punto in evidenza.

8.16. Cenno alle disequazioni in due incognite

Se per le disequazioni in un'incognita reale la tecnica grafica era un optional, per le disequazioni in due incognite diventa praticamente obbligatoria: non esistono metodi analitici semplici per scrivere le soluzioni delle disequazioni in due incognite, per cui di solito l'insieme di soluzioni viene graficamente rappresentato nel piano cartesiano. In questa breve introduzione tratteremo solo le disequazioni di primo e secondo grado e semplici sistemi.

8.16.1. Il principio generale

Consideriamo il luogo di punti del piano rappresentato da un'equazione del tipo $f(x,y)=0$: nei casi che a noi interesseranno questo luogo è una retta (se la funzione f a primo membro è un polinomio di primo grado in due incognite: $f(x,y)=ax+by+c$), oppure una conica (se la funzione è un polinomio di secondo grado in due incognite: $f(x,y)=ax^2+bxy+cy^2+dx+ey+f$).

Le disequazioni che considereremo saranno sempre del tipo $f(x,y)>0$ (oppure <0, oppure ≥ 0, oppure ≤ 0). Poiché la retta o la conica divide il piano in due (tre per il caso dell'iperbole, ma come vedremo potremo sempre ragionare come se fossero due) regioni disgiunte, la disequazione sarà sempre verificata in una delle due e non verificata nell'altra: basterà, come vedremo negli esempi, solo provare con un punto interno ad una delle regioni per scoprire qual è l'insieme delle soluzioni.

Per risolvere sistemi basterà trovare l'intersezione degli insiemi soluzione delle singole disequazioni, per disequazioni fratte o per prodotti, basterà usare la regola dei segni, esattamente come fatto nel caso di un'incognita.

8.16.2. Disequazioni di primo grado

Esempio 8.48. Risolvere la disequazione $2x+3y-6>0$. Rappresentata la retta $2x+3y-6=0$, se si prova a sostituire nella disequazione le coordinate dell'origine degli assi si ottiene $-6>0$ che è falsa: l'insieme delle soluzioni è allora l'altro semipiano individuato dalla retta, rappresentato in tratteggio nella figura 8.58, con l'esclusione della retta origine.

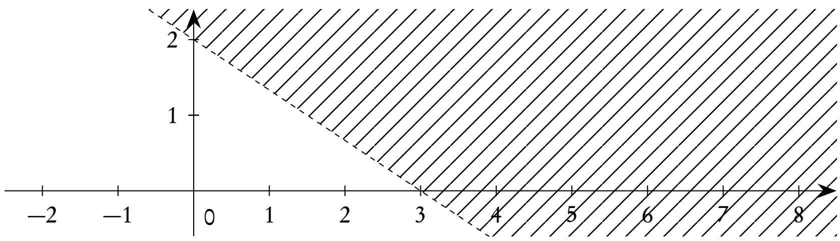


Figura 8.58.: *L'insieme* $2x+3y-6>0$

Esempio 8.49. Risolvere la disequazione $x-2y+1\leq 0$. Rappresentata la retta $x-2y+1=0$, se si prova a sostituire nella disequazione le coordinate dell'origine degli assi si trova $1\leq 0$, che è falsa: l'insieme di soluzioni è allora l'altro semipiano individuato dalla retta, rappresentato in tratteggio nella figura 8.59. Questa volta anche l'origine del semipiano, cioè la retta, fa parte dell'insieme delle soluzioni.

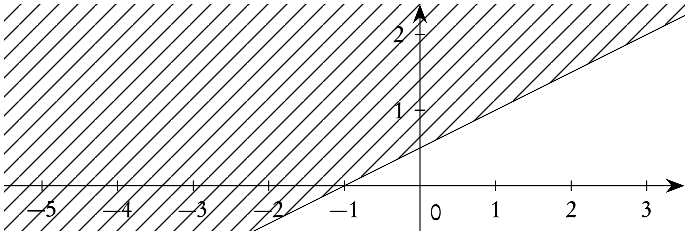


Figura 8.59.: *L'insieme* $x - 2y + 1 \leq 0$

8.16.3. Disequazioni di secondo grado

Esempio 8.50. Risolvere la disequazione $x^2 - 3y^2 + 4x - 6y - 4 < 0$. Rappresentata l'iperbole $x^2 - 3y^2 + 4x - 6y - 4 = 0$, se si sostituiscono nella disequazione le coordinate dell'origine si ottiene $-4 < 0$ che è vera: la disequazione è allora verificata in tutta la parte di piano individuata dall'iperbole e a cui appartiene l'origine (regione tratteggiata nella figura 8.60). L'iperbole in realtà divide il piano in tre regioni distinte, ma se si tiene conto di come è originata l'iperbole si possono considerare le due "interne" (quelle non tratteggiate nella figura) come un'unica regione che "si salda" all'infinito. L'iperbole stessa, bordo della regione tratteggiata, non è compresa nell'insieme.

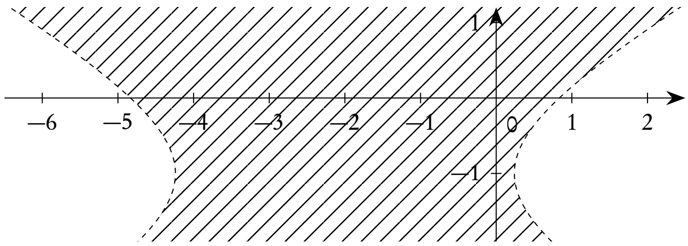


Figura 8.60.: *L'insieme* $x^2 - 3y^2 + 4x - 6y - 4 < 0$

Esempio 8.51. Risolvere la disequazione

$$8x^2 - 4xy + 5y^2 - 100 \leq 0.$$

Rappresentata l'ellisse $8x^2 - 4xy + 5y^2 - 100 = 0$, se si sostituiscono nella disequazione le coordinate dell'origine si trova $-100 \leq 0$ che è vera: la disequazione è allora verificata in tutta la parte interna all'ellisse, compresa l'ellisse stessa. Si veda la figura 8.61.

8.16.4. Sistemi di disequazioni e equazioni

Come per il caso di un'incognita si potranno considerare sistemi di disequazioni: basterà considerare l'intersezione tra gli insiemi soluzioni delle varie disequazioni. Il sistema, per esempio, tra le due disequazioni considerate negli esempi 8.50 e 8.51 ha come insieme di soluzioni la regione tratteggiata nella figura 8.62.

Sono molto importanti nelle applicazioni anche i sistemi tra equazioni e disequazioni in due incognite, che si risolveranno sempre intersecando l'insieme soluzione della equazione (una curva nei casi che ci interessano) e l'insieme soluzione della disequazione (un opportuno sottoinsieme del piano).

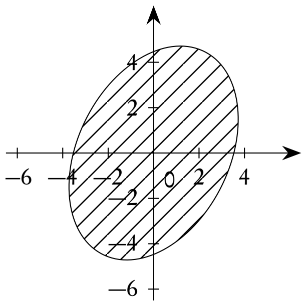


Figura 8.61.: *L'insieme* $8x^2 - 4xy + 5y^2 - 100 < 0$

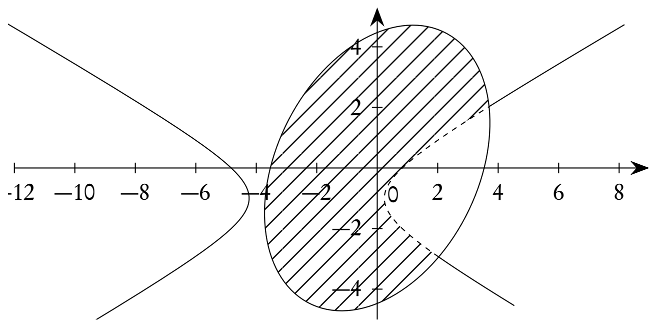


Figura 8.62.: *Esempio di sistema di disequazioni in due incognite*

Esempio 8.52. Risolvere il sistema

$$\begin{cases} x^2 + y^2 - 1 = 0 \\ 4x + 2y - 1 > 0 \end{cases}.$$

L'insieme soluzione è l'arco di circonferenza evidenziato nella figura 8.63. Esso è stato ottenuto intersecando la circonferenza (insieme delle soluzioni della equazione di secondo grado) e il semipiano tratteggiato (insieme delle soluzioni della disequazione di primo grado).

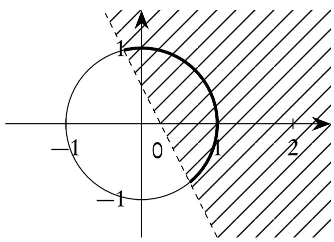


Figura 8.63.: *Esempio di sistema tra un'equazione e una disequazione in due incognite*

Se si dovesse risolvere una disequazione fratta, o contenente prodotti, basterà applicare la regola dei segni, come già accennato.

Esempio 8.53. Risolvere la disequazione

$$\frac{x^2+y^2-1}{4x+2y-1} \geq 0.$$

Nella figura 8.64 sono riportati, nell'ordine, gli insiemi di positività del numeratore (a sinistra) e del denominatore (al centro) e infine (a destra) l'insieme dove è verificata la disequazione: sono compresi i due archi di circonferenza, privati degli estremi, ed è esclusa la retta.

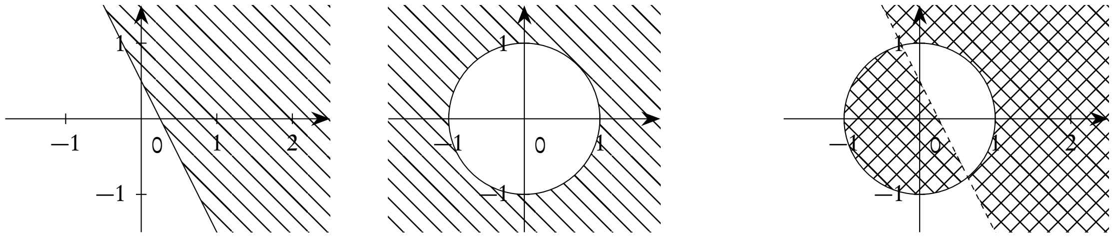


Figura 8.64.: *Esempio di disequazione fratta in due incognite*

8.17. Esercizi

Esercizio 8.1. *Un triangolo* ABC *ha il vertice* A *di coordinate* $(-2,3)$*; l'equazione dell'altezza* h *e della mediana* m *uscenti dal vertice* C *sono, rispettivamente,* $3x-2y-8=0$ *e* $4x-5y+1=0$*. Determinare le coordinate degli altri vertici del triangolo.*

Risoluzione. Il vertice C si trova chiaramente nell'intersezione tra la mediana e la bisettrice uscenti da C stesso.

$$\begin{cases} 3x-2y-8=0 \\ 4x-5y+1=0 \end{cases},$$

da cui si ottiene facilmente $C(6,5)$.

Cerchiamo ora l'equazione della retta AB: essa passa per A ed è perpendicolare ad h, dunque ha coefficiente angolare $-2/3$. Allora

$$\mathsf{AB}\colon y-3=-\frac{2}{3}(x+2) \quad \Rightarrow \quad 2x+3y-5=0.$$

Si poteva anche usare la (8.18) ottenendo lo stesso risultato in maniera forse anche più semplice.

Il punto medio M di $\overline{\mathsf{AB}}$ si trova sull'intersezione tra la retta AB e la retta m:

$$\begin{cases} 2x+3y-5=0 \\ 4x-5y+1=0 \end{cases},$$

da cui si ottiene $\mathsf{M}(1,1)$.

Per trovare il punto B possiamo usare le formule per il punto medio, oppure intersecare la circonferenza di centro M e raggio $|\overline{AM}|$ con la retta AB: scegliamo la prima opzione, più elementare e semplice dal punto di vista dei calcoli.

$$x_M = \frac{x_A + x_B}{2} \quad \Rightarrow \quad 1 = \frac{-2 + x_B}{2} \quad \Rightarrow \quad x_B = 4;$$
$$y_M = \frac{y_A + y_B}{2} \quad \Rightarrow \quad 1 = \frac{3 + y_B}{2} \quad \Rightarrow \quad y_B = -1$$
□

Esercizio 8.2. *Determinare l'equazione della circonferenza di centro* $C(-1, -2)$ *e tangente alla retta* $t\colon x + y - 2 = 0$.

Risoluzione. Si può trovare il punto T di tangenza intersecando la perpendicolare per C a t. La perpendicolare ha equazione, vedi la (8.18),

$$1(x+1) - 1(y+2) = 0 \quad \Rightarrow \quad x - y - 1 = 0.$$

Si trova facilmente $T(3/2, 1/2)$. Il raggio è dunque

$$r = |\overline{CT}| = \frac{5}{\sqrt{2}}.$$

Si poteva anche trovare direttamente il raggio come distanza tra C e t. La circonferenza ha equazione

$$(x+1)^2 + (y+2)^2 = \frac{25}{2} \quad \Rightarrow \quad 2x^2 + 2y^2 + 4x + 8y - 15 = 0.$$
□

Esercizio 8.3. *Sia data l'ellisse* $x^2 + 4y^2 = 4$ *e il suo punto* $P(1, \sqrt{3}/2)$. *Trovare le equazioni della tangente e della normale in* P *all'ellisse. Detti* F_1 *e* F_2 *i fuochi, provare poi che la normale é bisettrice dell'angolo* $F_1\widehat{P}F_2$, *ovvero che un raggio luminoso emesso da* F_1 *e passante per* P *si riflette sull'ellisse passando per* F_2.

Risoluzione. Per trovare la tangente possiamo usare la formula di sdoppiamento:

$$x + 2\sqrt{3}y - 4 = 0.$$

Per trovare la normale usiamo la formula (8.18):

$$2\sqrt{3}(x-1) - \left(y - \frac{\sqrt{3}}{2}\right) = 0 \quad \Rightarrow \quad 4\sqrt{3}x - 2y - 3\sqrt{3} = 0.$$

Essendo

$$c = \sqrt{|a^2 - b^2|} = \sqrt{4-1} = \sqrt{3},$$

i fuochi saranno

$$F_1 = (-\sqrt{3}, 0) \qquad F_2 = (\sqrt{3}, 0).$$

Le equazioni di F_1P e F_2P sono, rispettivamente,

$$\sqrt{3}x - 2(\sqrt{3}+1)y + 3 = 0 \quad \text{e} \quad \sqrt{3}x + 2(\sqrt{3}-1)y - 3 = 0.$$

Le equazioni delle bisettrici degli angoli individuati da queste due rette sono dunque

$$\frac{|\sqrt{3}x-2(\sqrt{3}+1)y+3|}{\sqrt{19+8\sqrt{3}}}=\frac{|\sqrt{3}x+2(\sqrt{3}-1)y-3|}{\sqrt{19-8\sqrt{3}}},$$

ovvero

$$\frac{\sqrt{3}x-2(\sqrt{3}+1)y+3}{\sqrt{19+8\sqrt{3}}}=\pm\frac{\sqrt{3}x+2(\sqrt{3}-1)y-3}{\sqrt{19-8\sqrt{3}}}.$$

Tenendo conto che

$$\sqrt{19\pm 8\sqrt{3}}=4\pm\sqrt{3},$$

ed eseguendo tutte le semplificazioni, si ottengono proprio le equazioni della tangente e della normale. La figura 8.65 illustra la situazione. □

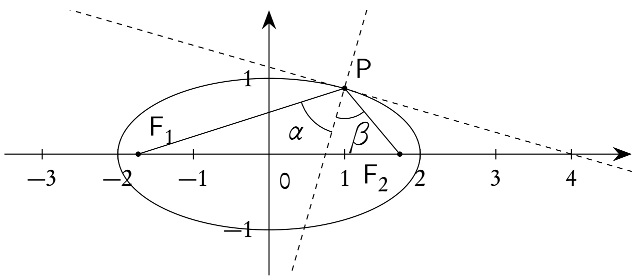


Figura 8.65.: *Proprietà di riflessione dell'ellisse*

Esercizio 8.4. *Si consideri il triangolo* ABC *di vertici* A(0,3), B(−2,0) *e* C(2,0)*. Scelto un punto* D *sul lato* $\overline{\mathsf{AB}}$ *e detta* H *la sua proiezione su* $\overline{\mathsf{BC}}$*, si prenda su* $\overline{\mathsf{BC}}$ *un punto* E *tale che* $|\mathsf{HE}|=2$*. Si provi che le perpendicolari per* E *a* DE *fanno parte di un fascio proprio di rette, di cui si chiede il centro.*

Risoluzione. La retta AB ha equazione $3x-2y+6=0$. Indicata con s l'ascissa di D, si ha

$$\mathsf{D}\left(s,\frac{3}{2}s+3\right),\quad \mathsf{H}(s,0),\quad \mathsf{E}(s+2,0).$$

Il coefficiente angolare di DE è

$$-\frac{\frac{3}{2}s+3}{2}=-\frac{3s+6}{4}.$$

La perpendicolare per E a DE avrà dunque equazione

$$y=\frac{4}{3s+6}(x-s-2)\quad\Rightarrow\quad 4x-6y-8-s(3y+4)=0.$$

Questo basta per concludere che le perpendicolari cercate fanno parte del fascio proprio individuato dalle due rette

$$2x-3y-4=0\quad \text{e}\quad 3y+4=0,$$

con centro in $\mathsf{Q}(0,-4/3)$. □

Esercizio 8.5. *Sono date le due circonferenze di centro* $\mathsf{O}(0,0)$ *e* $\mathsf{O}'(2,0)$ *e raggi* $r=1$ *e* $R=5$ *rispettivamente. Si dimostri che il luogo geometrico descritto dai centri delle circonferenze tangenti comuni ad entrambe le circonferenze date è costituito da una coppia di ellissi confocali (cioè aventi gli stessi fuochi). Generalizzare il problema considerando due circonferenze qualunque del piano e ricercando il luogo dei centri delle circonferenze tangenti ad entrambe. Per questo conviene considerare separatamente le varie situazioni che si possono presentare relativamente alle posizioni reciproche delle due circonferenze.*

Risoluzione. Non è difficile rendersi conto che ci sono due famiglie di circonferenze tangenti comuni alla due circonferenze date. Con riferimento alla figura 8.66, si tratta delle circonferenze come quella di centro C' che sono esternamente tangenti alla circonferenza piccola, e delle circonferenze come quella di centro C che sono internamente tangenti alla circonferenza piccola. Detto ϱ il raggio delle circonferenze incognite, per C' si ha

$$|\overline{\mathsf{C}'\mathsf{O}}|=\varrho+1,\quad |\overline{\mathsf{C}'\mathsf{O}'}|=5-\varrho.$$

Per C si ha, invece,

$$|\overline{\mathsf{CO}}|=\varrho-1,\quad |\overline{\mathsf{CO}'}|=5-\varrho.$$

Dunque

$$|\overline{\mathsf{C}'\mathsf{O}}|+|\overline{\mathsf{C}'\mathsf{O}'}|=6,\quad |\overline{\mathsf{CO}}|+|\overline{\mathsf{CO}'}|=4.$$

Questo basta per affermare che i luoghi descritti da C e C' rispettivamente sono ellissi di fuochi O e O', quindi con $2c=2$ e centro in $(1,0)$; la prima ha $2a=6$, la seconda $2a=4$. Dunque $b^2=8$ e $b^2=3$ rispettivamente. Le equazioni sono, rispettivamente,

$$\frac{(x-1)^2}{9}+\frac{y^2}{8}=1,\quad \frac{(x-1)^2}{4}+\frac{y^2}{3}=1.$$

La risoluzione proposta può essere generalizzata molto facilmente; si ottiene quanto segue.

— Se le due circonferenze date sono interne, ma non concentriche, si ha sempre una coppia di ellissi;
— se le due circonferenze date sono secanti si ha una ellisse e un'iperbole;
— se le due circonferenze date sono esterne si hanno due iperboli;
— se le due circonferenze date sono esternamente tangenti si ha un'iperbole e la retta che passa per i due centri;
— se le due circonferenze date sono internamente tangenti si ha un'ellisse e la retta che passa per i due centri;
— se le due circonferenze date sono concentriche si hanno due circonferenze.

□

Esercizio 8.6. *Data l'iperbole equilatera di equazione* $xy=1$*, si considerino tre suoi punti qualsiasi* A, B, C. *Provare che l'ortocentro del triangolo* ABC *appartiene ancora all'iperbole.*

Risoluzione. Si faccia riferimento alla figura 8.67.

Il coefficiente angolare della retta AB è

$$m_{\mathsf{AB}}=\frac{y_{\mathsf{B}}-y_{\mathsf{A}}}{x_{\mathsf{B}}-x_{\mathsf{A}}}=\frac{\dfrac{1}{x_{\mathsf{B}}}-\dfrac{1}{x_{\mathsf{A}}}}{x_{\mathsf{B}}-x_{\mathsf{A}}}=-\frac{1}{x_{\mathsf{A}}x_{\mathsf{B}}}.$$

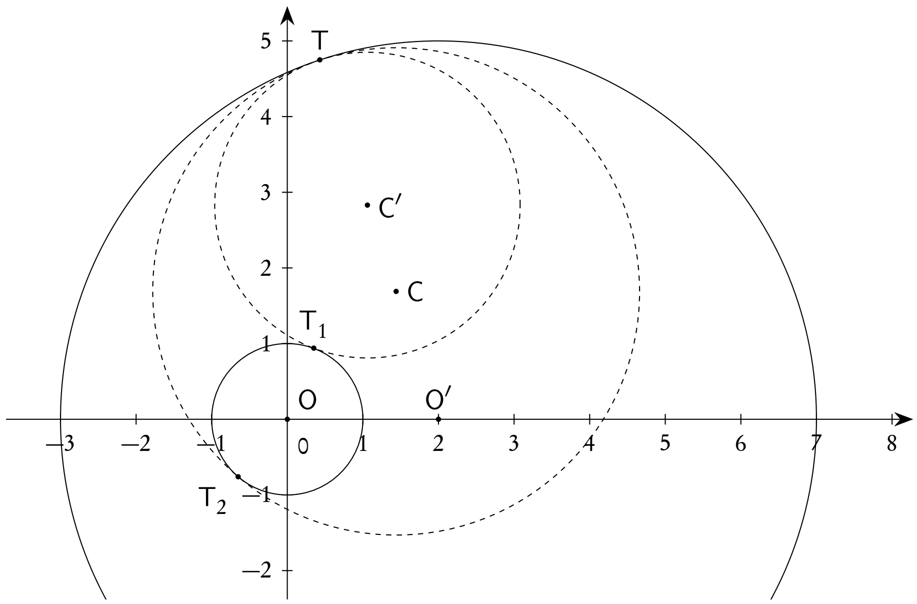


Figura 8.66.: *Circonferenze tangenti a due circonferenze*

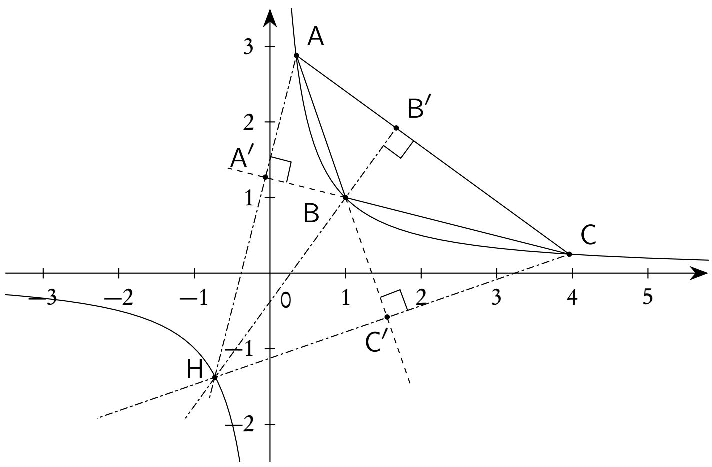


Figura 8.67.: *Ortocentro di un triangolo con vertici sull'iperbole $xy = 1$*

Il coefficiente angolare della perpendicolare $\mathsf{CC'}$ è dunque

$$m_{\mathsf{CC'}} = x_\mathsf{A} x_\mathsf{B}.$$

Analogamente si avrà

$$m_{\mathsf{AA'}} = x_\mathsf{B} x_\mathsf{C}.$$

Per trovare l'ortocentro H basterà mettere a sistema le equazioni delle due rette $\mathsf{AA'}$ e $\mathsf{CC'}$:

$$\begin{cases} y - y_\mathsf{A} = x_\mathsf{B} x_\mathsf{C} (x - x_\mathsf{A}) \\ y - y_\mathsf{C} = x_\mathsf{A} x_\mathsf{B} (x - x_\mathsf{C}) \end{cases}$$

Risolvendo questo sistema si trova

$$\mathsf{H}\left(\frac{-1}{x_\mathsf{A}x_\mathsf{B}x_\mathsf{C}}, -x_\mathsf{A}x_\mathsf{B}x_\mathsf{C}\right) \quad \Rightarrow \quad x_\mathsf{H}y_\mathsf{H} = 1,$$

ovvero il punto H appartiene all'iperbole. □

Esercizio 8.7. *È data la retta* r: $2x - 3y + 6 = 0$. *Si trovi per quali valori del parametro* $t \in \mathbb{R}$ *il punto*

$$\mathsf{P}(-3 + 3t^2, 2t)$$

giace, rispetto ad r *dalla stessa banda del punto* A(1, 1) *e per quali, invece, giace dalla banda opposta.*

Risoluzione. Troviamo l'equazione cartesiana del luogo descritto da P al variare di t:

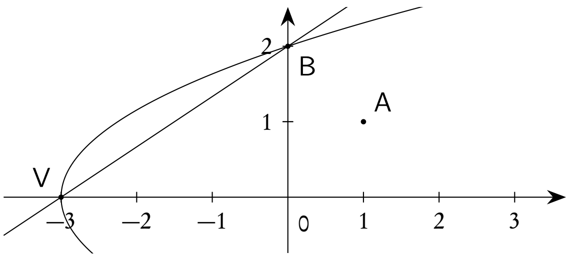
$$\begin{cases} x = -3 + 3t^2 \\ y = 2t \end{cases} .$$

Ricavando t dalla seconda equazione e sostituendo nella prima si ottiene

$$x = \frac{3}{4}y^2 - 3 :$$

si tratta dunque di una parabola con asse orizzontale e vertice nel punto V(−3, 0). La parabola e la retta si intersecano nei punti V(−3, 0) e A(0, 2), corrispondenti ai valori 0 e 1 del parametro t.

Con riferimento anche alla figura seguente, si conclude subito che il punto P sta dalla stessa banda di A se $t < 0 \vee t > 1$, dalla banda opposta se $0 < t < 1$.



□

Esercizio 8.8. *Dato un triangolo* ABC *e detto* H *il piede dell'altezza condotta da* A *a* BC, *si assuma un sistema di assi cartesiani in cui* BC *è l'asse* x *e* AH *è l'asse* y. *Si calcolino le ordinate dei piedi delle perpendicolari condotte da* A *alle* 4 *bisettrici (interne ed esterne) degli angoli in* B *e in* C. *Che cosa se ne deduce sui* 4 *punti citati?*

Risoluzione. Con riferimento alla figura 8.68, siano A(0, a), B(b, 0) e C(c, 0) i tre vertici del triangolo.

Le rette AB e AC hanno equazioni

$$\mathsf{AB}: ax + by - ab = 0 \quad \text{e} \quad \mathsf{AC}: ax + cy - ac = 0.$$

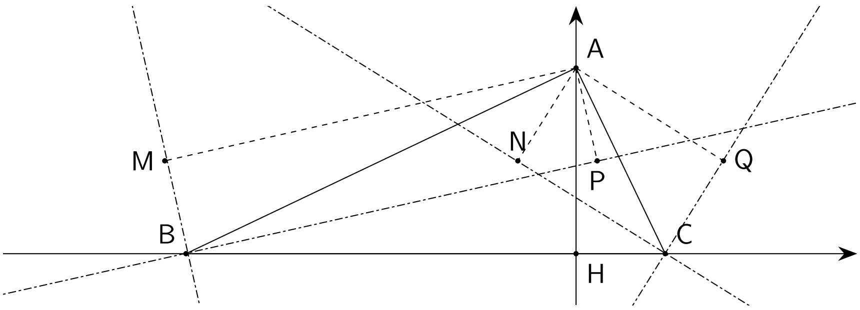


Figura 8.68.: *Perpendicolari condotte dal vertice* A *di un triangolo alle bisettrici degli angoli in* B *e* C

Tenendo conto che l'asse x (ovvero la retta che contiene il lato $\overline{BC}$ ha equazione $y=0$, le bisettrici degli angoli in B e C sono

$$\frac{|ax+by-ab|}{\sqrt{a^2+b^2}}=|y| \quad \text{e} \quad \frac{|ax+cy-ac|}{\sqrt{a^2+c^2}}=|y|.$$

Esaminiamo in particolare una di queste quattro rette e determiniamone la perpendicolare condotta da A: il sistema tra le due ci darà il primo dei piedi delle perpendicolari cercate.

$$\begin{cases} ax+\left(b-\sqrt{a^2+b^2}\right)y-ab=0 \\ \left(b-\sqrt{a^2+b^2}\right)x-ay+a^2=0 \end{cases}.$$

Moltiplicando la prima equazione per $b-\sqrt{a^2+b^2}$, la seconda per a e sottraendo le due, si trova facilmente $y={}^a/2$. Lo stesso identico risultato si ottiene esaminando le altre tre bisettrici con rispettive perpendicolari.

Se ne conclude che i quattro piedi sono allineati sulla parallela al lato $\overline{BC}$ che biseca sia $\overline{AB}$ che $\overline{AC}$. □

Esercizio 8.9. *Si consideri il triangolo* ABC *ove* $A=(-1,0)$, $B=(2,0)$, $C=(0,1)$. *Si provi che le tre bisettrici degli angoli esterni incontrano le rette dei lati opposti in tre punti allineati.*

Risoluzione. Si faccia riferimento alla figura 8.69.

Le equazioni di AC e CB sono, rispettivamente,

$$x-y+1=0 \quad \text{e} \quad x+2y-2=0.$$

Le equazioni delle due bisettrici degli angoli in A sono dunque

$$\frac{|x-y+1|}{\sqrt{2}}=|y|;$$

quella degli angoli esterni (cioè quella con coefficiente angolare negativo) è

$$x+y(\sqrt{2}-1)+1=0.$$

Analogamente si trovano le bisettrici degli angoli esterni in B e C:

$$x+(2-\sqrt{5})y-2=0, \quad (\sqrt{5}-\sqrt{2})x-(\sqrt{5}+2\sqrt{2})y+\sqrt{5}+2\sqrt{2}=0.$$

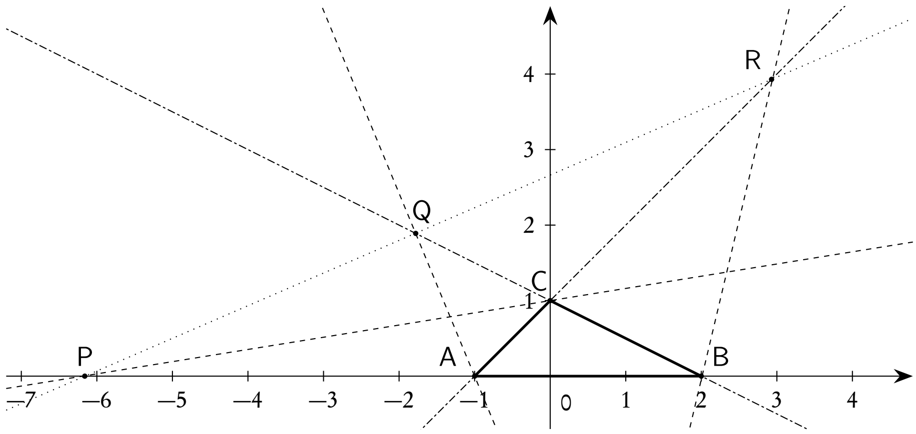


Figura 8.69.: *Proprietà delle bisettrici degli angoli esterni di un triangolo*

Le coordinate dei punti di intersezione P, Q e R si trovano risolvendo tre sistemi:

$$\mathsf{P}\colon \begin{cases} (\sqrt{5}-\sqrt{2})x-(\sqrt{5}+2\sqrt{2})y+\sqrt{5}+2\sqrt{2}=0 \\ y=0 \end{cases}, \quad \Rightarrow \quad \mathsf{P}\left(-\sqrt{10}-3\,,0\right);$$

$$\mathsf{Q}\colon \begin{cases} x+y(\sqrt{2}-1)+1=0 \\ x+2y-2 \end{cases}, \quad \Rightarrow \quad \mathsf{Q}\left(\frac{-4-6\sqrt{2}}{7},\frac{9+3\sqrt{2}}{7}\right);$$

$$\mathsf{R}\colon \begin{cases} x+(2-\sqrt{5})y-2=0 \\ x-y+1=0 \end{cases}, \quad \Rightarrow \quad \mathsf{R}\left(\frac{3\sqrt{5}+5}{4},\frac{9+3\sqrt{5}}{4}\right).$$

Si può ora trovare la retta per P e Q:

$$\left(x+\sqrt{10}+3\right)\frac{9+3\sqrt{2}}{7}=y\left(\frac{-4-6\sqrt{2}}{7}+\sqrt{10}+3\right).$$

Anche se con un po' di pazienza, si verifica ora che le coordinate di R verificano questa equazione: i tre punti sono allineati. □

Esercizio 8.10. *Di un triangolo* ABC *si conoscono i punti medi dei tre lati:* $M_1(-1,4)$, $M_2(5,0)$, $M_3(1,-2)$. *Trovare le coordinate dei vertici del triangolo.*

Risoluzione. In un triangolo la congiungente i punti medi di due lati è parallela al terzo lato. I coefficienti angolari delle congiungenti i punti medi sono:

$$m_{M_1M_2}=-\frac{2}{3}, \quad m_{M_1M_3}=-3, \quad m_{M_2M_3}=\frac{1}{2}.$$

Le equazioni delle rette contenenti i tre punti medi, e dunque i tre lati, saranno:

$$M_3(1,-2)\colon y+2=-\frac{2}{3}(x-1) \quad \text{ovvero} \quad 2x+3y+4=0;$$

$$\mathsf{M}_2(5,0)\colon y=-3(x-5) \quad \text{ovvero} \quad 3x+y-15=0;$$
$$\mathsf{M}_1(-1,4)\colon y-4=\frac{1}{2}(x+1) \quad \text{ovvero} \quad x-2y+9=0.$$

Intersecando queste rette a due a due si trovano i tre vertici: $\mathsf{A}(3,6)$, $\mathsf{B}(-5,2)$, $\mathsf{C}(7,-6)$. □

Esercizio 8.11. *Sia dato il punto* $\mathsf{P}(1,1)$ *del piano cartesiano. Condotte per* P *due rette perpendicolari, si indichi con* A *l'intersezione di una di esse con l'asse delle ascisse e con* B *l'intersezione dell'altra con l'asse delle ordinate. Si determini il luogo descritto dal punto medio* M *di* $\overline{\mathsf{AB}}$.

Risoluzione. Se m è il coefficiente angolare di una delle due rette per P (si può supporre $m \neq 0$ e $m \neq \infty$: in questi casi A e B non sarebbero definiti), le equazioni delle due rette per P sono

$$\mathsf{PA}\colon y-1=m(x-1), \quad \mathsf{PB}\colon y-1=-\frac{1}{m}(x-1).$$

Dunque

$$\mathsf{A}\left(\frac{m-1}{m},0\right), \quad \mathsf{B}\left(0,\frac{1+m}{m}\right).$$

Per il punto medio M di $\overline{\mathsf{AB}}$ si ha allora

$$x=\frac{m-1}{2m}, \quad y=\frac{1+m}{2m}.$$

Eliminando m tra queste due equazioni si ottiene

$$x+y-1=0,$$

che è l'equazione richiesta.

□

Esercizio 8.12. *Data, nel piano con riferimento cartesiano ortogonale monometrico, l'ellisse di equazione* $x^2+4y^2-9=0$, *determinare l'equazione della retta* r *sulla quale l'ellisse intercetta una corda il cui punto medio è* $\mathsf{M}(2,1)$.

Risoluzione. Una generica retta non verticale per P ha equazione $y=mx+q$. Le sue intersezioni con l'ellisse, che ci sono per tutti i valori di m e q visto che M è interno all'ellisse, si trovano risolvendo il sistema

$$\begin{cases} x^2+4y^2-9=0 \\ y=mx+q \end{cases},$$

di equazione risolvente

$$(1+4m^2)x^2+8mqx+4q^2-9=0.$$

Poiché il Δ è positivo, visto che ci sono sempre intersezioni, si ha

$$\frac{x_1+x_2}{2}=-\frac{b}{2a}=-\frac{4mq}{1+4m^2}.$$

Poiché $y_1 = mx_1 + q$ e $y_2 = mx_2 + q$, si trova

$$\frac{y_1+y_2}{2} = \frac{q}{1+4m^2}.$$

Si ha, infine,

$$\mathsf{M}(x_\mathsf{M}, y_\mathsf{M}) = \mathsf{M}\left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right).$$

Da qui, tenendo conto che $\mathsf{M} = (2,1)$, si trovano i valori

$$m = -1/2, \quad q = 2.$$

La retta richiesta ha equazione $x + 2y - 4 = 0$. □

Esercizio 8.13. *Si consideri la parabola* $y = x^2$, *tre suoi punti* A, B *e* C, *e le tre tangenti* r, s *e* t *nei punti* A, B *e* C *rispettivamente. Si provi che il rapporto tra l'area del triangolo* ABC *e quella del triangolo individuato dalle tre tangenti è costante.*

Risoluzione. Siano (a,a^2), (b,b^2) e (c,c^2) le coordinate dei tre punti dati: possiamo supporre $a \neq b \neq c$ per evitare triangoli degeneri. Per determinare l'area del triangolo ABC possiamo trovare la lunghezza di $\overline{\mathsf{AB}}$ e la distanza di C dalla retta AB.

$$|\overline{\mathsf{AB}}| = \sqrt{(a-b)^2 + (a^2-b^2)^2} = \sqrt{(a-b)^2 + (a-b)^2(a+b)^2} = |a-b|\sqrt{1+(a+b)^2}.$$

Per l'equazione della retta AB abbiamo

$$(x-a)(b^2-a^2) = (y-a^2)(b-a) \quad \Rightarrow \quad x(a+b) - y - ab = 0.$$

Dunque

$$d(\mathsf{C}, \mathsf{AB}) = \frac{|c(a+b) - c^2 - ab|}{\sqrt{1+(a+b)^2}}.$$

Infine:

$$\mathcal{A}(\mathsf{ABC}) = \frac{1}{2}|a-b|\sqrt{1+(a+b)^2}\frac{|c(a+b)-c^2-ab|}{\sqrt{1+(a+b)^2}} == \frac{1}{2}|a-b|\left|c(a+b)-c^2-ab\right|.$$

Tenendo conto della formula (8.57) relativa al coefficiente angolare della tangente ad una parabola in un suo punto, le tangenti in A, B e C hanno equazioni, rispettivamente,

$$y = 2ax^2 - a^2, \quad y = 2bx^2 - b^2, \quad y = 2cx^2 - c^2.$$

Intersecando a due a due le tre tangenti si trovano i punti

$$\mathsf{P}\left(\frac{a+b}{2}, ab\right), \quad \mathsf{Q}\left(\frac{a+c}{2}, ac\right), \quad \mathsf{R}\left(\frac{b+c}{2}, bc\right),$$

rispettivamente per le intersezioni tra le tangenti in A e B, in A e C, in B e C. La lunghezza del segmento $\overline{\mathsf{QR}}$ è

$$|\overline{\mathsf{QR}}| = \sqrt{\left(\frac{a+c}{2} - \frac{b+c}{2}\right)^2 + (ac.bc)^2} = \frac{|a-b|}{2}\sqrt{1+4c^2}.$$

La retta QR ha equazione

$$\left(x - \frac{a+c}{2}\right)(bc - ac) = (y - ac)\left(\frac{b+c}{2} - \frac{a+c}{2}\right) \quad \Rightarrow \quad 2cx - y - c^2 = 0.$$

La distanza tra P e la retta QR è

$$d(\mathsf{P}, \mathsf{QR}) = \frac{\left|2c\dfrac{a+b}{2} - ab - c^2\right|}{\sqrt{1+4c^2}} = \frac{|c(a+b) - ab - c^2|}{\sqrt{1+4c^2}}.$$

Infine

$$\mathcal{A}(\mathsf{PQR}) = \frac{1}{2}\frac{|a-b|}{2}\sqrt{1+4c^2}\frac{|c(a+b)-ab-c^2|}{\sqrt{1+4c^2}} = \frac{1}{4}|a-b|\left|c(a+b)-ab-c^2\right|.$$

Se ne conclude che

$$\frac{\mathcal{A}(\mathsf{ABC})}{\mathcal{A}(\mathsf{PQR})} = 2,$$

ovvero che il rapporto è costante. □

Esercizio 8.14. *Si trovino i punti di intersezione delle due coniche*

$$9x^2 + 16y^2 - 1 = 0, \quad 400x^2 - 225y^2 - 7 = 0,$$

verificando che essi appartengono ad una circonferenza di cui si chiede l'equazione. Si mostri anche che le tangenti alle due coniche nei punti comuni sono tra di loro perpendicolari.

Risoluzione. La prima conica è un'ellisse, in forma canonica, di semiassi $1/3$ e $1/4$; la seconda è un'iperbole, in forma canonica, con fuochi sull'asse x e semiassi $\sqrt{7}/20$ e $\sqrt{7}/15$.

Per trovare i punti di intersezione risolviamo il sistema

$$\begin{cases} 9x^2 + 16y^2 - 1 = 0 \\ 400x^2 - 225y^2 - 7 = 0 \end{cases}.$$

Si trova $x^2 = y^2 = 1/25$. I quattro punti di intersezione sono dunque

$$\mathsf{A}\left(\frac{1}{5}, \frac{1}{5}\right), \quad \mathsf{B}\left(-\frac{1}{5}, \frac{1}{5}\right), \quad \mathsf{C}\left(-\frac{1}{5}, -\frac{1}{5}\right), \quad \mathsf{D}\left(\frac{1}{5}, -\frac{1}{5}\right).$$

Semplicemente per questioni di simmetria si conclude che i quattro punti sono equidistanti dall'origine e dunque appartengono ad una circonferenza di centro l'origine. Per trovare il raggio basta trovare, per esempio, la distanza di A dall'origine: $|\overline{\mathsf{AO}}| = \sqrt{2}/5$. L'equazione della circonferenza è

$$x^2 + y^2 = \frac{2}{25}.$$

Sempre per questioni di simmetria basterà provare la perpendicolarità delle tangenti in uno dei quattro punti di intersezione: scegliamo il punto A. Per trovare le tangenti usiamo le formule di sdoppiamento, ottenendo

$$9\frac{1}{5}x + 16\frac{1}{5}y - 1 = 0, \quad \text{ovvero} \quad 9x + 16y - 5 = 0,$$

e

$$400\frac{1}{5}x - 225\frac{1}{5}y - 7 = 0, \quad \text{ovvero} \quad 80x - 45y - 7 = 0.$$

Essendo $9 \times 400 - 16 \times 225 = 0$, le due rette sono perpendicolari. □

9. Geometria euclidea piana

In questo capitolo presenteremo i concetti essenziali della geometria euclidea del piano. Segnaliamo subito che la presentazione degli argomenti non è strettamente sequenziale: abbiamo quasi sempre preferito raggruppare argomenti simili in un un'unica sezione, anche se logicamente dovrebbero far parte di sezioni diverse (e magari potrebbero richiedere prerequisiti diversi). Anche se non vogliamo proporre un semplice formulario, lo scopo di questo capitolo è principalmente quello di costituire un manuale di rapida consultazione e non quello di proporre un corso formale. Come per il resto di questo testo, non proporremo le dimostrazioni dei teoremi via via enunciati e, nei casi più semplici, enunceremo le proprietà senza nemmeno proporle sotto forma di teorema.

La geometria euclidea costituisce il primo, e forse più importante, esempio di scienza *ipotetico deduttiva*. Con riferimento a quanto già accennato nel paragrafo 1.4, si accettano alcuni *concetti primitivi* e si ritengono vere alcune affermazioni dette *assiomi* o *postulati*. Da qui si deducono poi logicamente tutte le affermazioni che formano il corpo della geometria, ovvero i *teoremi*.

Naturalmente la scelta dei concetti primitivi e degli assiomi è arbitraria: quella che noi adotteremo (anche senza farci esplicita e dettagliata menzione) è la scelta fatta da Euclide nei suoi *Elementi*, sistemata e completata da Hilbert nei suoi *Grundlagen der Geometrie* (Fondamenti della Geometria).

Riservandoci comunque di riprendere il discorso in seguito, tra i concetti primitivi segnaliamo quelli di *punto*, *retta*, *piano*, *spazio* (per i quali dunque, lo ripetiamo, non si dà alcuna definizione); tra gli assiomi segnaliamo, per la sua enorme importanza storica, il *Postulato delle parallele* (o "Quinto postulato"), che formuleremo nella maniera tradizionale, anche se diversa da quella di Euclide: "in un piano, per un punto fuori da una retta, si può condurre una sola parallela a una retta data".

Supporremo di disporre, fin dall'inizio, della possibilità di misurare le grandezze geometriche oggetto di indagine, anche se introdurremo in seguito in maniera formale i concetti relativi, in particolare quelli di *lunghezza* di un segmento, di *ampiezza* di un angolo, di *area* di una superficie.

Prima di addentrarci nel riepilogo dei concetti e teoremi fondamentali, spendiamo due parole sul concetto di *congruenza*. Hilbert assume questo concetto come primitivo, formulando una serie di assiomi che lo precisano: questi assiomi rendono formalmente accettabile il concetto di "uguaglianza mediante movimento rigido" usato da Euclide: *Cose che si possono sovrapporre una con l'altra sono uguali*. In realtà il concetto di movimento rigido non può essere fatto risalire agli altri assiomi della geometria e quindi è corretto assumerlo come primitivo. Segnaliamo che per la maggior parte i testi moderni di geometria usano il termine "congruente" per indicare due figure sovrapponibili mediante movimento rigido. Tuttavia molti usano anche il termine "uguale" per descrivere la stessa proprietà: è naturalmente solo questione di scelta, e anche noi di norma faremo questa seconda scelta.

9.1. Concetti fondamentali

I concetti su cui si basa la geometria sono quelli di *punto*, *retta*, *piano*, *spazio*. Questi concetti sono primitivi, nel senso che non si dà alcuna definizione per essi: saranno gli assiomi relativi a precisarne le caratteristiche. Tuttavia, anche seguendo Euclide, possiamo dire che ci si forma intuitivamente il concetto di punto osservando corpi via via più piccoli, fino ad arrivare ad un ente “privo di dimensioni”: *il punto è ciò che non ha parti*, sempre per dirla con Euclide[1]. Indichiamo i punti con lettere maiuscole dell’alfabeto latino: A, B, …In maniera analoga ci si forma il concetto di retta osservando un filo teso e immaginando di prolungarlo all’infinito: ancora per dirla con Euclide, *una linea è lunghezza senza larghezza*. Euclide stesso, formalizzando l’idea del filo teso e poi prolungato, sostanzialmente introduce la retta come prolungamento illimitato di una “retta finita”. Le rette si indicano con una lettera minuscola dell’alfabeto, o con due lettere maiuscole indicanti due loro punti: r, s, …, AB, …Ancora, ci si forma l’idea di piano pensando ad una superficie, oggetto che ha solo lunghezza e larghezza, prolungata all’infinito in tutte le direzioni e in modo che, sempre seguendo Euclide, *giaccia ugualmente rispetto alle rette su di essa*. I piani si indicano con le lettere minuscole dell’alfabeto greco: α, β, …Infine lo spazio è semplicemente l’insieme di tutti i punti.

Definizione 9.1. *Si chiama* figura geometrica *un qualunque sottoinsieme di punti dello spazio.*

È chiaro che formano oggetto di studio specifico della geometria solo alcune particolari figure geometriche.

9.1.1. Gli assiomi sulla retta

Per due punti distinti passa una e una sola retta. I punti di una retta sono ordinati in due versi distinti, opposti l’uno dell’altro, in modo che non v’è né un primo né un ultimo punto. Fra due punti vi sono infiniti punti intermedi. Per un punto fuori da una retta si può condurre, in un piano, una sola parallela alla retta data.

Definizione 9.2. *Sia* r *una retta e* O *un suo punto. Il punto* O *individua sulla retta due sottoinsiemi, detti* semirette, *costituiti ciascuno dal punto* O *stesso, detto* origine, *e dai punti che lo seguono in uno dei due versi possibili sulla retta. Due semirette situate sulla stessa retta e aventi in comune solo l’origine si dicono* semirette opposte.

Se su una semiretta si sceglie, oltre all’origine O, un altro punto P, la semiretta si potrà indicare con OP; si potranno anche usare le lettere minuscole dell’alfabeto latino, come per le rette.

9.1.2. Segmenti

Definizione 9.3. *Dati due punti distinti* A *e* B *di una retta* r, *dicesi* segmento $\overline{\mathsf{AB}}$ *il sottoinsieme di* r *costituito dai punti compresi tra* A *e* B.

I punti A e B si dicono *estremi* del segmento; ogni altro punto diverso da A e B si dice *interno* al segmento; gli altri punti della retta r cui appartiene il segmento si dicono *esterni* al segmento.

Due segmenti $\overline{\mathsf{AB}}$ e $\overline{\mathsf{BC}}$ aventi in comune solo l’estremo B si dicono *consecutivi*. Se due segmenti consecutivi appartengono alla stessa retta si dicono *adiacenti*. Anche per le semirette si usano i termini consecutive e adiacenti, con ovvio significato.

Più segmenti a due a due consecutivi e non adiacenti costituiscono una *poligonale*. Se il primo e l’ultimo estremo della poligonale sono distinti essa si dice *aperta*, altrimenti *chiusa*. Se due segmenti consecutivi hanno un punto in comune, diverso dai vertici, la poligonale si dice *intrecciata*.

[1] Non si prenda questa come *definizione* di punto: si tratta solo di un aiuto per arrivare a formarsi un concetto corretto. Analogo discorso per gli altri enti fondamentali.

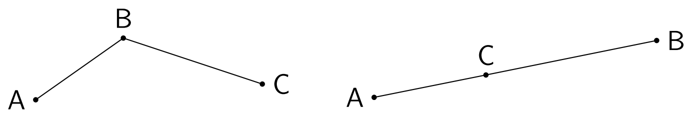


Figura 9.1.: *Segmenti consecutivi (a sinistra) e adiacenti (a destra)*

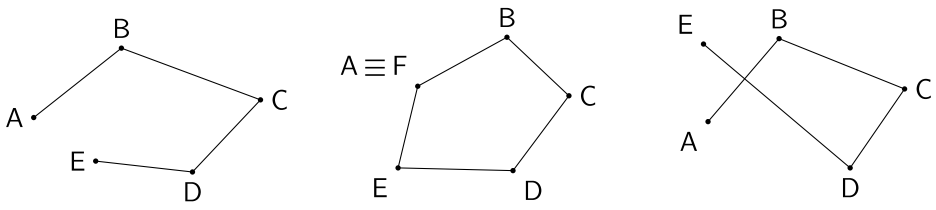


Figura 9.2.: *Poligonale aperta (a sinistra), chiusa (al centro), intrecciata (a destra)*

I segmenti si possono confrontare, sommare e sottrarre: non entriamo nei dettagli delle definizioni, che sono abbastanza intuitive.

9.1.3. Gli assiomi sul piano

Per tre punti a due a distinti e non allineati passa uno e un solo piano. Una retta r *qualsiasi di un piano π individua sul piano due sottoinsiemi, detti* semipiani, *ciascuno costituito dalla retta stessa, detta* origine *e da due sottoinsiemi π_1 e π_2 non aventi alcun punto in comune e tali che ogni segmento avente entrambi gli estremi in uno dei due sottoinsiemi π_1 e π_2 non ha alcun punto in comune con* r, *mentre ogni segmento avente un estremo in uno dei due sottoinsiemi e l'altro nell'altro ha un solo punto in comune con* r.

I semipiani si indicano con le lettere minuscole dell'alfabeto greco, come i piani.

Definizione 9.4. *Una figura si dice* convessa *se, comunque si scelgano due suoi punti, il segmento che li congiunge appartiene interamente alla figura; si dice invece* concava *se esistono segmenti che congiungono due punti della figura ma non sono sottoinsiemi della stessa.*

L'intersezione di due figure convesse è sempre una figura convessa. L'unione di due figure convesse, invece, può anche essere concava.

I piani, i semipiani, le rette, le semirette, i segmenti sono figure convesse. La figura costituita da due segmenti consecutivi ma non adiacenti è un esempio di figura concava.

9.1.4. Angoli e strisce

Definizione 9.5. *Date due semirette distinte di un piano, aventi l'origine in comune, si chiama* angolo *ciascuno dei due sottoinsiemi individuati dalle due semirette, unito con le semirette stesse.*

Le due semirette si chiamano *lati*, la comune origine si chiama *vertice*. Se le due semirette sono adiacenti, ciascuno dei due angoli (in realtà un semipiano) si chiama *angolo piatto*. Se le due semirette non sono adiacenti uno dei due angoli è una figura convessa e si chiama *angolo convesso*, l'altra è concava, e si chiama *angolo concavo*. Più in dettaglio un angolo è convesso se non contiene i prolungamenti dei suoi lati, concavo se invece li contiene. Per estensione, se le due semirette coincidono si chiama *angolo giro* l'intero piano, *angolo nullo* l'insieme costituito solo dalle due semirette, e in entrambi i casi i lati si dicono coincidenti. Si noti che un angolo giro e un angolo piatto sono figure convesse; tuttavia, parlando di angoli convessi generalmente non si fa riferimento ad essi.

Gli angoli si indicano con le lettere minuscole dell'alfabeto greco oppure con scritture del tipo $A\widehat{O}B$, ove A e B sono punti, diversi dal vertice, uno su un lato e uno sull'altro, e O è il vertice.

Osservazione 9.6. La definizione di angolo che abbiamo dato non è l'unica possibile. In trigonometria si preferisce di solito definire un angolo come la rotazione di una semiretta attorno alla sua origine: in questo modo è possibile considerare anche angoli più grandi di un angolo giro, cosa che torna molto utile proprio in trigonometria. Torneremo su questo argomento nel capitolo 11.

Due angoli $A\widehat{O}B$ e $B\widehat{O}C$, che hanno in comune solo il vertice e un lato, si dicono *consecutivi*. Due angoli consecutivi in cui i lati non comuni sono semirette opposte si dicono *adiacenti*. Due rette incidenti in un punto O dividono il piano in quattro angoli, tutti convessi: due di essi con i lati dell'uno costituiti dai prolungamenti dei lati dell'altro si dicono *opposti al vertice*.

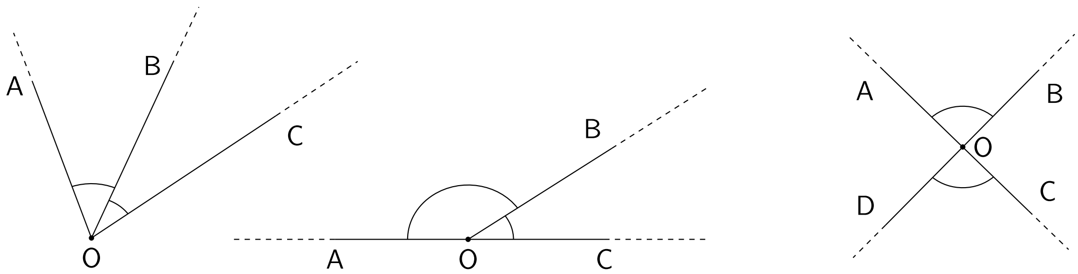


Figura 9.3.: *Angoli consecutivi (a sinistra), adiacenti (al centro), opposti al vertice (a destra)*

Gli angoli si possono sempre confrontare, mentre si possono sommare solo alla condizione che la somma non superi un angolo giro; si potrà poi sottrarre un angolo con uno minore: non entriamo nei dettagli delle definizioni che sono abbastanza intuitive.

Due angoli che hanno per somma un angolo giro si dicono *esplementari*.

Due angoli che hanno per somma un angolo piatto si dicono *supplementari*: essi possono essere sempre traslati rigidamente fino a farli diventare adiacenti.

Definizione 9.7. *Due rette incidenti si dicono* perpendicolari *se individuano quattro angoli uguali, detti angoli* retti.

Un angolo retto è la metà di un angolo piatto. Un angolo convesso si dice *acuto* se è minore di un angolo retto, si dice *ottuso* se è maggiore di un angolo retto. Due angoli acuti che hanno per somma un angolo retto si dicono *complementari*.

Data una retta r e un punto P di un piano (appartenente o no alla retta), esiste, nel piano, una unica perpendicolare s ad r condotta da P. Il punto H comune alla retta r e alla sua perpendicolare si chiama *piede* della perpendicolare s, o anche *proiezione* del punto P sulla retta r. Il segmento $\overline{PH}$ si dice *segmento di perpendicolare*.

Dato un segmento $\overline{AB}$, una retta r e le due proiezioni H e K di A e B su r, il segmento $\overline{HK}$ si chiama *proiezione* di $\overline{AB}$ su r.

Valgono i seguenti teoremi di uso continuo.

Teorema 9.8. *Angoli supplementari di angoli uguali sono uguali.*

Teorema 9.9. *Due angoli opposti al vertice sono uguali.*

Teorema 9.10. *Angoli complementari di angoli uguali sono uguali.*

Definizione 9.11. *Date due rette parallele* r *ed* s *distinte, si chiama* striscia *l'intersezione dei due semipiani di origine* r *e contenente* s *e di origine* s *e contenente* r*.*

Ogni striscia di piano è una figura convessa.

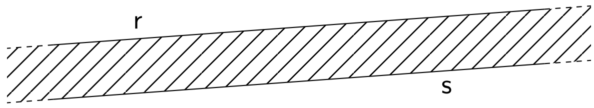


Figura 9.4.: *Una striscia di piano*

Le due rette che individuano una striscia si chiamano *lati della striscia*.

9.1.5. Rette parallele

Due rette parallele ad una terza sono tra di loro parallele. Se due rette sono parallele, ogni retta incidente una delle due è incidente anche l'altra.

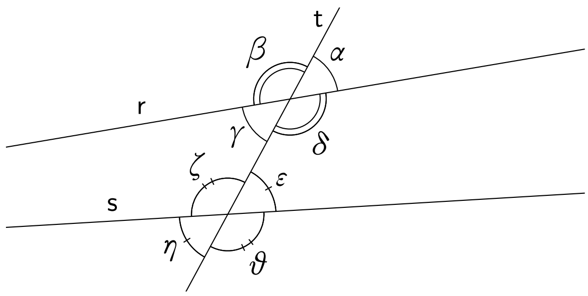


Figura 9.5.: *Angoli individuati da due rette tagliate da una trasversale*

Se due rette distinte r e s sono incontrate da una terza retta t in due punti distinti, esse formano con questa otto angoli a cui vengono dati, vedi la figura 9.5, particolari nomi:

- gli angoli δ, γ, ε, ζ si dicono *interni*;
- gli angoli α, β, η, ϑ si dicono *esterni*;
- gli angoli che si trovano da una stessa parte della retta t si dicono *coniugati*; in particolare le coppie γ e ζ, δ e ε sono angoli *coniugati interni*, le coppie β e η, α e ϑ sono angoli *coniugati esterni*;
- gli angoli che si trovano da bande opposte rispetto alla retta t si dicono *alterni*; in particolare le coppie γ e ε, δ e ϑ sono angoli *alterni interni*, le coppie β e ϑ, α e η sono angoli *alterni esterni*;
- gli angoli β e ϑ, γ e η, α e ε, δ e ϑ si dicono *corrispondenti* (essi sono coniugati, uno interno e uno esterno).

Se due rette formano

— o una coppia di angoli alterni uguali,
— o una coppia di angoli corrispondenti uguali,
— o una coppia di angoli coniugati supplementari,

allora risultano uguali tutte le coppie di angoli alterni e corrispondenti, e risultano supplementari tutte le coppie di angoli coniugati.

Il seguente teorema costituisce il criterio fondamentale di parallelismo di due rette.

Teorema 9.12. *Due rette* r *e* s *sono parallele se e solo se, tagliate da una trasversale* t *formano*

— *o una coppia di angoli alterni uguali;*
— *o una coppia di angoli corrispondenti uguali;*
— *o una coppia di angoli coniugati supplementari.*

9.2. I triangoli

Definizione 9.13. *Dati tre punti non allineati* A, B, C, *si dice* triangolo ABC *l'insieme dei punti comuni ai tre angoli convessi* $\mathsf{A\widehat{B}C}$, $\mathsf{B\widehat{C}A}$, $\mathsf{C\widehat{A}B}$. *In maniera equivalente un triangolo è l'intersezione dei semipiani di origine* AB *e contenente* C, *di origine* BC *e contenente* A, *di origine* CA *e contenente* B.

Un triangolo ABC, essendo intersezione di figure convesse, è sempre una figura convessa.

Nel triangolo ABC, i punti A, B, C si dicono *vertici*, i segmenti $\overline{\mathsf{AB}}$, $\overline{\mathsf{BC}}$, $\overline{\mathsf{CA}}$ si dicono *lati*. I tre lati complessivamente costituiscono il contorno del triangolo e il segmento somma dei tre lati è il *perimetro*. L'angolo $\mathsf{A\widehat{B}C}$ si dice *compreso* fra i lati $\overline{\mathsf{AB}}$ e $\overline{\mathsf{BC}}$ e *adiacente* a ciascuno di essi. Il lato $\overline{\mathsf{AC}}$ si dice *opposto* all'angolo $\mathsf{A\widehat{B}C}$ e al vertice B. Analogo discorso per gli altri angoli, lati e vertici. I lati e gli angoli di un triangolo si dicono complessivamente i suoi elementi.

Dato l'angolo $\mathsf{A\widehat{B}C}$ di un triangolo, ciascuno dei due angoli $\mathsf{A\widehat{B}D}$ e $\mathsf{C\widehat{B}E}$ (vedi la figura 9.6) ad esso adiacenti si chiama *angolo esterno* del triangolo.

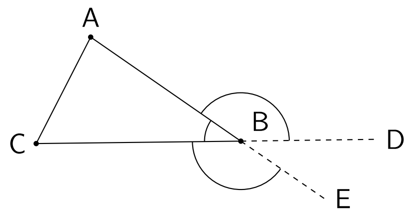


Figura 9.6.: *Angoli esterni in un triangolo*

9.2.1. I criteri di uguaglianza

Teorema 9.14 (Primo criterio di uguaglianza dei triangoli)**.** *Due triangoli che hanno rispettivamente uguali due lati e l'angolo tra essi compreso sono uguali.*

Teorema 9.15 (Secondo criterio di uguaglianza dei triangoli)**.** *Due triangoli che hanno rispettivamente uguali un lato e i due angoli ad esso adiacenti sono uguali.*

Teorema 9.16 (Terzo criterio di uguaglianza dei triangoli)**.** *Due triangoli che hanno rispettivamente uguali i tre lati sono uguali.*

Si noti, nei tre criteri, che per l'uguaglianza di due triangoli serve l'uguaglianza di tre elementi scelti tra i tre lati e i tre angoli, purché almeno uno sia un lato.

9.2.2. Relazioni tra gli elementi di un triangolo

Teorema 9.17. *La somma degli angoli interni di un triangolo è uguale ad un angolo piatto e ogni angolo esterno è uguale alla somma dei due angoli interni non adiacenti.*

Segue da questo teorema che in ogni triangolo almeno due angoli interni sono acuti, dunque al massimo un angolo può essere retto od ottuso.

Teorema 9.18. *In un triangolo con due lati disuguali, anche gli angoli opposti sono disuguali e a lato maggiore sta opposto angolo maggiore. Viceversa in un triangolo con due angoli disuguali, anche i lati opposti sono disuguali e ad angolo maggiore sta opposto lato maggiore.*

Teorema 9.19. *In un triangolo un lato è minore della somma e maggiore della differenza degli altri due.*

9.2.3. Varie specie di triangoli

Un triangolo con i tre lati (e quindi i tre angoli) disuguali si dice *scaleno*. Un triangolo si dice *isoscele* se ha due lati uguali. L'angolo individuato dai due lati uguali si chiama *angolo al vertice*, il lato opposto si chiama *base*. Un triangolo è isoscele se e solo se ha due angoli uguali. Un triangolo che ha tutti i lati uguali o, equivalentemente, tutti gli angoli uguali, si dice *equilatero* o *equiangolo*.

Un triangolo che ha tutti gli angoli acuti si dice *acutangolo*. Un triangolo che ha un angolo retto si dice *rettangolo*. Un triangolo che ha un angolo ottuso si dice *ottusangolo*.

In un triangolo rettangolo i lati che individuano l'angolo retto si chiamano *cateti*, il lato opposto si chiama *ipotenusa*.

Per i triangoli rettangoli i criteri di uguaglianza si possono riassumere nel seguente enunciato: due triangoli rettangoli sono uguali se hanno due elementi, che non siano i due angoli acuti, e diversi dall'angolo retto, ordinatamente uguali.

Si noti come questo criterio abbia una portata più generale dei criteri per i triangoli generici. Per esempio, per due triangoli generici avere uguali due lati e un angolo non garantisce l'uguaglianza dei triangoli: nella figura 9.7 i triangoli ABC_1 e ABC_2 hanno due lati e un angolo uguali, ma sono diversi.

9.2.4. Punti notevoli

Esiste uno ed un solo punto che divide un dato segmento in due segmenti uguali: esso si chiama il *punto medio* del segmento; analogamente esiste una ed una sola semiretta avente origine nel vertice di un angolo e che divide l'angolo in due angoli uguali: essa si chiama *bisettrice* dell'angolo.

Considerato in un triangolo ABC il punto medio M di un lato, si chiama *mediana* il segmento che unisce il punto M con il vertice opposto. Dunque in un triangolo vi sono tre mediane. Esse sono sempre

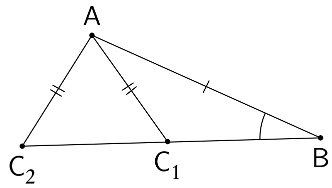


Figura 9.7.: *Triangoli diversi con due lati e un angolo uguali*

interne al triangolo e si incontrano in un punto detto *baricentro*. Il baricentro divide ciascuna mediana in due parti tali che quella contenente il vertice è doppia dell'altra.

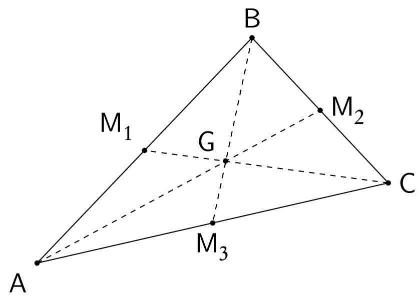


Figura 9.8.: *Mediane di un triangolo*

L'origine del nome baricentro è legata al fatto che, in fisica, se il triangolo ABC è una lamina omogenea, il baricentro come punto di intersezione delle mediane coincide con il centro di gravità o centro di massa della lamina[2] stessa.

Per ciascuno dei tre angoli interni del triangolo è possibile considerare la relativa semiretta bisettrice: essa incontra il lato opposto al vertice dell'angolo in un punto. Il segmento individuato sulla semiretta bisettrice dalla sua origine e dall'intersezione con il lato opposto si chiama semplicemente *bisettrice*. Dunque in un triangolo vi sono tre bisettrici. Esse sono sempre interne al triangolo e si incontrano in un punto detto *incentro*. L'incentro è equidistante dai tre lati dell'angolo: per questo motivo esso è il centro di una particolare circonferenza detta *inscritta* nel triangolo, e a questo fatto è dovuto il nome del punto stesso.

Da ciascuno dei tre vertici di un triangolo ABC si può condurre la perpendicolare alla retta contenente il lato opposto. Se indichiamo con A', B', C' i piedi delle perpendicolari, i segmenti $\overline{AA'}$, $\overline{BB'}$, $\overline{CC'}$ si chiamano *altezze* del triangolo. Le tre altezze, o i loro prolungamenti, si incontrano in un punto detto *ortocentro*. Diversamente dalle mediane e bisettrici le altezze non sono sempre interne e precisamente si hanno le tre situazioni rappresentate nella figura 9.10: se il triangolo è acutangolo le tre altezze e l'ortocentro sono interni; se il triangolo è rettangolo due delle tre altezze coincidono con i cateti e l'ortocentro coincide con il vertice dell'angolo retto; se il triangolo è ottusangolo l'altezza condotta dal

[2] Si presti attenzione al fatto che questa coincidenza si ha solo nella situazione descritta: per esempio il centro di massa di un telaio triangolare omogeneo *non* coincide con il punto di intersezione delle mediane del triangolo, a meno che il triangolo non sia equilatero.

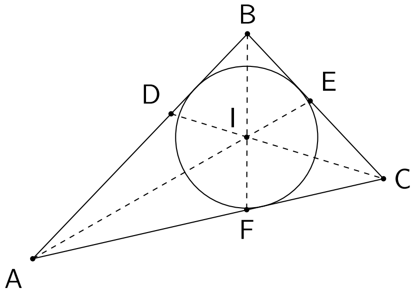


Figura 9.9.: *Bisettrici di un triangolo*

vertice dell'angolo ottuso è interna al triangolo, le altre due sono esterne e tale è anche l'ortocentro (in questo caso si incontrano nell'ortocentro i prolungamenti delle altezze).

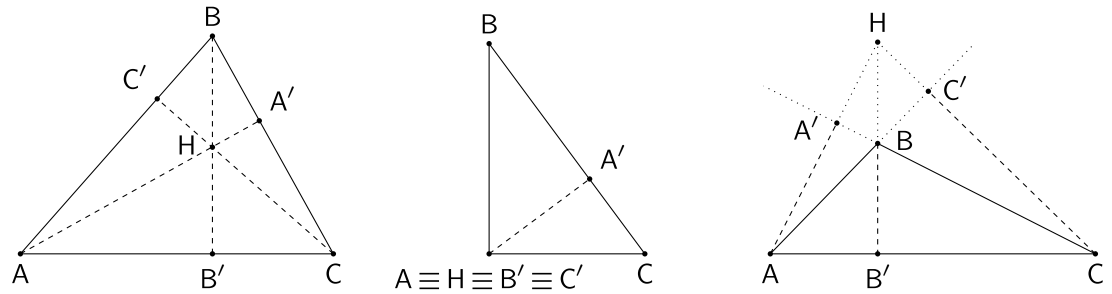


Figura 9.10.: *Altezze di un triangolo acutangolo (a sinistra), rettangolo (al centro), ottusangolo (a destra)*

Definizione 9.20. *Si dice* asse *di un segmento la perpendicolare al segmento condotta dal suo punto medio.*

Teorema 9.21. *L'asse di un segmento è il luogo geometrico dei punti del piano equidistanti dagli estremi del segmento stesso.*

In un triangolo si possono considerare gli assi[3] dei tre lati; essi si incontrano in un punto detto *circocentro*, che è equidistante dai tre vertici del triangolo e dunque è il centro della circonferenza che passa per i tre vertici del triangolo, detta circonferenza *circoscritta* al triangolo: da qui il nome. Il circocentro è interno ai triangoli acutangoli, coincide con il punto medio dell'ipotenusa nei triangoli rettangoli, è esterno nei triangoli ottusangoli.

Teorema 9.22. *In un triangolo isoscele la bisettrice dell'angolo al vertice è perpendicolare alla base e la divide in due parti uguali, dunque è anche altezza e bisettrice e la retta che la contiene è asse della base.*

Segue da questo che il baricentro, l'incentro, l'ortocentro e il circocentro sono allineati e stanno sull'asse della base. In un triangolo equilatero questa proprietà è valida per tutti e tre i vertici: i quattro punti notevoli coincidono in un punto che può essere chiamato *centro* del triangolo.

[3] Si noti che gli assi sono rette, non segmenti come le mediane bisettrici o altezze

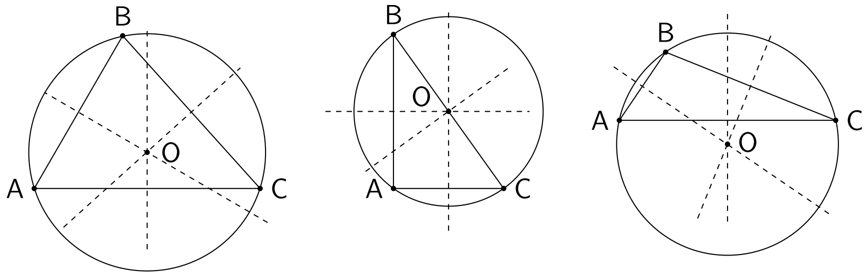


Figura 9.11.: *Assi di un triangolo acutangolo (a sinistra), rettangolo (al centro), ottusangolo (a destra)*

Teorema 9.23. *Le semirette bisettrici dei quattro angoli formati da due rette incidenti sono a due a due in linea retta e perpendicolari. Inoltre esse costituiscono il luogo geometrico dei punti del piano equidistanti dalle due rette date.*

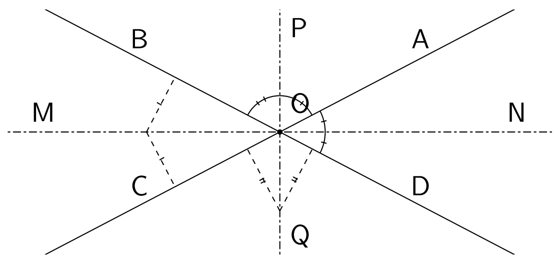


Figura 9.12.: *Bisettrici di due rette incidenti*

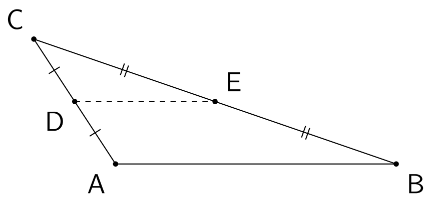


Figura 9.13.: *Proprietà di una speciale corda in un triangolo*

Teorema 9.24. *In un triangolo il segmento che unisce i punti medi di due lati è parallelo al terzo lato ed uguale alla metà di questo. Viceversa se dal punto medio di un lato si conduce la parallela ad uno degli altri due lati, tale parallela divide il lato rimanente in due parti uguali.*

9.3. I poligoni

Ci occuperemo di poligoni convessi: la trattazione dei poligoni concavi è decisamente complessa ed esula dagli scopi di un normale corso di geometria elementare. Faremo solo un breve cenno alla loro trattazione in fondo a questo paragrafo, al solo scopo di evidenziare le difficoltà connesse con questo problema.

Definizione 9.25. *Siano dati nel piano n punti (con $n > 2$)* $A_1, A_2, \ldots, A_n$*, tali che tre consecutivi non siano in linea retta e che la congiungente due a due i punti, nell'ordine in cui sono dati, lasci gli altri* $n-2$ *punti tutti in uno stesso semipiano. Si dice* poligono convesso *la parte di piano comune agli angoli* $A_n\widehat{A_1}A_2, A_1\widehat{A_2}A_3, \ldots, A_{n-1}\widehat{A_n}A_1$ *o, equivalentemente, la parte di piano comune ai semipiani di origine* $A_1A_2, A_2A_3, \ldots, A_nA_1$ *e contenenti gli altri punti.*

Poiché un poligono convesso è intersezione di figure convesse, è una figura convessa. In genere, parlando di poligono senza altre specificazioni, ci si riferisce a un poligono convesso. La poligonale $A_1A_2\ldots A_n$ che determina il poligono si dice *contorno* dello stesso, i punti $A_1, A_2, \ldots, A_n$ si dicono *vertici* e i lati della poligonale *lati* del poligono. Il segmento somma dei lati si dice *perimetro* del poligono. Gli angoli (convessi) $A_n\widehat{A_1}A_2, A_1\widehat{A_2}A_3, \ldots, A_{n-1}\widehat{A_n}A_1$ si dicono *angoli interni* del poligono o, semplicemente, angoli del poligono. I lati e gli angoli di un poligono si dicono complessivamente i suoi elementi.

I $2n$ angoli adiacenti agli angoli interni si dicono *angoli esterni*. È tradizione, nel costruire gli angoli esterni di un poligono convesso, orientare il poligono stesso in uno dei due versi possibili e prendere, per ciascun angolo interno, come angolo esterno quello che ha come primo lato il prolungamento del primo lato dell'angolo interno e come secondo lato lo stesso secondo lato dell'angolo interno.

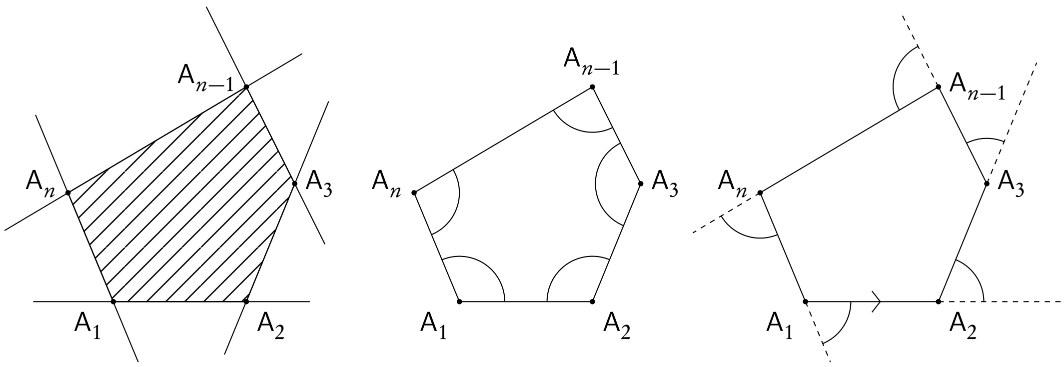


Figura 9.14.: *Poligono convesso, angoli interni (al centro), angoli esterni (a destra)*

Ciascuno dei segmenti che congiungono due vertici non consecutivi di un poligono si dice una *diagonale*. Le diagonali di un poligono di n lati sono sempre in numero di $n(n-3)/2$. Più generalmente si dice *corda* un segmento che congiunge due punti del contorno che non appartengono allo stesso lato.

I poligoni hanno nomi particolari, a seconda del numero dei lati (e quindi degli angoli): *quadrilatero* o *quadrangolo* è un poligono di 4 lati, *pentagono* uno di 5 lati, *esagono* uno di 6 lati, *eptagono* uno di 7 lati, ecc.

9.3.1. Relazioni fra lati e fra angoli in un poligono

Teorema 9.26. *In un poligono un lato è minore della somma di tutti gli altri.*

Teorema 9.27. *In un poligono convesso di n lati la somma degli angoli interni è uguale* $(n-2)$ *angoli piatti.*

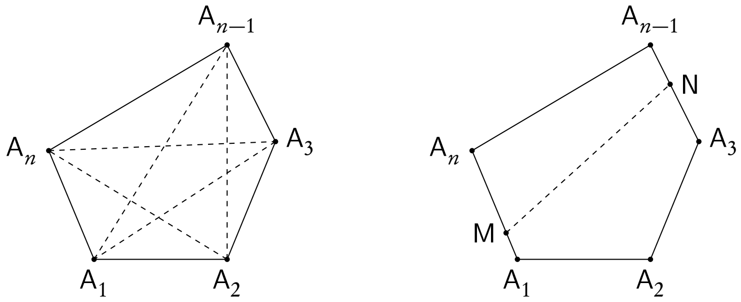


Figura 9.15.: *Diagonali e corda in un poligono*

Teorema 9.28. *In un poligono convesso la somma degli angoli esterni (uno per ciascun angolo interno) è uguale a due angoli piatti.*

9.3.2. I parallelogrammi

In un quadrilatero si dicono *opposti* due lati che non hanno vertici in comune e due angoli che non hanno lati in comune.

Definizione 9.29. *Si dice* parallelogramma *un quadrilatero in cui i lati opposti sono paralleli.*

Un segmento di perpendicolare condotto da un punto di un lato al lato opposto si dice *altezza* relativa alla coppia di lati opposti che si dicono *basi*.

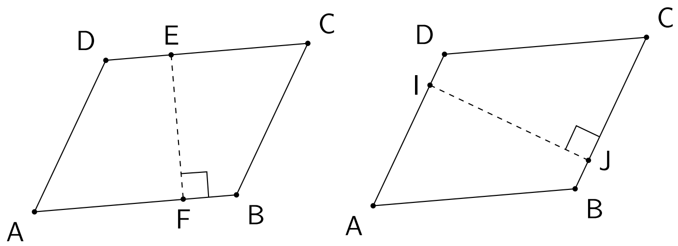


Figura 9.16.: *Parallelogrammi, altezze e basi*

Teorema 9.30. *In ogni parallelogramma*

— *ciascuna diagonale lo divide in due triangoli uguali;*
— *i lati opposti sono uguali;*
— *gli angoli opposti sono uguali;*
— *gli angoli adiacenti a ciascun lato sono supplementari;*
— *le due diagonali si dividono scambievolmente a metà.*

Teorema 9.31. *Valgono le seguenti condizioni per riconoscere se un quadrilatero è un parallelogramma.*

— *Un quadrilatero convesso con le due coppie di lati opposti uguali è un parallelogramma.*

— *Un quadrilatero convesso con le due coppie di angoli opposti uguali è un parallelogramma.*
— *Un quadrilatero convesso con due lati opposti uguali e paralleli è un parallelogramma.*
— *Un quadrilatero convesso in cui le diagonali si dividono scambievolmente a metà è un parallelogramma.*

Parallelogrammi particolari

In un parallelogramma i lati e gli angoli sono uguali a due a due. Situazioni particolari riguardano i casi in cui

— i quattro angoli sono tutti uguali: si ottiene un *rettangolo*;
— i quattro lati sono tutti uguali: si ottiene un *rombo* o *losanga*;
— sia i quattro lati che i quattro angoli sono uguali: si ottiene un *quadrato*.

In un rettangolo, se uno dei lati è una base, i suoi due lati consecutivi sono altezze. La base e l'altezza si dicono anche le *dimensioni* del rettangolo. In un rettangolo le due diagonali sono uguali e, viceversa, se in un parallelogramma le due diagonali sono uguali esso è un rettangolo.

In un rombo le diagonali sono perpendicolari e bisettrici degli angoli e, viceversa, se in un parallelogramma le diagonali o sono perpendicolari o bisecano gli angoli allora il parallelogramma è un rombo.

Il quadrato riunisce in sè le proprietà del rettangolo e del rombo; dunque le diagonali sono uguali, perpendicolari e bisecano gli angoli.

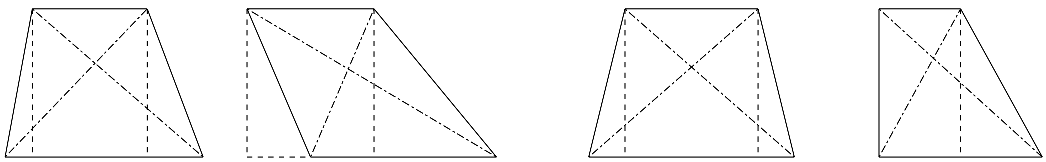
9.3.3. I trapezi

Definizione 9.32. *Un quadrangolo convesso con due soli lati opposti paralleli e gli altri non paralleli si dice* trapezio.

I due lati paralleli si chiamano *basi* del trapezio e si distinguono in *base maggiore* e *base minore*; i restanti due lati si chiamano *lati obliqui*; il segmento di perpendicolare condotto da un punto di una base all'altra si chiama *altezza*.

Un trapezio con i lati obliqui disuguali si dice *scaleno*; un trapezio in cui i lati obliqui sono uguali si dice *isoscele*; un trapezio in cui uno dei lati obliqui è perpendicolare alla base si dice *rettangolo*.

In ogni trapezio gli angoli adiacenti ai lati obliqui, in quanto coniugati interni rispetto a due rette parallele, sono supplementari. In un trapezio isoscele gli angoli adiacenti a ciascuna base sono uguali, le diagonali sono uguali, le proiezioni dei lati obliqui sulla base maggiore sono uguali.



(a) Trapezio scaleno (b) Trapezio scaleno (c) Trapezio isoscele (d) Trapezio rettangolo

Figura 9.17.: *Trapezi con evidenziate le altezze e le diagonali.*

9.3.4. Cenno ai poligoni concavi

Come già accennato, i problemi relativi ai poligoni concavi, in particolare i concetti di angolo interno ed esterno, sono molto più complessi che nel caso di poligoni convessi. Daremo qui solo un breve cenno del problema, rimandando a trattati più specialistici un approfondimento della questione (si può consultare, per esempio, [2]). Per trattare i poligoni concavi occorre innanzitutto modificare la definizione di poligono: seguiremo la definizione di L.Poinsot.

Definizione 9.33. *Si dice* poligono *la figura composta da n* (> 2) *punti* $\mathsf{A}_1, \mathsf{A}_2, \dots, \mathsf{A}_n$*, detti* vertici*, assunti ordinatamente nel piano e dai segmenti, detti* lati*, che congiungono il primo con il secondo, il secondo con il terzo, ..., l'ultimo con il primo. Se esistono punti diversi dai vertici nei quali concorrono due o più lati il poligono si dice* intrecciato*, altrimenti si dice ordinario.*

Nella figura 9.18 sono rappresentati alcuni pentagoni non convessi.

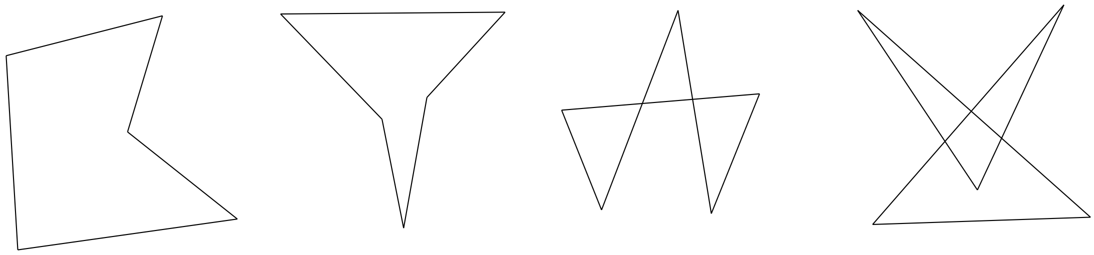


Figura 9.18.: *Pentagoni concavi: ordinari i primi due, intrecciati gli altri*

Per poter parlare di angoli interni ed esterni occorre innanzitutto orientare il piano: in senso puramente intuitivo parleremo di orientamento *orario* ed *antiorario*, assumendo come verso positivo uno dei due, per esempio quello antiorario. Orientato il piano si deve orientare il poligono, in uno dei due versi possibili. Per i poligoni ordinari si può scegliere, come noi faremo, di orientare anche il poligono in verso antiorario, mentre il concetto non ha più senso nel caso di poligoni intrecciati. Infine bisogna orientare gli angoli, immaginati come rotazione di una semiretta attorno alla sua origine: saranno positivi gli angoli orientati concordemente al piano, negativi gli altri.

Considerato ora un angolo del poligono, immaginiamo di far ruotare il secondo lato attorno al vertice in modo che descriva un angolo positivo (antiorario nella nostra convenzione) per sovrapporsi al primo lato: l'angolo descritto si chiamerà *angolo interno*. Si noti che, se si cambia il verso del poligono, ciascun angolo si cambia nel suo esplementare.

Nella figura 9.19 sono rappresentati gli angoli interni nei casi possibili per un quadrilatero (orientato in senso antiorario se è ordinario). Nei primi due casi la somma degli angoli interni è di un angolo giro, nel terzo caso di due angoli giri. Non si è dunque mantenuta con questa definizione la proprietà della somma degli angoli valida per un poligono convesso; tuttavia rimane valida la proprietà che la somma è data da un certo numero di angoli piatti e che rimane costante, per ogni tipo di poligono.

Con le stesse convenzioni di prima si dice angolo *esterno* l'angolo, minore di un angolo piatto, di cui deve ruotare il prolungamento del primo lato per sovrapporsi al secondo lato anche nel verso. In questo modo si mantiene una caratteristica valida per i poligoni convessi e cioè che la somma di un angolo

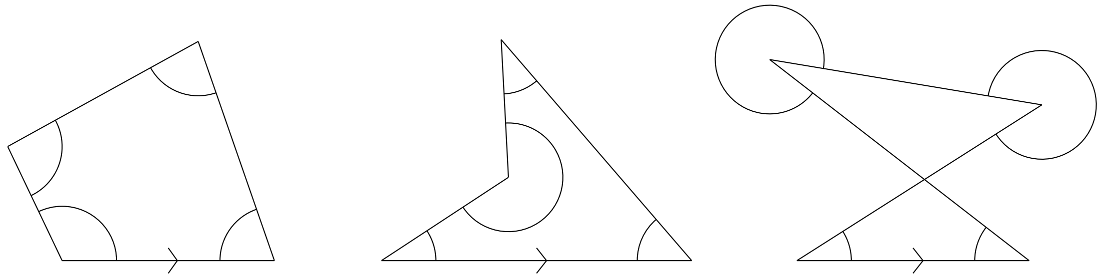


Figura 9.19.: *Angoli interni in un quadrilatero*

interno coll'angolo esterno ad esso adiacente è sempre un angolo piatto (se si tiene anche conto del fatto che alcuni angoli, quelli antiorari, sono positivi, altri, quelli orari, negativi).

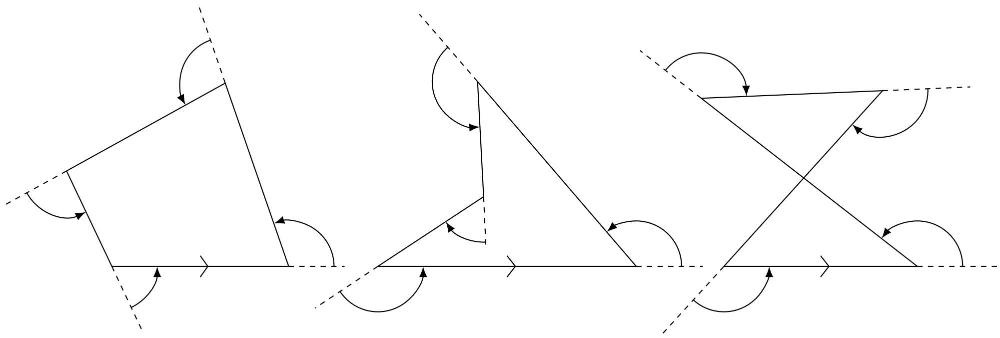


Figura 9.20.: *Angoli esterni in un quadrilatero*

Nelle figure 9.21, 9.22 e 9.23 sono rappresentate le situazioni possibili, relativamente agli angoli interni, per un pentagono.

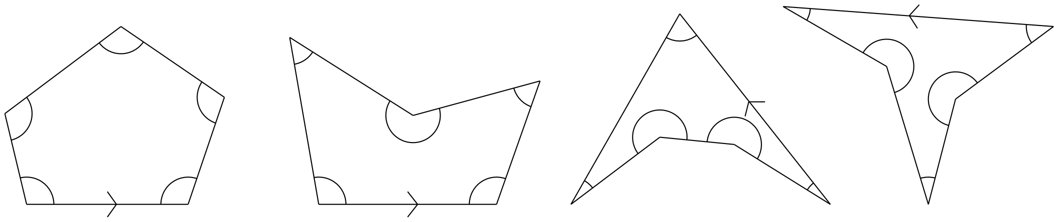


Figura 9.21.: *Angoli interni nei pentagoni*

Le situazioni diventano via via più complesse, man mano che cresce il numero dei lati: per esempio in un esagono si devono esaminare anche i casi in cui in un punto diverso dai vertici concorrono tre lati. Solo a titolo di curiosità una di queste situazioni è rappresentata nella figura 9.24, in cui sono rappresentati anche gli angoli interni.

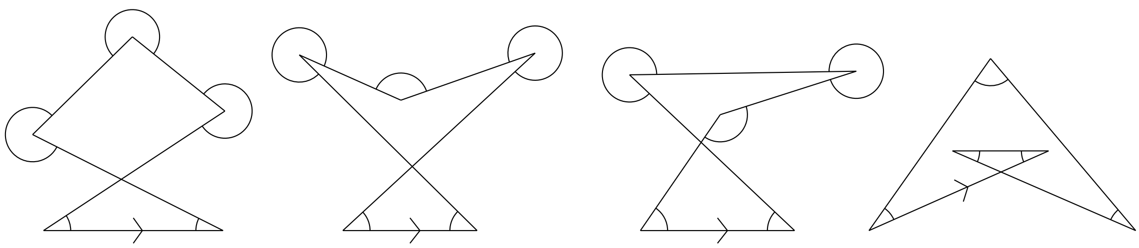


Figura 9.22.: *Ancora angoli interni nei pentagoni*

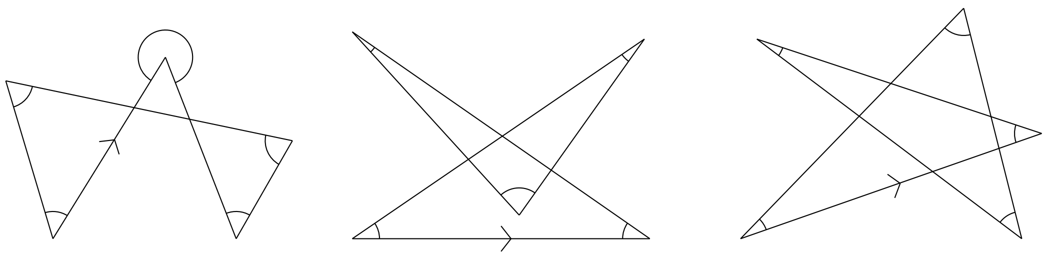


Figura 9.23.: *Ancora angoli interni nei pentagoni*

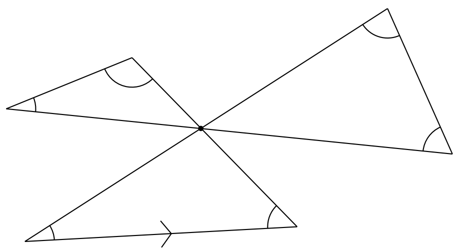


Figura 9.24.: *Un esempio di esagono intrecciato*

Questa semplice introduzione rende evidente il perché, nel trattare i poligoni, di solito ci si limita ai poligoni convessi.

9.4. La circonferenza e il cerchio

Definizione 9.34. *Si dice* circonferenza *o* circolo *di centro* O *e di raggio* r, *il luogo geometrico dei punti del piano che hanno distanza da* O *uguale ad* r.

Definizione 9.35. *Si dice* cerchio *di centro* O *e di raggio* r, *il luogo geometrico dei punti del piano che hanno distanza da* O *minore o uguale ad* r.

La circonferenza di centro O e di raggio r dunque appartiene al cerchio di stesso centro e raggio e ne costituisce il *contorno*.

Tutti i segmenti congiungenti O con un punto della circonferenza si dicono *raggi* e sono tutti tra di loro uguali. I punti del piano la cui distanza da O è minore di r si dicono *interni* alla circonferenza e al cerchio, i punti la cui distanza da O è maggiore di r si dicono *esterni* alla circonferenza e al cerchio. Il cerchio è una figura convessa, la circonferenza no.

Due cerchi o due circonferenze sono uguali se e solo se hanno lo stesso raggio.

9.4.1. Corde

Definizione 9.36. *Si chiama* corda *ogni segmento che unisce due punti di una circonferenza. Una corda passante per il centro si chiama* diametro.

Ogni diametro divide la circonferenza e il cerchio in due parti dette, rispettivamente, *semicirconferenza* e *semicerchio*. Tutti i diametri sono somma di due raggi e sono tra di loro uguali. Ogni diametro è maggiore di qualsiasi corda non passante per il centro.

L'asse di una corda passa per il centro e il diametro perpendicolare ad una corda la divide in due parti uguali.

Teorema 9.37. *Corde uguali di una stessa circonferenza distano ugualmente dal centro e, viceversa, corde che distano ugualmente dal centro sono uguali.*

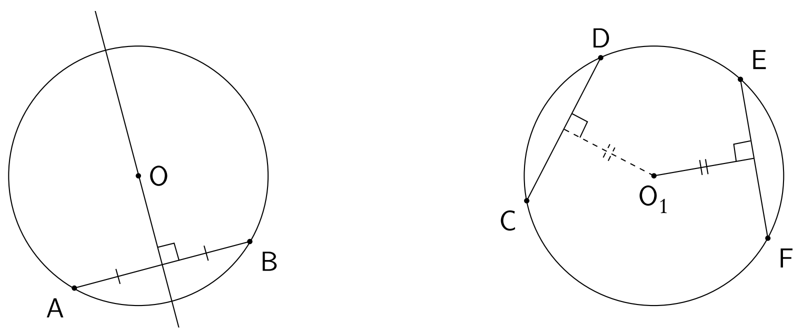


Figura 9.25.: *Proprietà delle corde di una circonferenza*

Teorema 9.38. *Per tre punti non allineati passa una e una sola circonferenza. Quindi due circonferenze distinte non possono avere più di due punti in comune e una circonferenza non può avere tre punti allineati.*

9.4.2. Posizioni relative di una retta e una circonferenza

Una retta la cui distanza dal centro è maggiore del raggio ha tutti i suoi punti esterni alla circonferenza e si dice *esterna alla circonferenza*. Viceversa una retta esterna a una circonferenza ha distanza dal centro maggiore del raggio.

Una retta la cui distanza dal centro è uguale al raggio ha un solo punto in comune con la circonferenza e si dice *tangente alla circonferenza*. Viceversa se una retta è tangente a una circonferenza, cioè ha un solo punto in comune con la circonferenza, la sua distanza dal centro è uguale al raggio. Il punto comune si dice *punto di tangenza* o *punto di contatto*. Una retta è tangente in P a una circonferenza di centro O se e solo se è perpendicolare al raggio $\overline{OP}$. Per un punto di una circonferenza si può condurre una sola tangente alla circonferenza stessa.

Una retta la cui distanza dal centro è minore del raggio ha due punti in comune con la circonferenza e si dice *secante alla circonferenza*. Viceversa se una retta è secante a una circonferenza ha due punti in comune con la circonferenza.

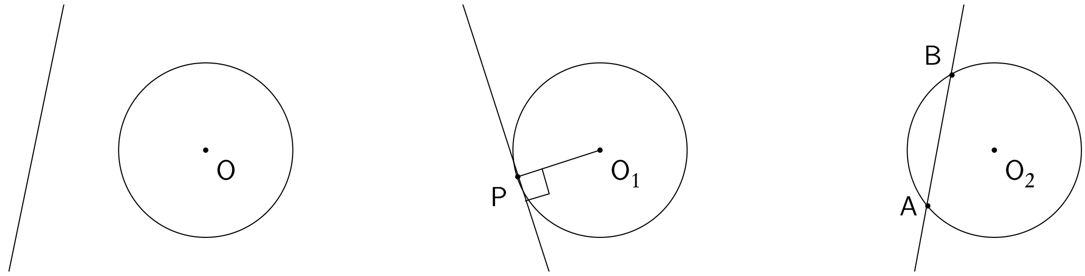


Figura 9.26.: *Retta esterna, tangente, secante a un circonferenza*

9.4.3. Parti della circonferenza e del cerchio

Si dice *angolo al centro* un angolo con il vertice nel centro di una circonferenza. Si dice poi *arco di circonferenza* l'intersezione di una circonferenza con un angolo al centro. La corda che unisce gli estremi di un arco si dice *sottesa* dall'arco. Dati un angolo al centro e l'arco da esso individuato si dice che l'angolo *insiste* sull'arco.

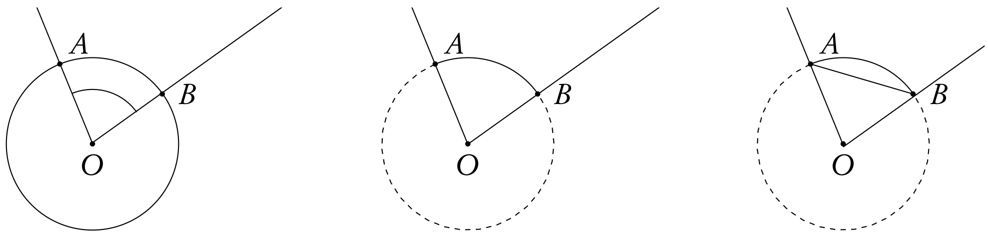


Figura 9.27.: *Angolo al centro, arco, corda sottesa*

Si dice *settore circolare* l'intersezione di un cerchio con un angolo al centro. Si dice *segmento circolare ad una base* l'intersezione di un cerchio e di un semipiano contenente una corda di una circonferenza. Si dice *segmento circolare a due basi* l'intersezione di un cerchio e di una striscia i cui lati contengono due corde parallele.

Nello stesso cerchio, o in cerchi uguali, ad archi uguali corrispondono angoli al centro uguali, settori uguali e corde uguali; ad angoli al centro uguali corrispondono archi uguali, settori uguali, corde uguali.

Definizione 9.39. *Si dice* angolo alla circonferenza *un angolo convesso con il vertice sulla circonferenza e i due lati entrambi secanti o uno secante e uno tangente alla circonferenza.*

Un angolo alla circonferenza individua sulla circonferenza un arco, precisamente quello contenuto nell'angolo e che ha gli estremi sui lati dell'angolo (angolo con i lati secanti) o uno su un lato e l'altro sul

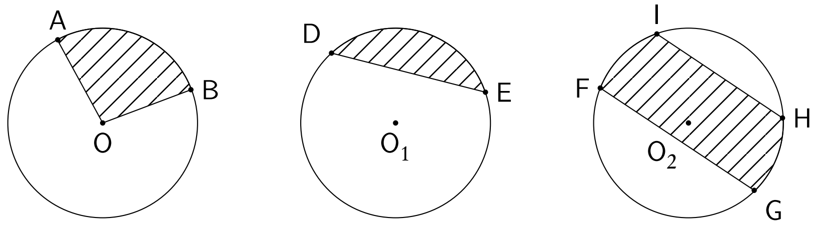


Figura 9.28.: *Settore circolare e segmenti circolari a una e due basi*

vertice (angolo con un lato secante e uno tangente) e si dice che l'angolo *insiste* sull'arco. Un angolo alla circonferenza che insiste su un dato arco si dice anche *inscritto* nell'arco di circonferenza rimanente. Un angolo alla circonferenza si dice *corrispondente* all'angolo al centro che insiste sullo stesso arco.

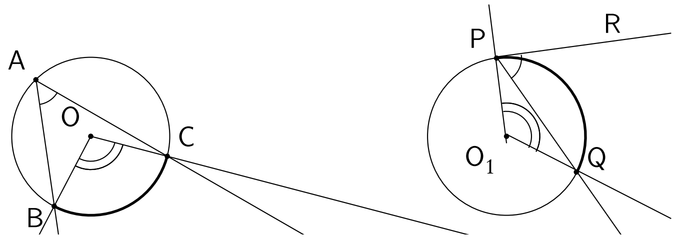


Figura 9.29.: *Angoli alla circonferenza, corrispondenti angoli al centro e archi su cui insistono*

Ad ogni angolo alla circonferenza corrisponde un unico angolo al centro, mentre ad ogni angolo al centro corrispondono infiniti angoli alla circonferenza. Vale il seguente importante teorema.

Teorema 9.40. *Un angolo al centro è doppio di ogni angolo alla circonferenza che insiste sullo stesso arco. Conseguentemente angoli alla circonferenza che insistono sullo stesso arco o su archi uguali sono uguali. Angoli alla circonferenza inscritti in una semicirconferenza sono retti.*

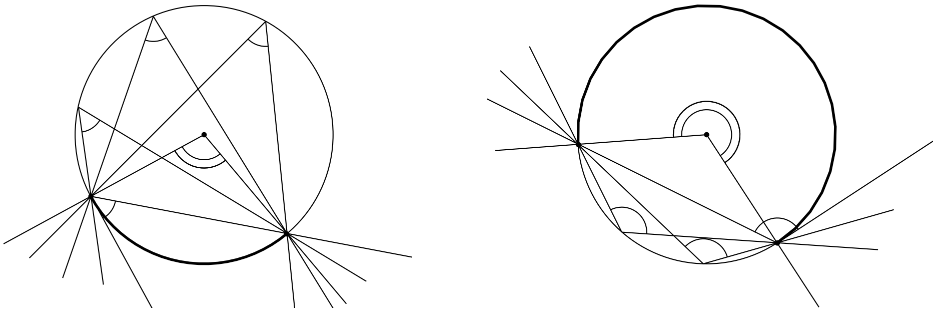


Figura 9.30.: *Proprietà degli angoli al centro e alla circonferenza*

9.4.4. Posizioni relative di due circonferenze

Due circonferenze distinte del piano possono avere le seguenti posizioni relative:

1. esterne una rispetto all'altra, o semplicemente *esterne*, quando ogni punto dell'una è esterno all'altra;
2. *esternamente tangenti*, quando hanno un solo punto in comune e ogni altro punto dell'una è esterno all'altra;
3. *secanti*, quando hanno solo due punti in comune;
4. *internamente tangenti*, quando, avendo raggi diversi, hanno un solo punto in comune e ogni altro punto di quella di raggio minore è interno a quella di raggio maggiore;
5. *interne una all'altra*, quando, avendo raggi diversi, non hanno nessun punto in comune e ogni punto di quella di raggio minore è interno a quella di raggio maggiore.

La condizione necessaria e sufficiente affinché due circonferenze risultino

1. esterne, è che la distanza dei loro centri sia maggiore della somma dei raggi;
2. esternamente tangenti è che la distanza dei loro centri sia uguale alla somma dei raggi;
3. secanti è che la distanza dei loro centri sia maggiore del modulo della differenza e minore della somma dei raggi;
4. internamente tangenti è che i raggi siano diversi e la distanza dei loro centri sia uguale al modulo della differenza dei raggi;
5. interne è che i raggi siano diversi e la distanza dei loro centri sia minore del modulo della differenza dei raggi.

Nell'ultimo caso, se la distanza dei centri è nulla (e le circonferenze hanno raggio diverso), esse si dicono *concentriche* e la regione piana compresa tra le due circonferenze si dice *corona circolare*.

Nel caso di circonferenze tangenti, internamente o esternamente, la perpendicolare alla retta passante per i centri, condotta dal punto di tangenza, è tangente comune alle due circonferenze. Nel caso di circonferenze secanti la retta per i due punti comuni è perpendicolare alla retta passante per i centri.

9.4.5. Tangenti a una circonferenza per un punto esterno

Abbiamo già osservato che per un punto di una circonferenza si può condurre una sola tangente alla circonferenza stessa. Risulta evidente dalle proprietà delle secanti che per un punto interno non si può condurre alcuna tangente alla circonferenza. Per un punto esterno vale il seguente teorema:

Teorema 9.41. *Per un punto esterno* P *a una circonferenza si possono condurre due e due sole tangenti alla circonferenza stessa.*

I segmenti che congiungono il punto P con i due punti A e B di contatto si dicono *segmenti tangenti* o *segmenti di tangenza* e sono uguali.

La semiretta PO che congiunge il punto P con il centro O della circonferenza è bisettrice dell'angolo $\mathsf{A}\widehat{\mathsf{P}}\mathsf{B}$. Inoltre la retta PO è asse del segmento $\overline{\mathsf{AB}}$.

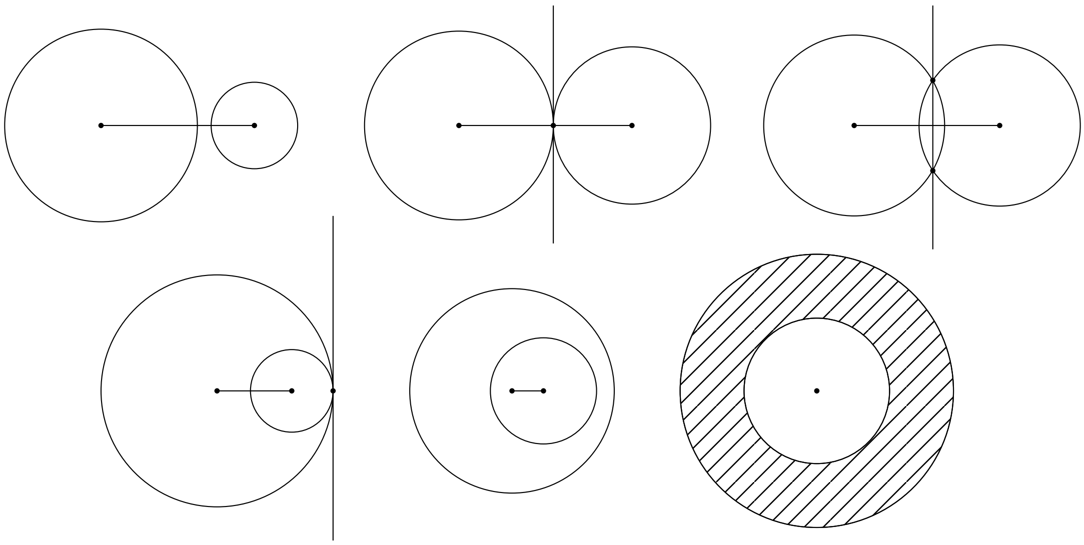


Figura 9.31.: *Circonferenze esterne, esternamente tangenti, secanti, internamente tangenti, interne, concentriche, con evidenziata la corona circolare*

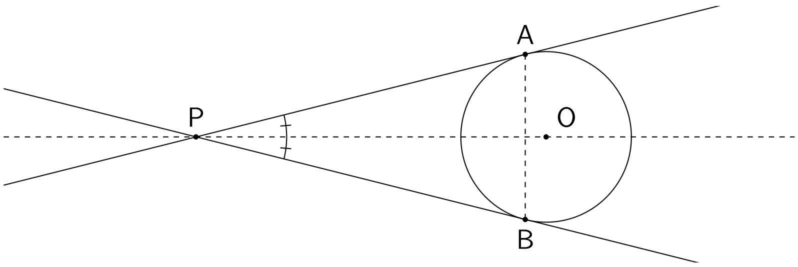


Figura 9.32.: *Tangenti a una circonferenza per un punto esterno*

9.4.6. Poligoni inscritti e circoscritti

Definizione 9.42. *Un poligono si dice* inscritto *in una circonferenza quando tutti i suoi vertici appartengono alla circonferenza; la circonferenza si dice allora* circoscritta *al poligono. Un poligono si dice* circoscritto *ad una circonferenza quando tutti i suoi lati sono tangenti alla circonferenza.*

I poligoni inscritti e circoscritti possono anche essere concavi o addirittura intrecciati, come mostra la figura 9.33: ci interesserà però principalmente il caso dei poligoni convessi.

Sappiamo già, vedi le figure 9.9 e 9.12, che in un triangolo si può sempre inscrivere una circonferenza il cui centro è il punto di intersezione delle bisettrici e che ad un triangolo si può sempre circoscrivere una circonferenza il cui centro è il punto di intersezione degli assi dei tre lati.

Nel caso del triangolo, oltre alla circonferenza inscritta si considerano anche tre circonferenze, dette *ex-inscritte*, e che sono tangenti a un lato e al prolungamento degli altri due. Esse hanno come centri i tre *ex-centri*, ottenuti come intersezione delle bisettrici di due angoli esterni e dell'angolo interno ad essi non adiacente. Complessivamente la circonferenza inscritta e le tre circonferenze ex-inscritte costituiscono

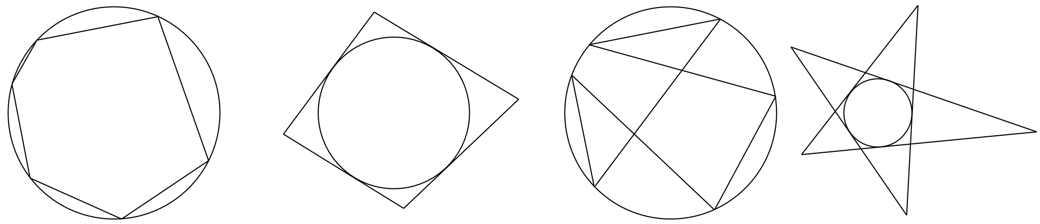


Figura 9.33.: *Poligoni inscritti e circoscritti, convessi (i primi due) e intrecciati (gli ultimi due)*

le quattro circonferenze tangenti a tre rette a due a due non parallele e non aventi un unico punto in comune.

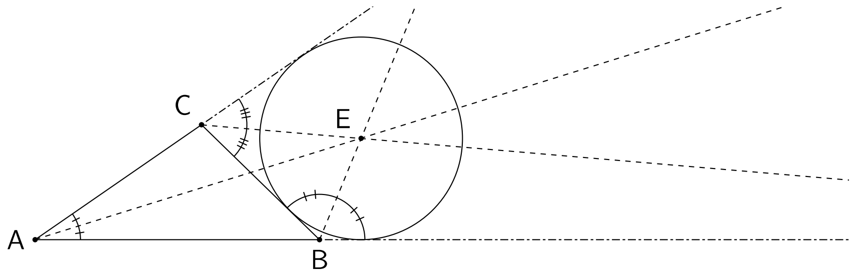


Figura 9.34.: *Ex-centro e circonferenza ex-inscritta in un triangolo*

Per i quadrilateri la situazione è completamente diversa: un quadrilatero arbitrario non è, di norma né inscrivibile né circoscrivibile ad una circonferenza. Valgono però i seguenti due teoremi, relativi ai quadrilateri convessi.

Teorema 9.43. *Un quadrilatero convesso è inscrivibile in una circonferenza se e solo se gli angoli opposti sono supplementari.*

Teorema 9.44. *Un quadrilatero convesso è circoscrivibile ad una circonferenza se e solo se la somma di due lati opposti è uguale alla somma degli altri due.*

9.4.7. Poligoni regolari

L'argomento "poligoni regolari" fa parte del capitolo sulla circonferenza e il cerchio, in quanto è strettamente legato al cosiddetto *problema della ciclotomia*, ovvero della divisione di una circonferenza in un dato numero di parti uguali, problema particolarmente studiato dai greci.

Definizione 9.45. *Un poligono (convesso) si dice* regolare *se ha tutti i lati e tutti gli angoli uguali.*

Si noti come mentre per un triangolo avere i lati uguali è condizione necessaria e sufficiente per avere gli angoli uguali, ciò non è più vero per poligoni con più di 3 lati: per esempi un rettangolo ha gli angoli ma non i lati uguali, un rombo ha i lati ma non gli angoli uguali.

Se una circonferenza è divisa in n archi uguali, il poligono che si ottiene congiungendo successivamente i punti di divisione è regolare (ed è inscritto nella circonferenza); il poligono individuato dalle tangenti alla circonferenza condotte nei punti di divisione è regolare (ed è circoscritto alla circonferenza).

Ogni poligono regolare è inscrivibile e circoscrivibile a una circonferenza. Il punto O, centro delle circonferenze inscritta e circoscritta, è detto *centro* del poligono; il raggio della circonferenza circoscritta è detto *raggio* del poligono; il raggio della circonferenza inscritta (che coincide con la distanza del centro da uno dei lati) è detto *apotema* del poligono.

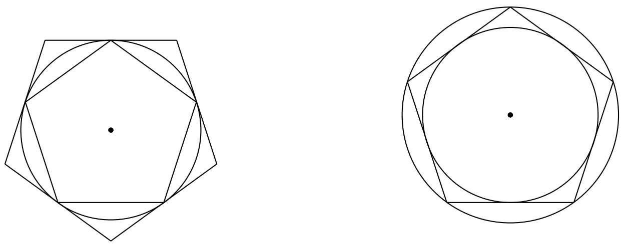


Figura 9.35.: *Pentagono regolare inscritto e circoscritto a una circonferenza (a sinistra). Circonferenze inscritta e circoscritta ad un pentagono regolare (a destra)*

È ovvio che, data una circonferenza, vi si possono iscrivere (o circoscrivere) infiniti poligoni con n lati; tuttavia tutti questi poligoni sono tra di loro uguali. Per questo motivo si parla *del* triangolo equilatero inscritto anziché *di un* triangolo equilatero inscritto, *del* quadrato inscritto, anziché *di un* quadrato inscritto, ecc.

Teorema 9.46. *Il lato dell'esagono regolare inscritto in una circonferenza è uguale al raggio della circonferenza.*

9.5. Equivalenze di superfici. Pitagora. Euclide

Ammettiamo come primitivi i concetti di superficie piana e di *estensione superficiale*.

Definizione 9.47. *Due superfici di uguale estensione si dicono* equivalenti.

L'equivalenza sostituisce, nel caso delle superfici, la sovrapponibilità: si confrontano tra di loro le estensioni di due superfici per stabilire se sono uguali o, in caso contrario, quale sia maggiore e, analogamente, si sommano e si sottraggono le estensioni di due superfici. Si ammette come postulato che l'equivalenza di superfici goda delle proprietà riflessiva, simmetrica e transitiva e che la somma goda delle proprietà commutativa e associativa. Inoltre somme e differenze di figure equivalenti sono equivalenti. Due figure scomponibili in figure ordinatamente equivalenti sono equivalenti.

9.5.1. Il caso dei poligoni

Nel caso dei poligoni lo studio dell'equivalenza può essere notevolmente semplificato, utilizzando il concetto di equiscomponibilità: due poligoni si dicono *equiscomponibili* quando possono essere decomposti in uno stesso numero di poligoni rispettivamente uguali. È ovvio che poligoni equiscomponibili sono equivalenti e anzi la quasi totalità dei teoremi sull'equivalenza dei poligoni si basa sulla verifica della equiscomponibilità. Elenchiamo i principali risultati, che saranno poi indispensabili per calcolare le aree dei poligoni.

— Un parallelogramma e un rettangolo, aventi basi ed altezze relative uguali, sono equivalenti.
— Due parallelogrammi, aventi uguali le basi e le altezze relative, sono equivalenti.
— Un triangolo è equivalente ad un parallelogramma che abbia per base la metà della base del triangolo e uguale altezza.
— Due triangoli di basi ed altezze uguali sono equivalenti.
— Un trapezio è equivalente ad un triangolo di uguale altezza e avente per base la somma delle basi del trapezio.

Teorema 9.48. *Un poligono si può trasformare in un altro equivalente, avente un lato di meno.*

Dimostrazione. Siano infatti A, B, C, D quattro vertici consecutivi del poligono e si tracci la diagonale $\overline{AC}$. Per D si tiri la parallela ad $\overline{AC}$. Sia E l'intersezione di BC con questa parallela.

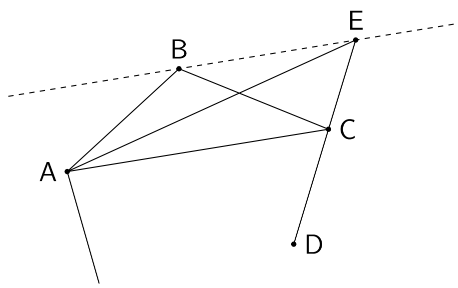


Figura 9.36.: *Trasformazione di un poligono*

I triangoli ABC e AEC sono equivalenti (stessa base e stessa altezza). Il poligono ottenuto dal poligono dato sostituendo i vertici A, B, C, D con A, E, D è equivalente al poligono dato ma ha un vertice e quindi un lato di meno. □

9.5.2. I teoremi di Pitagora ed Euclide

Anche i teoremi di Pitagora ed Euclide, probabilmente i teoremi più famosi della geometria, sono teoremi sull'equivalenza di superfici piane. Per questi teoremi, in particolare per il teorema di Pitagora, esistono decine di dimostrazioni diverse di cui le più interessanti sono, a nostro avviso, quelle basate sulla equiscomposizione.

Teorema 9.49 (Teorema di Pitagora)**.** *In un triangolo rettangolo il quadrato costruito sull'ipotenusa è equivalente alla somma dei quadrati costruiti sui cateti.*

Teorema 9.50 (Primo teorema di Euclide)**.** *In un triangolo rettangolo il quadrato costruito su un cateto è equivalente al rettangolo che ha per lati l'ipotenusa e la proiezione del cateto sull'ipotenusa.*

Teorema 9.51 (Secondo teorema di Euclide)**.** *In un triangolo rettangolo il quadrato costruito sull'altezza relativa all'ipotenusa è equivalente al rettangolo che ha per lati le proiezioni dei cateti sull'ipotenusa.*

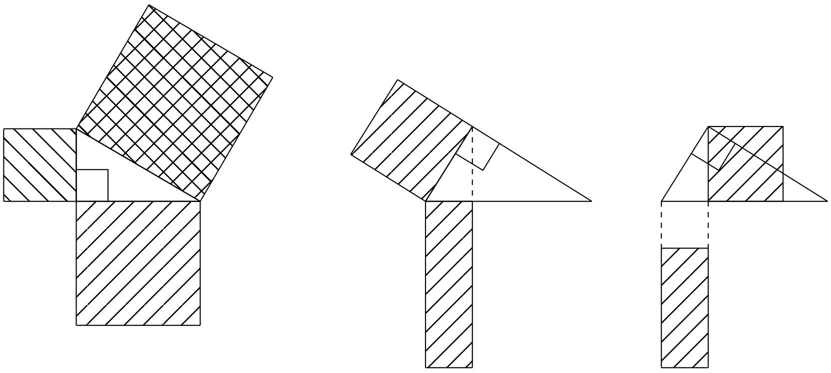


Figura 9.37.: *I teoremi di Pitagora ed Euclide*

9.6. Misura delle grandezze. Proporzionalità

Abbiamo supposto di disporre, fin dall'inizio, della possibilità di misurare le grandezze geometriche: ora presenteremo, molto sinteticamente, i principali concetti relativi.

Trattando i segmenti o gli angoli abbiamo visto che due di essi si possono confrontare tra di loro, in modo da riconoscere se sono uguali o, in caso contrario, quale dei due è maggiore. Per i segmenti e per gli angoli si sono poi introdotte le operazioni di addizione e sottrazione, e la somma gode delle proprietà commutativa e associativa. Analogo discorso per le estensioni delle superfici. Questo ci conduce alla seguente definizione formale.

Definizione 9.52. *Un insieme di figure costituisce una* classe di grandezze *se per esse è possibile definire una relazione di uguaglianza e di diseguaglianza, e se sono definite le operazioni di addizione e sottrazione, in modo che l'addizione goda delle proprietà commutativa e associativa. Devono inoltre valere le consuete relazioni che legano la somma all'uguaglianza e diseguaglianza. Grandezze di una stessa classe si dicono* omogenee.

I segmenti, gli angoli, le estensioni delle superfici costituiscono esempi di classi di grandezze omogenee.

Se una grandezza A è somma di n grandezze uguali alla grandezza B, diremo che A e *multiplo* di B secondo il numero n e scriveremo $A = nB$. Diremo anche che B è *sottomultiplo* di A e scriveremo $B = {}^1/_n A$. Avranno anche senso scritture del tipo $A = {}^m/_n B$ che si può leggere: A è multiplo secondo m del sottomultiplo secondo n di B. In questo caso

Valgono i seguenti due postulati: *Ogni grandezza geometrica è divisibile in un numero qualunque di parti uguali* (postulato della divisibilità); *date due grandezze omogenee e disuguali, esiste un multiplo della minore che supera la maggiore* (postulato di Eudosso-Archimede).

Due classi di grandezze omogenee si dicono *separate* se ogni elemento della prima è minore di ogni elemento della seconda. Se due classi di grandezze sono separate esiste almeno una grandezza maggiore o uguale a ogni elemento della prima classe e minore o uguale a ogni elemento della seconda: essa si chiama *elemento separatore* delle due classi (postulato di Dedekind o della continuità). Se poi, comunque si fissi una grandezza, esiste una grandezza della seconda classe e una grandezza della prima la cui differenza sia minore della grandezza data, le due classi si dicono *contigue* e in questo caso l'elemento separatore è *unico*. Si noti che nel caso delle superfici l'unicità dell'elemento separatore non è da intendersi nel senso che c'è un'unica superficie che funge da elemento separatore, ma un'unica estensione di superficie.

Date due grandezze omogenee A e B, se esistono m ed n tali che $A = {}^m\!/_n B$, diremo che le grandezze A e B sono fra loro *commensurabili*. La scrittura precedente si può anche mettere nella forma ${}^A\!/_m = {}^B\!/_n$: due grandezze sono commensurabili se hanno un sottomultiplo comune; essa si può anche scrivere come $nA = mB$: due grandezze sono commensurabili se hanno un multiplo comune. Scriveremo anche

$$\frac{A}{B} = \frac{m}{n} \quad \text{oppure} \quad A : B = \frac{m}{n},$$

e chiameremo il numero ${}^m\!/_n$ *rapporto* delle due grandezze commensurabili. Si noti che il rapporto così definito tra due grandezze commensurabili è un numero razionale (positivo).

Se, date due grandezze A e B, non esiste alcun sottomultiplo comune, le due grandezze si dicono *incommensurabili*. Per due grandezze incommensurabili la definizione di rapporto che abbiamo dato non è applicabile. Tuttavia, utilizzando le classi contigue di grandezze, è possibile estendere questa definizione, in maniera non dissimile da quella che ci ha condotto all'introduzione dei numeri irrazionali; non entreremo nei dettagli di questo processo, segnaliamo solo che si perviene alla conclusione che il rapporto di due grandezze incommensurabili è sempre espressa da un numero reale irrazionale. Il più famoso esempio (il primo della storia) di due grandezze incommensurabili è costituito dal lato e dalla diagonale di un quadrato: il rapporto della diagonale e del lato vale $\sqrt{2}$.

9.6.1. La misura delle grandezze

I risultati citati relativi al rapporto di due grandezze (omogenee) consentono di dare un preciso significato al concetto di *misura* di una data grandezza. Per ogni classe di grandezze si sceglie convenzionalmente una grandezza U, detta campione, e di ogni singola grandezza di quella classe si definisce *misura rispetto alla* U il rapporto $A : U$. Poiché la misura di U rispetto a se stessa, cioè il rapporto $U : U$ è uguale a 1, la U stessa si chiama *unità di misura*.

Il risultato fondamentale sulle misure è espressa dal seguente teorema, che esprime il fatto che esiste una corrispondenza biunivoca tra le grandezze di una qualsiasi classe e i numeri reali (positivi).

Teorema 9.53. *Ogni grandezza, di qualsiasi classe, ammette come misura rispetto a una prefissata unità un numero reale razionale o irrazionale a seconda che la grandezza stessa sia o no commensurabile con l'unità; viceversa scelto ad arbitrio un numero reale (positivo) esiste sempre una e una sola grandezza della classe che ha per misura quel numero reale rispetto a una prefissata unità, e la grandezza è commensurabile o incommensurabile con l'unità a seconda che il numero reale sia razionale o irrazionale.*

Le proprietà della misura di grandezze sono elencate nei teoremi che seguono.

Teorema 9.54. *Date due grandezze omogenee A e B, e considerate le loro misure rispetto ad una unità qualsiasi, se A è maggiore, minore o uguale a B, anche la misura di A è rispettivamente maggiore, minore o uguale a quella di B; viceversa se la misura di A è maggiore, minore o uguale a quella di B, anche la grandezza A è rispettivamente maggiore, minore o uguale alla grandezza B.*

Teorema 9.55. *La somma e la differenza di due grandezze omogenee hanno per misura la somma e la differenza delle rispettive misure rispetto a una prefissata unità.*

Teorema 9.56. *Il rapporto di due grandezze omogenee ha per misura il quoziente delle rispettive misure rispetto a una prefissata unità.*

Nei casi particolari, e particolarmente importanti, dei segmenti, degli angoli e delle (estensioni delle) superfici, le misure prendono nomi particolari; precisamente si chiamano *lunghezze* le misure dei segmenti, *ampiezze* le misure degli angoli, *aree* le misure delle estensioni superficiali. Altre misure si incontreranno nello studio della geometria dello spazio. Per le lunghezze l'unità di misura è, di norma, il *metro*, con i suoi multipli e sottomultipli decimali; per gli angoli l'unità di misura è il *grado*, o meglio il *grado sessagesimale*, che è un angolo pari alla 360ª parte dell'angolo giro, con i suoi sottomultipli sessagesimali, il *primo* (60ª parte del grado) e *secondo* (60ª parte del primo), e successivamente con sottomultipli decimali; per le aree l'unità è il *metro quadrato*, con i suoi multipli e sottomultipli secondo 100.

9.6.2. Proporzionalità fra grandezze

Date 4 grandezze A, B, C e D, a due a due omogenee, se il rapporto delle prime due è uguale al rapporto delle seconde due, si dice che le grandezze formano una *proporzione* e si scrive

$$A:B=C:D,$$

che si legge "A sta a B, come C sta a D". Le coppie (A,B) e (C,D) si possono anche chiamare coppie di grandezze proporzionali.

Si usa, per le proporzioni tra grandezze, tutta la terminologia già nota per le proporzioni numeriche: termini della proporzione, estremi e medi, antecedente e conseguente.

Il seguente teorema permette di estendere alla proporzionalità fra grandezze le proprietà delle proporzioni numeriche.

Teorema 9.57. *Condizione necessaria e sufficiente affinché quattro grandezze, a due a due omogenee, siano in proporzione, è che siano in proporzione le loro misure.*

Tra le proprietà delle proporzioni segnaliamo in particolare quella espressa dal seguente teorema.

Teorema 9.58. *Date tre grandezze A, B e C, le prime due omogenee, esiste una ed una sola grandezza X, omogenea con C, tale che valga la proporzione $A:B=C:X$.*

Consideriamo ora due classi di grandezze $A, B, \dots$ e $A', B', \dots$, tali che quelle di ciascuna classe siano tra di loro omogenee e che esista una corrispondenza biunivoca tra le due classi. Le due classi si dicono *direttamente proporzionali* o, semplicemente, *proporzionali*, se, prese due qualunque grandezze A e B della prima classe e considerate le corrispondenti A' e B' della seconda, si ha

$$A:B=A':B'.$$

Un esempio di classi di grandezze direttamente proporzionali è data dai rettangoli di data altezza e dalle rispettive basi o, di data base e dalle rispettive altezze: raddoppiando la base di un rettangolo, a parità di altezza, raddoppia l'area.

Nelle stesse condizioni di prima se si ha invece

$$A:B=B':A',$$

le due classi di grandezze si dicono *inversamente proporzionali*.

Un esempio di classi di grandezze inversamente proporzionali è data dalle basi e altezze dei rettangoli aventi la stessa area: raddoppiando la base di un rettangolo, a parità di area, dimezza l'altezza.

Tenendo conto della definizione di proporzionalità, i teoremi di Euclide si possono enunciare come segue.

1° teorema di Euclide In un triangolo rettangolo un cateto è medio proporzionale fra l'ipotenusa e la proiezione del cateto sull'ipotenusa.

2° teorema di Euclide In un triangolo rettangolo l'altezza relativa all'ipotenusa è media proporzionale tra le proiezioni dei cateti sull'ipotenusa.

Vale anche il seguente teorema per gli archi di circonferenza e i settori circolari.

Teorema 9.59. *Gli archi di una circonferenza, o di circonferenze uguali, i corrispondenti angoli al centro e i corrispondenti settori circolari formano tre classi di grandezze, a due a due direttamente proporzionali.*

9.6.3. Aree dei poligoni

Avendo ora a disposizione la misura delle grandezze, possiamo richiamare le formule per le aree dei principali poligoni, tenendo conto di quanto già noto a proposito dell'equivalenza.

Area del rettangolo L'area di un rettangolo è uguale al prodotto della lunghezza della base per la lunghezza dell'altezza.

Area del quadrato L'area di un quadrato è uguale al quadrato della lunghezza del suo lato.

Area del parallelogramma L'area di un parallelogramma è uguale al prodotto della lunghezza della base per la lunghezza dell'altezza.

Area del triangolo L'area di un triangolo è uguale al semiprodotto della lunghezza della base per la lunghezza dell'altezza.

Area del trapezio L'area di un trapezio è uguale al semiprodotto della somma delle lunghezze delle basi per la lunghezza dell'altezza.

Area di un poligono regolare L'area di un poligono regolare è uguale al semiprodotto della lunghezza del perimetro per la lunghezza dell'apotema.

9.6.4. Misure di circonferenza e cerchio

Riterremo come primitivo il concetto di *linea piana*, concetto la cui idea ci può venire fornita, per esempio, da un filo sottilissimo comunque disteso su un piano, o dal contorno di una superficie piana qualunque. Segnaliamo comunque che questo concetto è, in molti casi, tutt'altro che intuitivo, come mostra, tra l'altro, la famosa "curva di Hilbert", che riempie un intero quadrato. Anche le rette e i segmenti sono linee: le linee diverse dai segmenti sono spesso chiamate linee curve o, semplicemente, *curve*.

Anche il concetto di *estensione lineare* è un concetto primitivo, e, ancora una volta, la sua idea ci può venire fornita dal filo sottilissimo che immaginiamo di tendere in modo da fargli assumere la forma di un segmento (ma per esempio la curva di Hilbert sopra citata, e non solo, avrebbe un'estensione infinita!).

Due linee aventi uguale estensione si dicono *equivalenti*; il segmento equivalente ad una linea (che abbia estensione finita) si dice *linea rettificata*. La *lunghezza* di una linea è la lunghezza della linea rettificata.

La più importante curva della geometria euclidea piana è la circonferenza: la determinazione della sua lunghezza è tutt'altro che semplice e solo a seguito di un articolo di C.L.F. Lindemann del 1882, dopo oltre due millenni di tentativi, si è riusciti definitivamente a dimostrare che non è possibile costruire con riga e compasso un segmento equivalente ad una circonferenza. Strettamente connesso è il problema di determinare l'area del cerchio, anzi, storicamente, è stato proprio il problema della "quadratura del cerchio" a risultare più noto[4] e famoso.

La determinazione della lunghezza della circonferenza e dell'area del cerchio si basa sui seguenti teoremi.

Teorema 9.60. *La circonferenza rettificata è minore del perimetro di un qualunque poligono regolare circoscritto e maggiore del perimetro di una qualunque poligono regolare inscritto. Un cerchio è minore di ogni poligono regolare circoscritto e maggiore di ogni poligono regolare inscritto.*

Teorema 9.61. *Prefissato ad arbitrio un segmento σ, è possibile determinare due poligoni regolari, uno circoscritto e uno inscritto in una data circonferenza, tali che la differenza dei loro perimetri sia minore di σ. Ne consegue che la circonferenza rettificata è l'unico segmento maggiore del perimetro di un qualunque poligono regolare inscritto e minore del perimetro di un qualunque poligono regolare circoscritto ad una data circonferenza. Prefissata ad arbitrio una superficie Σ è possibile determinare due poligoni regolari, uno circoscritto e uno inscritto in un dato cerchio, tali che la loro differenza sia minore di Σ. Ne consegue che il cerchio è l'unica superficie maggiore di un qualunque poligono regolare inscritto e minore di un qualunque poligono regolare circoscritto ad una dato cerchio.*

Teorema 9.62. *Il rapporto di una circonferenza rettificata al suo diametro è costante. Tale rapporto si indica con* π (pi greco).

Teorema 9.63. *Un cerchio è equivalente ad un triangolo avente per base la circonferenza rettificata e per altezza il raggio.*

Dunque, detta c la lunghezza di una circonferenza, r la lunghezza del suo raggio ed $\mathcal{A}$ l'area del cerchio, si ha

$$c = 2\pi r, \quad \mathcal{A} = \pi r^2. \tag{9.1}$$

Storicamente una delle prime valutazioni di π risale ai babilonesi, che utilizzavano l'approssimazione $\pi \simeq 3$. Molto migliore la valutazione degli egiziani che giunsero all'approssimazione $\pi \simeq 3.16$. Nell'antichità la migliore valutazione è dovuta ad Archimede che, usando un poligono di 96 lati, giunse al risultato $3.1408 < \pi < 3.1429$. Oggi si conosce il valore di π con un'enormità di cifre decimali esatte. Fino alla decima si ha:

$$\pi = 3,1415926535\dots \tag{9.2}$$

Per quanto riguarda la lunghezza di un arco e l'area di un settore circolare basta ricordare il teorema 9.59 per determinarne le misure, una volta che sia noto il corrispondente angolo al centro.

[4] Il problema è citato anche da Dante, nel Paradiso, canto XXXIII, versi 133 — 135: Qual'è il geomètra che tutto s'affige per misurar lo cerchio, e non ritrova, pensando, quel principio ond'elli indige,...

9.7. La similitudine

Teorema 9.64 (Teorema di Talete)**.** *Un fascio di rette parallele determina sopra due trasversali due classi di segmenti proporzionali.*

Con riferimento alla figura 9.38, e tenendo conto che in questo caso le due classi sono costituite da grandezze omogenee, si ha

$$\overline{\mathsf{AB}} : \overline{\mathsf{A'B'}} = \overline{\mathsf{AC}} : \overline{\mathsf{A'C'}} = \overline{\mathsf{BC}} : \overline{\mathsf{B'C'}} = \overline{\mathsf{CD}} : \overline{\mathsf{C'D'}} = \dots$$

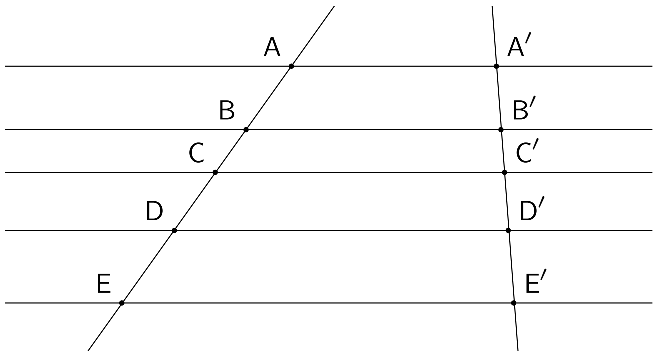


Figura 9.38.: *Teorema di Talete*

Conseguenza immediata di questo teorema è il seguente.

Teorema 9.65. *Una retta parallela al lato di un triangolo determina sugli altri due lati, o sui loro prolungamenti, segmenti proporzionali; viceversa se una retta determina su due lati di un triangolo, o sui loro prolungamenti, segmenti proporzionali, è parallela la terzo lato.*

Definizione 9.66 (Poligoni simili)**.** *Due poligoni con lo stesso numero di lati, i cui vertici siano presi in un conveniente ordine e che abbiano ordinatamente uguali gli angoli e proporzionali i lati, si dicono* simili.

Il rapporto costante k di due lati omologhi si chiama *rapporto di similitudine*.

In due poligoni simili si dicono vertici *omologhi* i vertici di angoli uguali, lati *omologhi* i lati che hanno per vertici due coppie di vertici omologhi. È ovvio che due poligoni uguali sono anche simili, con rapporto di similitudine 1.

Per i triangoli la definizione 9.66 è sovrabbondante in quanto si dimostra che se due triangoli hanno gli angoli ordinatamente uguali, hanno anche i lati in proporzione; viceversa se hanno i lati ordinatamente proporzionali hanno anche gli angoli uguali. Si dimostra anche che se due triangoli hanno un angolo uguale e i lati che lo comprendono in proporzione, allora hanno gli angoli uguali e anche il terzo lato in proporzione. Tutti questi risultati si sogliono compendiare nei seguenti criteri di similitudine dei triangoli.

1° criterio di similitudine dei triangoli Se due triangoli hanno due angoli ordinatamente uguali, e quindi tutti gli angoli uguali, sono simili.

2° criterio di similitudine dei triangoli Se due triangoli hanno un angolo uguale e i due lati che lo comprendono in proporzione sono simili.

3° criterio di similitudine dei triangoli Se due triangoli hanno i lati ordinatamente proporzionali sono simili.

Valgono le seguenti proprietà per i triangoli e i poligoni simili.

- In due triangoli simili le altezze corrispondenti a due lati omologhi sono proporzionali a questi lati.
- Due poligoni regolari di ugual numero di lati sono simili.
- Due poligoni simili vengono decomposti dalle diagonali condotte da due vertici omologhi in triangoli simili.
- I perimetri di due poligoni simili hanno come rapporto il rapporto di similitudine.
- I perimetri di due poligoni regolari dello stesso numero di lati stanno fra loro come i rispettivi raggi e le rispettive apoteme.
- Due poligoni simili stanno fra loro come i quadrati costruiti su due lati omologhi.

9.7.1. Applicazioni della similitudine

Teorema 9.67 (Teorema della bisettrice dell'angolo interno)**.** *La bisettrice di un angolo interno di un triangolo divide il lato opposto in parti proporzionali agli altri due lati.*

Teorema 9.68 (Teorema della bisettrice dell'angolo esterno)**.** *La bisettrice di un angolo esterno di un triangolo incontra il prolungamento del lato opposto, quando non gli è parallela, in un punto, le cui distanze dagli estremi di quel lato sono proporzionali agli altri due lati.*

Con riferimento alla figura 9.39, si ha, sia per l'angolo interno che per l'angolo esterno,

$$\overline{\mathsf{BD}} : \overline{\mathsf{DC}} = \overline{\mathsf{AB}} : \overline{\mathsf{AC}}.$$

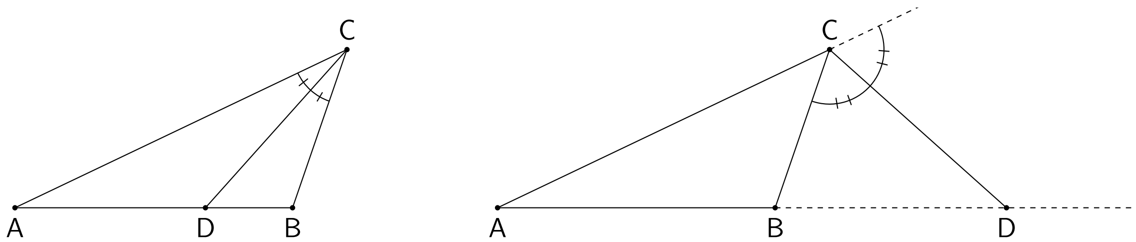


Figura 9.39.: *Teoremi della bisettrice dell'angolo interno ed esterno*

Teorema 9.69 (Teorema delle due corde)**.** *Condotte da un punto interno ad una circonferenza due corde, esse si dividono in modo che le due parti dell'una formano i medi, le due parti dell'altra gli estremi di una proporzione.*

I nomi "segmento di secante" e " segmento di tangente" sono usati nei successivi due teoremi in un senso che appare chiaro dalla figura 9.40.

Teorema 9.70 (Teorema delle due secanti)**.** *Condotte da un punto esterno ad una circonferenza due secanti, si ha che un intero segmento di secante e la sua parte esterna formano i medi e l'altro intero segmento di secante e la sua parte esterna gli estremi di una proporzione.*

Teorema 9.71 (Teorema della secante e della tangente)**.** *Condotte da un punto esterno ad una circonferenza una secante e una tangente, il segmento di tangente è medio proporzionale tra l'intero segmento di secante e la sua parte esterna.*

Con riferimento alla figura 9.40 si ha

$$\overline{AC} : \overline{AD} = \overline{AE} : \overline{AB},$$

per i primi due casi, e

$$\overline{AC} : \overline{AD} = \overline{AD} : \overline{AB},$$

per il terzo caso. Il terzo caso può essere considerato come caso particolare del secondo:la secante AD si è spostata fino a diventare tangente, di modo che i punti D e E diventano coincidenti.

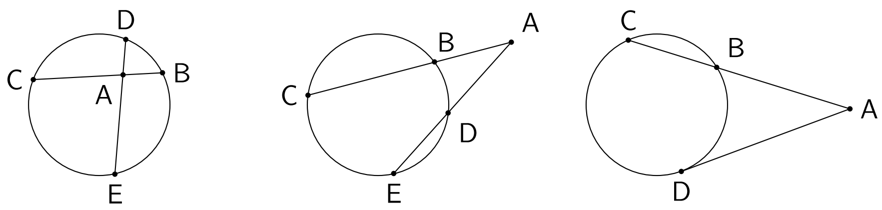


Figura 9.40.: *Teoremi delle due corde, delle due secanti, della secante e della tangente*

Definizione 9.72 (Parte aurea di un segmento)**.** *Si dice* parte aurea *o* sezione aurea *di un segmento la parte di esso che è media proporzionale fra l'intero segmento e la parte restante.*

$$\overline{AB} : \overline{AC} = \overline{AC} : \overline{CB}$$

Figura 9.41.: *Sezione aurea di un segmento*

Teorema 9.73. *Il lato del decagono regolare inscritto in un cerchio è la sezione aurea del raggio.*

9.8. Costruzioni con riga e compasso

Le costruzioni con riga e compasso sono state un problema chiave della matematica fin dal tempo dei greci. Eseguire una costruzione con questi strumenti vuol dire determinare oggetti geometrici a partire da altri oggetti dati, utilizzando solo la riga e il compasso. Con *riga* si intende solo uno strumento in grado di tracciare una retta, a partire da due suoi punti, e dunque non una "riga graduata" in grado di prendere misure o segnare distanze. Con *compasso* intendiamo uno strumento in grado di tracciare una circonferenza avente un centro assegnato e passante per un punto e anche uno strumento in grado di rilevare una certa lunghezza (attenzione: *non* misurare misurare una lunghezza, ma semplicemente rilevarla da un segmento già costruito) e poterla trasportare rigidamente in modo da costruire una circonferenza avente quel raggio e un opportuno centro. Come per tutte le figure geometriche non è importante il disegno costruito nella pratica su un foglio di carta, quanto la correttezza del procedimento seguito: i punti non dovrebbero avere dimensioni, le linee non dovrebbero avere spessore.

I problemi di costruzione più famosi, direttamente tramandatici dai greci, sono quelli della *duplicazione del cubo* (costruire un cubo che abbia volume doppio rispetto a un cubo assegnato), della *trisezione di un angolo generico* (dividere in tre parti uguali un dato angolo), della *quadratura di un cerchio* (costruire un quadrato equivalente a cerchio dato), ai quali va aggiunto il problema cosiddetto della *ciclotomia*, ovvero della divisione di una circonferenza in un numero n di parti uguali, problema sostanzialmente equivalente a quello di costruire un poligono regolare di n lati. I problemi della duplicazione del cubo e della quadratura del cerchio sono sempre irresolubili con riga e compasso, quello della trisezione di un angolo è risolubile solo per alcuni pochi angoli particolari, quello della ciclotomia solo per alcuni valori di n, da cui sono esclusi per esempio il caso $n = 7$ e $n = 9$. Sorprende il fatto che in molti testi di disegno tecnico siano proposte le costruzioni di un eptagono e di un ennagono con riga e compasso, senza precisare che tali costruzioni sono necessariamente approssimate.

In questo paragrafo proporremo alcune delle costruzioni più importanti, rimandando a testi specialistici una trattazione più dettagliata (vedi, per esempio, [7]). È un utile esercizio dimostrare la correttezza delle costruzioni proposte, utilizzando i teoremi presentati.

Dividere per metà un segmento

Dato un segmento $\overline{\mathsf{AB}}$, si faccia centro prima in A e poi in B e si costruiscano due circonferenze aventi lo stesso raggio, superiore alla metà del segmento dato: le circonferenze si intersecano in C e D. La retta CD incontra $\overline{\mathsf{AB}}$ in un punto M che è punto medio del segmento dato.

La retta CD è anche l'asse del segmento dato. Si veda la figura 9.42.

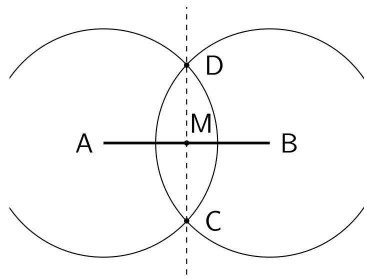


Figura 9.42.: *Divisione di un segmento a metà*

Dividere un angolo a metà

Sia $\mathsf{A\hat{B}C}$ l'angolo dato. Con centro in B e apertura a piacere si descriva un arco che intersechi i lati dell'angolo in E e D. Con la medesima apertura di compasso si traccino due archi di centri E e D che si intersecano in F ($\mathsf{F} \neq \mathsf{B}$). BF è la bisettrice dell'angolo dato e lo divide a metà. Si veda la figura 9.43.

Condurre la perpendicolare ad un retta da un punto

Sia che il punto P stia sulla retta o ne sia esterno, con centro in P si tracci una circonferenza che individui sulla retta due punti A e B: la perpendicolare richiesta è l'asse del segmento $\overline{\mathsf{AB}}$.

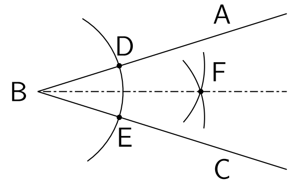


Figura 9.43.: *Divisione di un angolo a metà*

Condurre la parallela ad una retta per un punto

Sia r la retta e P un punto ad essa esterno. Con centro in un punto A di r e raggio $\overline{AP}$ si tracci un arco che intersechi in B la retta r. Con centro in P e in B e la stessa apertura di compasso si descrivano due archi che si intersecano in C. La retta PC è la parallela richiesta. Si veda la figura 9.44.

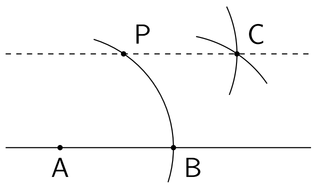


Figura 9.44.: *Costruzione della parallela a una retta per un punto*

Dividere a metà un arco di circonferenza

Se $\overset{\frown}{ABC}$ è l'arco, basta tracciare l'asse del segmento (corda) $\overline{AC}$.

Costruire la circonferenza passante per 3 punti

Se A, B, C sono i tre punti (distinti e non allineati), si tracciano gli assi di due dei tre segmenti $\overline{AB}$, $\overline{BC}$ e $\overline{AC}$: il loro punto di intersezione è il centro della circonferenza cercata.

Tangenti a una circonferenza

Sia C il centro della circonferenza data e P un punto esterno. Con centro nel punto medio M di $\overline{PC}$ si costruisca la circonferenza di raggio $\overline{CM}$: i punti A e B di intersezione delle due circonferenze sono i punti di contatto delle tangenti richieste.

Se poi il punto P appartiene alla circonferenza, basta tracciare la perpendicolare per P alla retta CP. Si veda la figura 9.45.

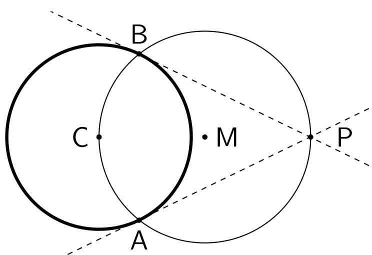


Figura 9.45.: *Tangenti a una circonferenza da un punto esterno*

Luogo dei punti che vedono un segmento sotto un dato angolo

Siano $\overline{AB}$ il segmento e α l'angolo dati. Si trasporti l'angolo in modo che il suo vertice coincida con A e un suo lato con la semiretta AB; sia AC il secondo lato dell'angolo trasportato. Per A si tracci la perpendicolare ad AC e sia O la sua intersezione con l'asse di $\overline{AB}$. L'arco di centro O ed estremi A e B unitamente al suo simmetrico rispetto al segmento $\overline{AB}$ costituiscono il luogo richiesto. Si veda la figura 9.46.

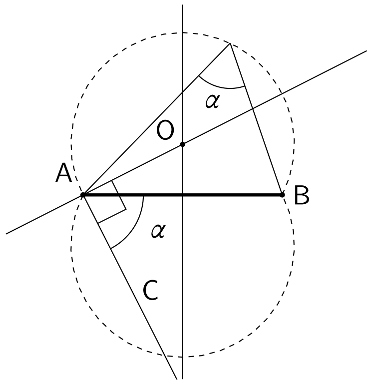


Figura 9.46.: *Punti che vedono un segmento sotto un dato angolo*

Triangolo equilatero, quadrato, esagono regolare inscritti in una circonferenza

Per inscrivere un quadrato in una circonferenza basta tracciare due diametri tra di loro perpendicolari; per inscrivere un esagono basta ricordare che il suo lato è uguale al raggio della circonferenza; per inscrivere un triangolo equilatero basta unire tre vertici, a due due non consecutivi, di un esagono regolare.

Costruiti il quadrato e l'esagono, per successive divisioni a metà dei vari archi si possono costruire i poligoni di 8, 16, 32, ..., 12, 24, ... lati. Si veda la figura 9.47.

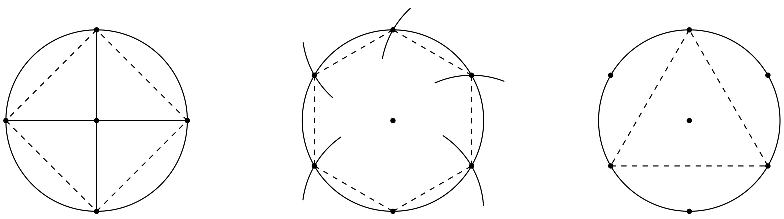


Figura 9.47.: *Quadrato, esagono regolare, triangolo equilatero inscritti*

Quadratura di un poligono

Poiché ogni poligono si può trasformare in un poligono con un lato in meno, vedi la costruzione della figura 9.36, ripetendo successivamente questa costruzione si può trasformare un poligono in un triangolo ad esso equivalente. Il triangolo può successivamente essere trasformato in un rettangolo che abbia la stessa altezza e metà base. Successivamente il rettangolo si può trasformare in un quadrato equivalente ricorrendo al primo o al secondo teorema di Euclide: scegliamo qui il secondo. Se ABCD è il rettangolo, di base $\overline{AB}$ e altezza $\overline{BC}$, sul prolungamento di $\overline{AB}$ si prenda il segmento $\overline{BE} = \overline{BC}$. Con centro nel punto medio F di $\overline{AE}$ si costruisca una semicirconferenza di diametro $\overline{AE}$. Tirata da B la perpendicolare ad AE, sia G il suo punto di intersezione con la semicirconferenza. Il triangolo AGE è rettangolo in G e $\overline{BG}$ ne è l'altezza relativa all'ipotenusa: il quadrato costruito su di essa è equivalente al rettangolo dato. Vedi la figura 9.48.

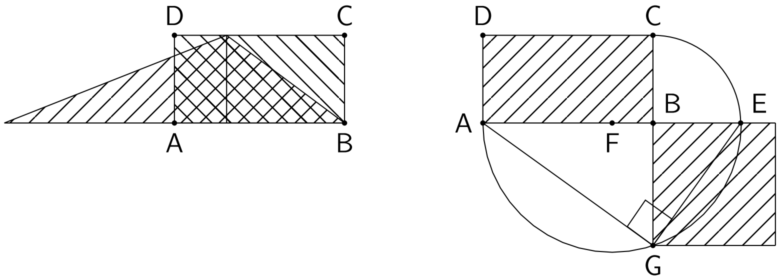


Figura 9.48.: *Quadratura di un triangolo*

Dividere un segmento in parti proporzionali a più segmenti dati, o in parti uguali

Sia $\overline{AB}$ il segmento dato e siano m, n, p gli altri segmenti. Si tracci una semiretta qualsiasi, ma non passante per B, di origine A e su di essa, a partire da A si riportino consecutivamente i segmenti m, n, p in $\overline{AC}$, $\overline{CD}$, $\overline{DE}$. Si tracci poi la retta DE e le parallele ad essa per C e D, che incontrano in C′ e D′ il segmento $\overline{AB}$: i punti C′ e D′ risolvono il problema. Se i segmenti m, n, p sono tutti uguali, il segmento $\overline{AB}$ risulta diviso in 3 parti uguali. Analoga costruzione per un numero qualunque di segmenti. Si veda la figura 9.49.

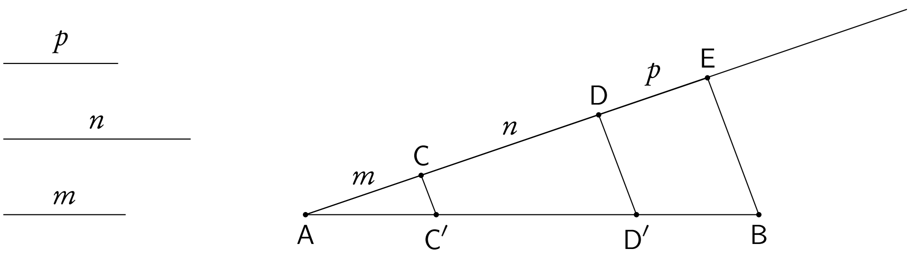


Figura 9.49.: *Divisione di un segmento*

Costruire il quarto proporzionale dopo tre segmenti dati

Siano m, n, p i segmenti dati. Si consideri an angolo acuto $A\widehat{B}C$ qualsiasi. Su lato BA, a partire da B si riportino consecutivamente m ed n in $\overline{BD}$ e $\overline{DE}$; sull'altro lato BC, a partire da B, si riporti il segmento p in $\overline{BF}$. Congiunto F con D si tracci per E la parallela a FD, che incontra BC in G: il segmento $\overline{FG}$ risolve il problema. Si veda la figura 9.50.

Costruzione del medio proporzionale

Dati due segmenti m ed n, li si riporti consecutivamente su una retta in $\overline{AB}$ e $\overline{BC}$. Costruita una semicirconferenza di diametro $\overline{AC}$, si tracci per B la perpendicolare a $\overline{AC}$, e sia D la sua intersezione

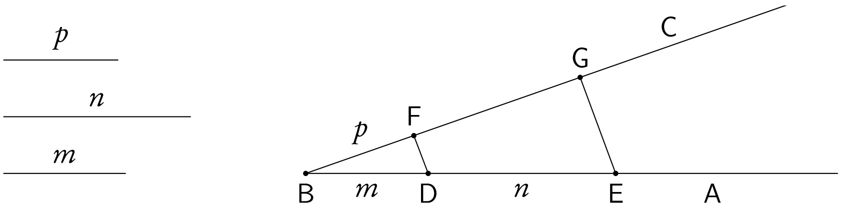


Figura 9.50.: *Costruzione del quarto proporzionale*

con la semicirconferenza. Il segmento $\overline{BD}$ è il medio proporzionale cercato: il triangolo ACD è infatti rettangolo, m ed n sono le proiezioni dei cateti sull'ipotenusa e $\overline{BD}$ è l'altezza relativa all'ipotenusa.

Costruzione della sezione aurea di un segmento

Dato un segmento $\overline{AB}$, si tracci una circonferenza con centro O sulla perpendicolare per B ad $\overline{AB}$ e raggio $\overline{OB} = \overline{AB}/2$. Congiunto A con O, sia D il punto di intersezione tra $\overline{AO}$ e la circonferenza. $\overline{AD}$ è la sezione aurea richiesta: basterà riportarla su $\overline{AB}$ con il compasso. Si veda la figura 9.51.

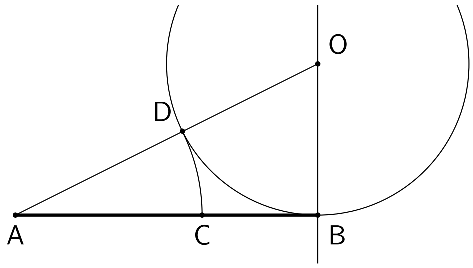


Figura 9.51.: *La sezione aurea di un segmento*

Costruzione del decagono, del pentagono e del pentadecagono regolari inscritti in un cerchio

Per costruire il decagono, in base al teorema 9.73, basta determinare la sezione aurea del raggio. Prendendo un punto si e uno no si costruisce poi il pentagono. Per il pentadecagono si può osservare quanto segue. Detto O il centro del cerchio, se $\overline{AB}$ è il lato dell'esagono (cioè il raggio del cerchio) e $\overline{AC}$ ne è la sezione aurea (cioè il lato del decagono), $\overline{CB}$ è il lato del pentadecagono, in quanto l'angolo al centro $C\hat{O}B$ è di 24°. Si veda la figura 9.52.

Costruzione di un segmento $\sqrt{n}$ volte l'unità

Se u è un segmento che assumiamo come unità, è possibile costruire facilmente con riga e compasso il segmento $\sqrt{n}u$, per ogni naturale n. Si comincia con il costruire una poligonale di due lati uguali ad u e tra di loro perpendicolari: il segmento che congiunge gli estremi è $\sqrt{2}\,u$. Riportando di nuovo u perpendicolarmente a quest'ultimo e congiungendo i nuovi estremi si ottiene $\sqrt{3}\,u$. Basta ripetere il procedimento per ottenere successivamente $\sqrt{4}\,u$, $\sqrt{5}\,u$, $\sqrt{6}\,u$, ecc. La figure 9.53 rende evidente questo procedimento. Non serve però partire sempre da una coppia di segmenti perpendicolari uguali ad u: si può partire da un segmento uguale ad u e uno uguale a $\sqrt{m}\,u$, dove m è il quadrato perfetto immediatamente precedente n. Per esempio per $\sqrt{17}\,u$ si può partire da u e da $4u = \sqrt{16}\,u$. Si veda la figura 9.53.

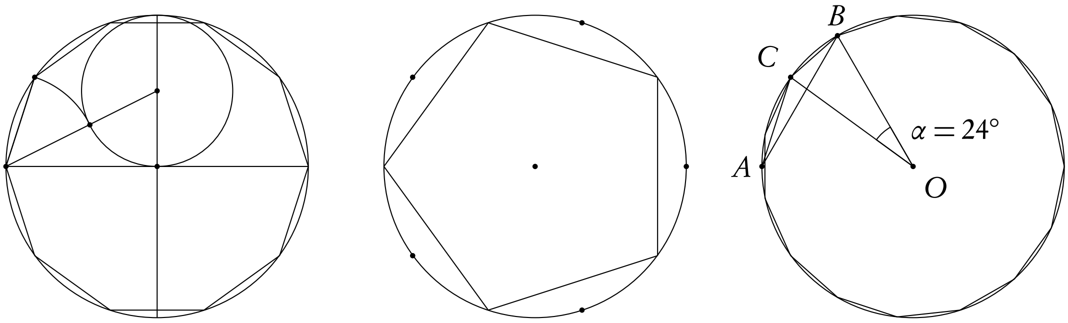


Figura 9.52.: *Decagono, pentagono e pentadecagono regolari*

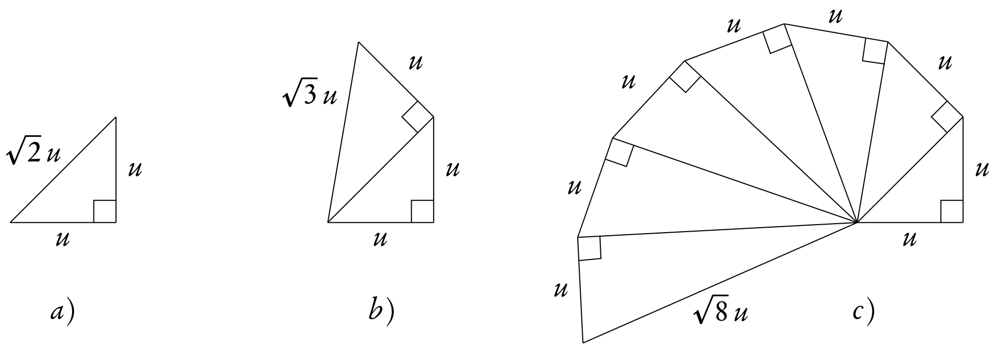


Figura 9.53.: *Costruzione di* $\sqrt{n}\,u$

9.9. Applicazioni dell'algebra alla geometria

Riportiamo in questa sezione alcune formule importanti della geometria piana, la maggior parte delle quali è conseguenza dei teoremi già noti, oppure immediata applicazione degli stessi.

Relazione fra lato e diagonale di un quadrato

Se l e d sono rispettivamente il lato e la diagonale di un quadrato, si ha

$$d = l\sqrt{2}. \tag{9.3}$$

Relazione fra gli elementi di un triangolo equilatero

Se l è il lato di un triangolo equilatero, h è l'altezza (che è anche mediana e bisettrice), e $\mathcal{A}$ è l'area, si ha

$$h = \frac{l}{2}\sqrt{3}, \quad \mathcal{A} = \frac{l^2}{4}\sqrt{3}. \tag{9.4}$$

Lati dei poligoni regolari inscritti in un cerchio

Se r è il raggio del cerchio e l_n, $n = 3, 4, 5, 6, 10$ il lato del poligono regolare di n lati inscritto nel cerchio, si ha

$$l_3 = r\sqrt{3}; \tag{9.5}$$

$$l_4 = r\sqrt{2}; \tag{9.6}$$

$$l_5 = \frac{r}{2}\sqrt{10-2\sqrt{5}}; \tag{9.7}$$

$$l_6 = r; \tag{9.8}$$

$$l_{10} = r\,\frac{\sqrt{5}-1}{2}. \tag{9.9}$$

Area di un triangolo in funzione dei lati

Se a, b, c sono i lati di un triangolo e $2p$ ne è il semiperimetro[5], si ha

$$\mathcal{A} = \sqrt{p(p-a)(p-b)(p-c)}, \tag{9.10}$$

detta *formula di Erone*.

Altezze di un triangolo in funzione dei lati

Dalla formula di Erone si deducono immediatamente le formule per le altezze h_a, h_b, h_c di un triangolo relative ai lati a, b, c rispettivamente, in funzione dei lati.

$$h_a = \frac{2}{a}\sqrt{p(p-a)(p-b)(p-c)}, \tag{9.11}$$
$$h_b = \frac{2}{b}\sqrt{p(p-a)(p-b)(p-c)}, \quad h_c = \frac{2}{c}\sqrt{p(p-a)(p-b)(p-c)}.$$

Mediane di un triangolo in funzione dei lati

Dette m_a, m_b, m_c le mediane di un triangolo relative ai lati a, b, c rispettivamente, si ha

$$m_a = \sqrt{2b^2+2c^2-a^2}, \quad m_b = \sqrt{2c^2+2a^2-b^2}, \quad m_c = \sqrt{2a^2+2b^2-c^2}. \tag{9.12}$$

Bisettrici di un triangolo in funzione dei lati

Dette s_a, s_b, s_c le bisettrici di un triangolo relative ai lati a, b, c rispettivamente, si ha

$$s_a = \frac{2}{b+c}\sqrt{bcp(p-a)}, \quad s_b = \frac{2}{c+a}\sqrt{cap(p-b)}, \quad s_c = \frac{2}{a+b}\sqrt{abp(p-c)}. \tag{9.13}$$

Raggio del cerchio inscritto e circoscritto a un triangolo

Se r ed R sono, rispettivamente, i raggi dei cerchi inscritto e circoscritto a un triangolo, a, b, c sono i lati, p è il semiperimetro e $\mathcal{A}$ è l'area del triangolo, si ha

$$r = \frac{\mathcal{A}}{p}; \tag{9.14}$$

$$R = \frac{abc}{4\mathcal{A}}. \tag{9.15}$$

[5] Il motivo per cui si indica, di solito, con $2p$ il perimetro di un triangolo è proprio legato al fatto che in alcune formule, tra cui quella qui presentata, compare il semiperimetro, e non il perimetro.

Raggio dei cerchi ex-inscritti a un triangolo

Se r_a, r_b e r_c sono i raggi dei cerchi tangenti ai lati a, b, c rispettivamente, p è il semiperimetro e $\mathcal{A}$ è l'area del triangolo, si ha

$$r_a = \frac{\mathcal{A}}{p-a}; \quad r_b = \frac{\mathcal{A}}{p-b}; \quad r_c = \frac{\mathcal{A}}{p-c}. \tag{9.16}$$

10. Geometria euclidea solida

Lo studio della geometria dello spazio, o *stereometria* si limita spesso solo alla trattazione dei solidi fondamentali e dei loro volumi, e questo rende abbastanza difficile, a livello degli studi universitari, una adeguata comprensione per esempio della teoria delle funzioni reali di due variabili reali o degli spazi lineari di dimensione 3 o superiori. Per questo nel presente capitolo richiameremo le principali nozioni relative alla geometria euclidea dello spazio, seppure in forma molto sintetica e senza proporre le dimostrazioni dei teoremi, anzi abitualmente senza nemmeno presentare le varie proprietà sotto forma di teorema.

10.1. Rette e piani nello spazio

10.1.1. Posizione reciproca di due rette

Due rette distinte dello spazio sono *parallele* se giacciono sullo stesso piano e non hanno alcun punto in comune.

Due rette dello spazio sono *incidenti* se hanno un solo punto in comune. Due rette incidenti sono sempre complanari. Due rette incidenti sono perpendicolari se individuano, nel loro piano, quattro angoli uguali.

Due rette distinte dello spazio che non siano né parallele né incidenti si dicono *sghembe*. Due rette sghembe sono necessariamente non complanari: presi du punti su di una e due punti sull'altra, non esiste alcun piano passante per i quattro punti.

Si noti che, a differenza di quanto accade nel piano, due rette che non hanno punti in comune possono essere parallele (e quindi complanari) o sghembe (e quindi non appartenenti allo stesso piano).

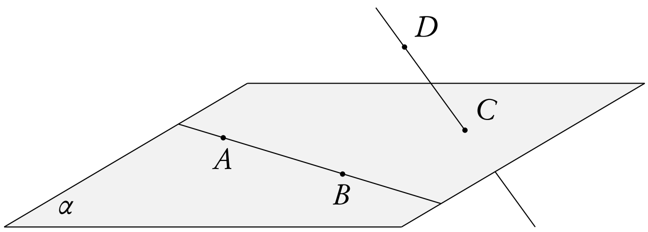


Figura 10.1.: *Rette sghembe*

10.1.2. Semispazi

Ogni piano α individua nello spazio due sottoinsiemi, detti *semispazi*, tra di loro opposti, ciascuno costituito dal piano stesso, detto *origine* e da due sottoinsiemi S_1 e S_2 non aventi alcun punto comune e tali che ogni segmento avente entrambi gli estremi in uno dei due sottoinsiemi S_1 e S_2 non ha alcun punto in comune con α, mentre ogni segmento avente un estremo in uno dei due sottoinsiemi e l'altro nell'altro ha un solo punto in comune con α.

Tutti i semispazi sono uguali; ogni semispazio è una figura convessa; una retta incidente un piano α in un punto P, ma non appartenente al piano, è divisa da P in due semirette che giacciono rispettivamente nei due semispazi opposti di origine α.

10.1.3. Intersezioni tra piani e tra piani e rette

Se due piani distinti hanno un punto in comune, hanno in comune una retta che passa per quel punto: la retta si chiama *retta intersezione* o semplicemente *intersezione* dei due piani. Due piani che non hanno alcun punto in comune si dicono *paralleli*.

Definizione 10.1. *Dati nello spazio due semipiani aventi retta origine in comune, essi dividono lo spazio in due parti, ciascuna delle quali è detta* angolo diedro*, o semplicemente* diedro.

I due semipiani si dicono *facce* del diedro, mentre la comune origine si dice *spigolo*. Se i due semipiani non appartengono allo stesso piano, uno dei due diedri è convesso, l'altro è concavo. Normalmente, a meno di diversa precisazione, parlando di diedro faremo sempre riferimento a quello convesso. Per i diedri si può ripetere quanto già detto a proposito degli angoli nel piano: i diedri si possono confrontare e sommare, si possono considerare diedri consecutivi e adiacenti, si parla di diedro piatto (che coincide con un semispazio) quando le due facce appartengono allo stesso piano, si parla di diedri opposti allo spigolo, che sono uguali, ecc.

Due piani si dicono *perpendicolari* se intersecandosi formano quattro diedri uguali.

Figura 10.2.: *Diedro. Piani perpendicolari*

Se due rette sono parallele, un piano che ne interseca una (senza contenerla) interseca anche l'altra.

Una retta e un piano si dicono *paralleli* se non hanno alcun punto in comune. Una retta è parallela ad un piano se e solo se è parallela ad una retta del piano. Da un punto fuori da un piano si possono condurre infinite rette parallele ad un piano: tutte le parallele appartengono ad uno stesso piano, che risulta essere parallelo al piano dato.

10.1.4. Perpendicolarità e parallelismo tra rette e piani

Esattamente come nel piano, per un punto fuori da una retta si può condurre una sola perpendicolare ad una retta data. Per un punto su una retta si possono invece condurre infinite rette perpendicolari alla retta stessa: tutte le perpendicolari appartengono ad uno stesso piano. Questo ci consente di definire la perpendicolarità tra retta e piano: una retta è *perpendicolare ad un piano* in un punto A se la retta è perpendicolare a tutte le rette del piano passanti per A.

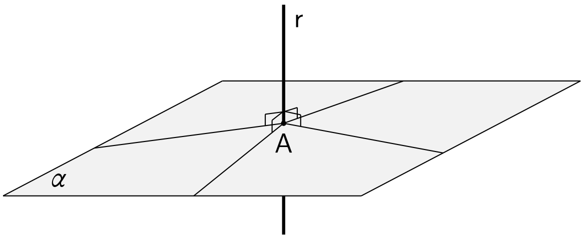


Figura 10.3.: *Perpendicolari ad una retta per un suo punto*

Una retta r che incontri un piano in un punto A, senza essere perpendicolare, si dice *obliqua* rispetto al piano. Una retta obliqua rispetto ad un piano forma angoli diversi con le rette del piano passanti per A. Fra tutte queste rette ce n'è sempre una che è perpendicolare alla r.

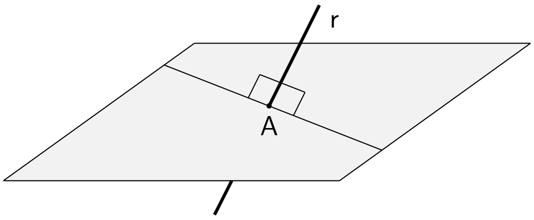


Figura 10.4.: *Retta obliqua rispetto a un piano*

Per ogni punto dello spazio si può condurre uno ed un solo piano perpendicolare a una retta data. Per ogni punto dello spazio si può condurre una ed una sola retta perpendicolare ad un dato piano. Il punto comune ad una piano e a una sua retta perpendicolare si chiama, come nella geometria piana, *piede* della perpendicolare.

Il più importante teorema sulla perpendicolarità tra rette e piani nello spazio è il seguente.

Teorema 10.2 (Teorema delle tre perpendicolari)**.** *Se dal piede* P *di una retta* r *perpendicolare ad un piano* α *si conduce la perpendicolare* s *ad una retta* t *qualunque del piano, quest'ultima è perpendicolare al piano di* r *ed* s.

Due rette perpendicolari allo stesso piano sono parallele. Viceversa se due rette sono parallele, ogni piano perpendicolare all'una è perpendicolare anche all'altra.

Due piani perpendicolari alla stessa retta sono paralleli.

Se un piano interseca due piani paralleli, le rette di intersezione sono parallele.

Se due rette incidenti sono parallele ad un piano, il loro piano è parallelo al piano dato.

Definizione 10.3. *Due piani paralleli dividono lo spazio in tre regioni: due semispazi e una* striscia solida*. I due piani si chiamano* facce *della striscia.*

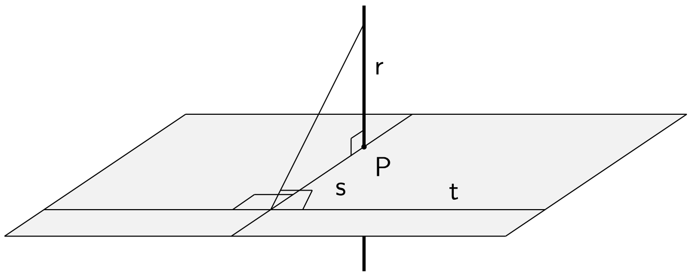


Figura 10.5.: *Il teorema delle tre perpendicolari*

10.2. Proiezioni, distanze ed angoli

Si chiama *proiezione di un punto su un piano* il piede della perpendicolare condotta da quel punto al piano.

Si chiama *proiezione di una retta su un piano* l'insieme delle proiezioni dei punti della retta sul piano. Se la retta non è perpendicolare al piano tale proiezione è una retta, se la retta è perpendicolare al piano la proiezione coincide con il piede della perpendicolare. Se la retta è parallela al piano, la retta e la sua proiezione sono parallele. La proiezione di un segmento su un piano si definisce in modo analogo e valgono proprietà analoghe.

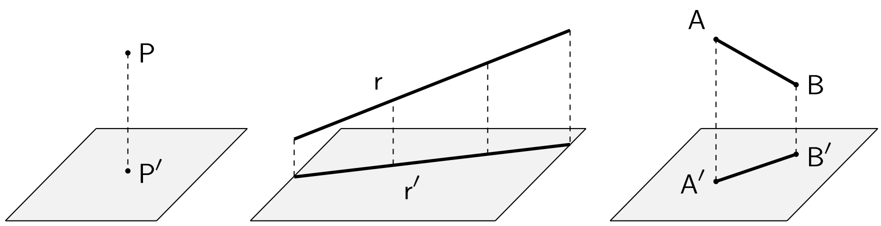


Figura 10.6.: *Proiezione di un punto, una retta, un segmento su un piano*

Si dice *distanza di un punto da un piano* il segmento di perpendicolare condotto da quel punto al piano.

Date una retta e un piano paralleli si dice distanza della retta dal piano la distanza di un punto della retta dal piano; dati due piani paralleli si dice distanza tra i piani la distanza di un punto di un piano dall'altro piano. La distanza di due piani paralleli si chiama anche altezza della striscia solida.

Data una retta r obliqua rispetto a un piano α, detta r' la sua proiezione su α e P la sua intersezione con α, ciascuno dei due angoli acuti tra r e r' si chiama *angolo tra la retta e il piano*. Questo angolo è il minimo tra gli angoli acuti formati da r e da una retta del piano uscente da P.

Date due rette sghembe e considerate le parallele alle rette condotte da un punto qualunque dello spazio, si chiama *angolo tra le due rette sghembe* l'angolo acuto o retto formato dalle due parallele. Se l'angolo è retto, le due rette sghembe si dicono *perpendicolari* o ortogonali, anche se le due rette non sono perpendicolari nel senso usualmente dato a questo concetto.

Date due rette sghembe r ed s, esiste una e una sola retta p perpendicolare ad entrambe ed il segmento di p compreso tra le due rette r ed s è minore di ogni altro segmento che congiunga due punti di r ed s:

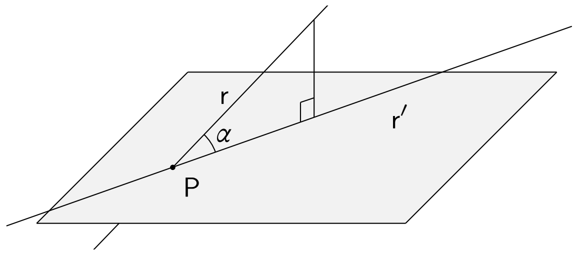


Figura 10.7.: *Angolo tra retta e piano*

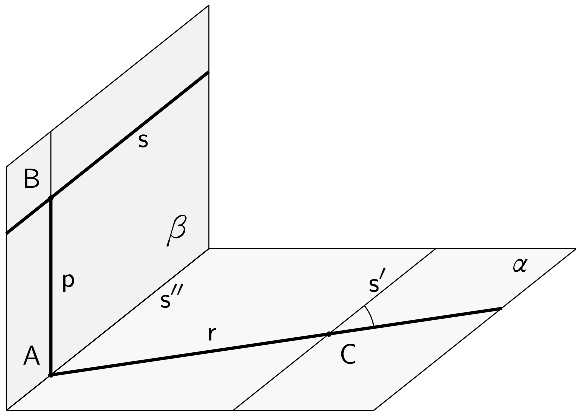


Figura 10.8.: *Angolo e distanza tra due rette sghembe*

esso si chiama *distanza delle due rette sghembe*. Con riferimento alla figura 10.8, consideriamo la parallela s' ad s per un punto C di r. Indichiamo con α il piano di r ed s', che risulta essere parallelo ad s. Per s conduciamo il piano β perpendicolare ad α, e sia s'' l'intersezione tra i due piani. Detto A il punto di intersezione tra r ed s'', conduciamo per A la perpendicolare p, nel piano β, ad s'', che incontra in B la s. Tale retta risulta perpendicolare comune alle due rette date e il segmento $\overline{AB}$ realizza la minima distanza. L'angolo tra r ed s' evidenziato nella figura 10.8 è l'angolo tra le due rette sghembe.

10.3. Angoloidi. Poliedri

Definizione 10.4. *Date in un certo ordine n ($n > 2$) semirette aventi l'origine in comune, a tre a tre non complanari, e tali che il piano individuato da due successive di esse lasci tutte le altre dalla stessa parte, si dice* angoloide convesso[(1)] *o, semplicemente,* angoloide, *l'intersezione degli n semispazi che hanno per origine quei piani e che contengono tutte le semirette date.*

[1] Un angoloide è un caso particolare di *angolo solido*, di cui possiamo dare la seguente definizione semplificata. Consideriamo una sfera e una curva semplice chiusa tracciata su di essa. Tracciamo poi le semirette con origine nel centro della sfera e passanti per i punti di questa curva: la superficie da esse individuata divide lo spazio in due regioni, ciascuna delle quali è un angolo solido. Si assume come misura di questo angolo il rapporto tra la porzione di superficie sferica racchiusa entro l'angolo e il quadrato del raggio della sfera. L'unità di misura si chiama *steradiante*.

L'origine comune delle semirette si dice *vertice*, le n semirette si chiamano *spigoli*, gli angoli (convessi) individuati da due semirette consecutive sono le *facce* dell'angoloide.

Se un angoloide ha tre facce (il minimo possibile) si chiama *triedro*, se ha quattro facce *angoloide tetraedro*, ecc.

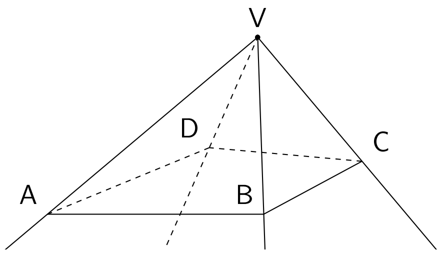


Figura 10.9.: *Un angoloide tetraedro*

Un angoloide si può definire anche in altri modi, tra cui riportiamo il seguente perché molto intuitivo. Dati un poligono non intrecciato, $ABCD\dots$, in un piano e un punto V fuori dal piano, un angoloide è individuato da tutte le semirette di origine V e passanti per i punti A, B, C, D, … Se il poligono è convesso si ottiene un angoloide convesso come quello della definizione 10.4, se il poligono è concavo si ottiene invece un angoloide concavo, di cui comunque non ci occuperemo nel seguito.

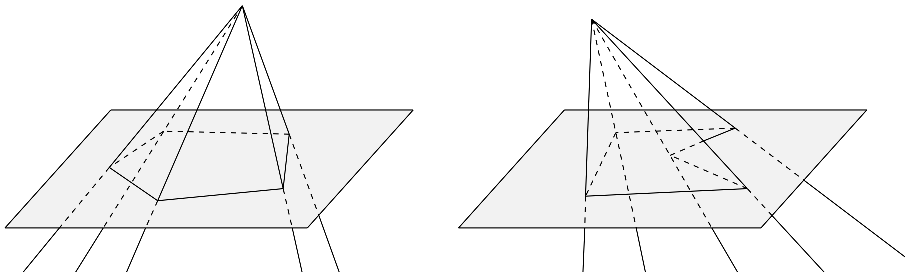


Figura 10.10.: *Angoloide convesso, a sinistra, e angoloide concavo, a destra*

La somma delle facce di un angoloide è minore di un angolo giro e ciascuna faccia è minore della somma di tutte le altre.

10.3.1. La piramide

Definizione 10.5. *Dato un angoloide (convesso) di vertice* V *ed un piano* α *che non passa per il vertice e interseca tutti gli spigoli, si dice* piramide *l'intersezione dell'angoloide*[2] *e del semispazio di origine* α *e che contiene* V.

Il piano α interseca l'angoloide secondo un poligono convesso che si dice *base* della piramide. I triangoli intersezione delle facce dell'angoloide con il semispazio che definisce la piramide si chiamano *facce laterali* della piramide, e la loro unione si dice *superficie laterale*; l'unione della superficie laterale

[2] Si potrebbero anche considerare piramidi concave, delle quali però non ci occuperemo.

e della base si chiama *superficie totale* della piramide. I lati delle facce laterali e della base si chiamano *spigoli* e il vertice dell'angoloide è anche *vertice* della piramide. La distanza del vertice V dal piano α si chiama *altezza* della piramide.

È utile tenere conto della seguente proprietà delle piramidi: se dal piede H dell'altezza si tira la perpendicolare alla retta di uno degli spigoli della base e K è il punto di intersezione, $\overline{\mathsf{VK}}$ è l'altezza della corrispondente faccia triangolare della piramide relativa allo spigolo della base.

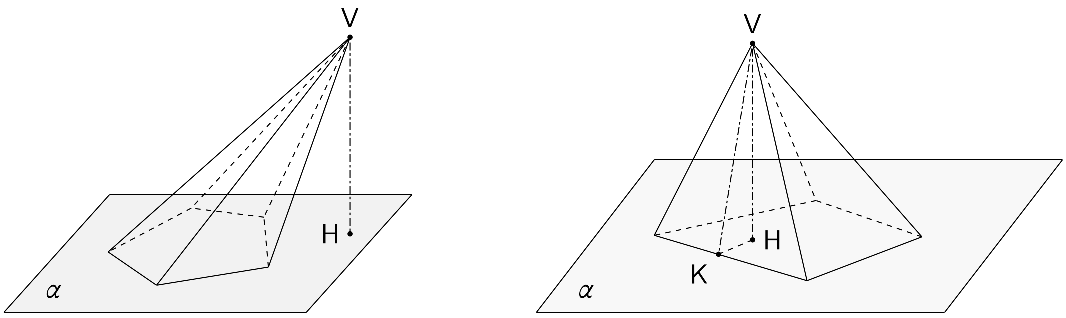


Figura 10.11.: *Piramide e sua proprietà*

Teorema 10.6. *Il poligono che si ottiene sezionando una piramide con un piano parallelo alla base è simile a questa; i perimetri del poligono sezione e della base stanno fra loro come le distanze dal vertice; l'area del poligono sezione e della base stanno fra loro come i quadrati delle distanze dal vertice*

Definizione 10.7. *Una piramide si dice* retta *se ha per base un poligono circoscrittibile*[3] *ad un cerchio e l'altezza cade nel centro di questo cerchio.*

Le altezze[4] delle facce laterali di una piramide retta relative agli spigoli della base sono uguali. Ciascuna di queste altezze uguali si chiama *apotema* della piramide. Si noti che in ogni piramide retta l'altezza h, l'apotema a, e il raggio r del cerchio inscritto nella base sono lati di un triangolo equilatero, per cui

$$a^2 = h^2 + r^2$$

Una piramide retta che ha per base un poligono regolare è talvolta detta piramide regolare, ma la locuzione è impropria perché l'attributo "regolare" è usato, in geometria solida, con un altro significato (vedi il paragrafo 10.3.5).

La determinazione delle aree della superficie laterale e totale di una piramide è molto facile applicando le già note formule per le aree di triangoli e poligoni.

Definizione 10.8. *Una piramide retta la cui base e le cui tre facce laterali sono triangoli equilateri si dice un* tetraedro regolare.

In un tetraedro regolare si usa il termine faccia anche per la base e, anzi, ciascuna delle quattro facce può essere assunto come base e il vertice ad essa opposto come vertice del tetraedro. Inoltre tutti gli spigoli e tutti gli angoli diedri tra due facce sono uguali. Se si sviluppa la superficie totale di un tetraedro regolare su un piano si può ottenere un triangolo equilatero, di lato doppio di uno spigolo.

[3] Si noti che se il poligono di base non fosse circoscrittibile non avrebbe senso parlare di centro.
[4] Questa proprietà è legata al fatto che il poligono di base è circoscrittibile a un cerchio.

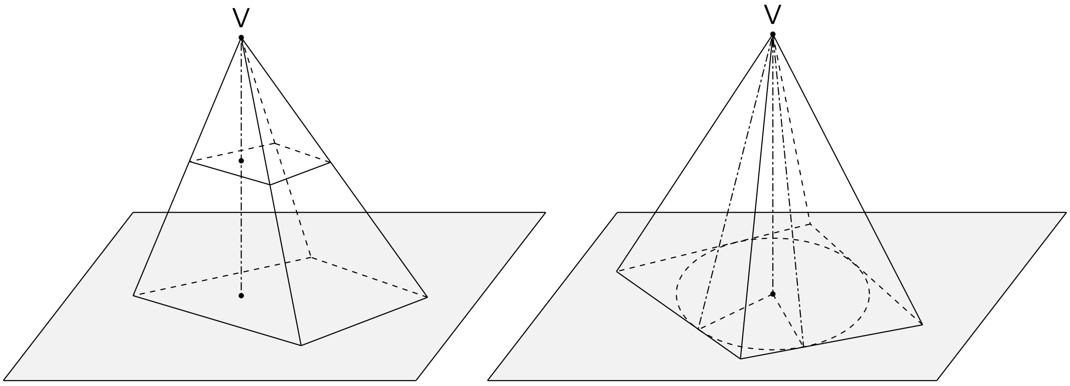


Figura 10.12.: *Sezioni di una piramide, a sinistra, e piramide retta, a destra*

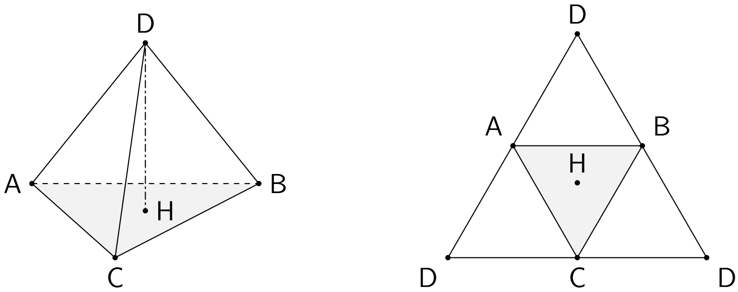


Figura 10.13.: *Tetraedro regolare e sviluppo della sua superficie totale*

10.3.2. Il tronco di piramide

Definizione 10.9. *Sezionando una piramide con un piano parallelo alla base si divide la piramide in due parti, di cui quella che contiene il vertice è ancora una piramide, mentre quella che non contiene il vertice si chiama* tronco di piramide *a basi parallele, o semplicemente tronco di piramide.*

Sono evidenti i significati di basi del tronco (che sono poligoni simili), facce laterali (che sono trapezi), altezza.

Se il tronco è ottenuto sezionando una piramide retta si dice *tronco di piramide retta*. In questo caso le facce laterali sono trapezi di uguale altezza che si dice *apotema* e, naturalmente, entrambi i poligoni di base sono circoscrittibili ad un cerchio. La determinazione delle aree delle superfici laterale e totale di un tronco di piramide è molto facile tenendo conto delle note formule per le aree dei trapezi e dei poligoni.

10.3.3. Il prisma

Definizione 10.10. *Date, in un certo ordine, n ($n > 2$) rette parallele a tre a tre non complanari e tali che il piano individuato da due successive di esse lasci tutte le altre dalla stessa parte, si dice* prisma convesso

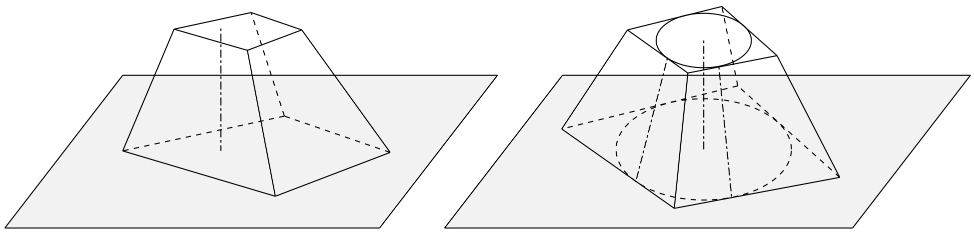


Figura 10.14.: *Tronco di piramide e tronco di piramide retta*

indefinito[5] *o semplicemente* prisma indefinito, *l'insieme comune agli n semispazi che hanno per origine i piani individuati da due semirette consecutive e che lasciano tutte le altre rette dalla stessa parte.*

Le rette parallele si dicono *spigoli* e le strisce comprese fra due rette parallele consecutive si dicono *facce* del prisma indefinito. Il poligono che si ottiene intersecando il prisma indefinito con un piano non parallelo agli spigoli si dice *sezione* del prisma.

Definizione 10.11. *Si dice* prisma finito *o, semplicemente,* prisma *l'intersezione di un prisma indefinito con una striscia solida le cui facce non sono parallele agli spigoli.*

Sono evidenti i significati di basi (che sono poligoni simili), di facce laterali (che sono parallelogrammi), di altezza di un prisma. Si usa il nome di spigoli per tutti i lati sia delle facce laterali che delle basi. Tutti i vertici delle facce laterali e delle basi si dicono anche *vertici del prisma*.

Definizione 10.12. *Un prisma*[6] *si dice* retto *se i piani delle basi sono perpendicolari agli spigoli laterali. Un prisma non retto si dice* obliquo.

La determinazione delle aree della superficie laterale o della superficie totale di un prisma si fa tenendo conto delle già note formule per le aree dei poligoni e parallelogrammi.

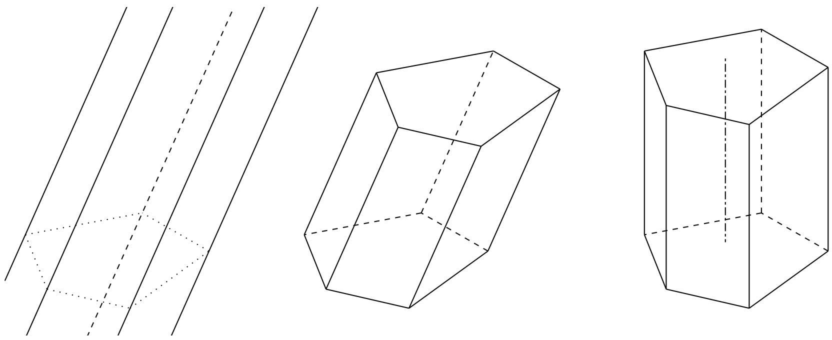


Figura 10.15.: *Prisma indefinito, a sinistra, prisma finito, al centro, prisma retto, a destra*

[5]Si potrebbero, in analogia a quanto fatto con gli angoloidi, considerare anche prismi concavi, ma non tratteremo questo argomento nel seguito.

[6]Si noti che, a differenza delle piramidi, per un prisma retto non si fa alcuna ipotesi sul tipo di poligono di base.

Il parallelepipedo e il cubo

Definizione 10.13. *Un prisma avente per basi due parallelogrammi si dice un* parallelepipedo.

In questo caso si usa il termine *facce* anche per le basi. Dunque un parallelepipedo ha sei facce e tutte sono parallelogrammi. Due vertici non appartenenti alla stessa faccia si dicono *opposti*. Ogni vertice ne ha uno opposto e il segmento che li unisce si dice *diagonale*. Ogni parallelepipedo ha quattro diagonali che si incontrano in uno stesso punto, che è medio per ciascuna di esse. Due facce che non hanno vertici comuni si dicono *opposte* e sono uguali (oltreché parallele). Due facce qualunque possono essere assunte come basi.

Definizione 10.14. *Si dice* parallelepipedo rettangolo *un parallelepipedo retto avente per basi due rettangoli.*

Ogni faccia di un parallelepipedo rettangolo è un rettangolo e tre spigoli che concorrono in uno stesso vertice si dicono le *dimensioni*. Le quattro diagonali di un parallelepipedo rettangolo sono tra loro uguali.

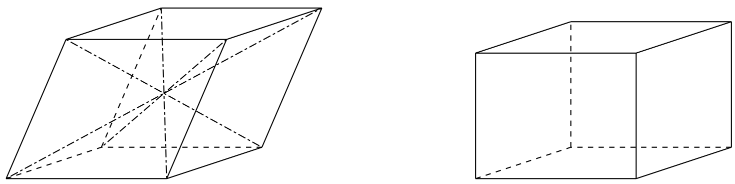


Figura 10.16.: *Parallelepipedo e sue diagonali, a sinistra, e parallelepipedo rettangolo, a destra*

Definizione 10.15. *Un parallelepipedo rettangolo avente le tre dimensioni uguali si dice* cubo *o* esaedro regolare.

10.3.4. Poliedri in generale

Definizione 10.16. *Si dice* superficie poliedrica *(convessa) la figura formata da poligoni convessi a due a due non complanari, tali che ogni lato di uno sia comune ad un altro poligono e che il piano di ogni poligono lasci tutti gli altri poligoni nello stesso semispazio. La regione di spazio racchiusa da una superficie poliedrica si dice* poliedro *(convesso).*

I poligoni, i loro lati e i loro vertici si dicono facce, spigoli e vertici della superficie poliedrica o del poliedro. Gli angoloidi (convessi) individuati dalle semirette degli spigoli uscenti da un vertice sono gli angoloidi della superficie poliedrica o del poliedro; i diedri (convessi) individuati dai piani di due facce aventi uno spigolo in comune sono i diedri della superficie poliedrica o del poliedro. Le piramidi e i prismi sono esempi di poliedri.

Teorema 10.17. *Se f è il numero delle facce n quello dei vertici, s quello degli spigoli di un poliedro, si ha*

$$f+n=s+2, \tag{10.1}$$

detta formula di Eulero per i poliedri.

Come già visto nel caso del tetraedro regolare, la superficie totale di un poliedro può essere "sviluppata" su un piano: si può immaginare di appoggiare il poliedro su un piano e di ruotare una qualunque altra faccia, adiacente a quella prescelta, intorno al lato comune fino a che si trovi sullo stesso piano e procedendo similmente per le altre facce successivamente adiacenti. Si ottiene in questo modo un poligono formato dall'unione di tanti poligoni uguali alle facce del poliedro.

10.3.5. Poliedri regolari

Definizione 10.18. *Un poliedro convesso si dice* regolare *quando le sue facce sono poligoni regolari uguali e i suoi diedri sono tutti uguali.*

A proposito dei poliedri regolari vale il seguente teorema, di cui forniamo anche la dimostrazione, in quanto si tratta di uno dei risultati classici più famosi della geometria.

Teorema 10.19. *Esistono solo cinque poliedri regolari (convessi).*

Dimostrazione. Ricordiamo che la somma delle facce di un angoloide convesso è minore di un angolo giro, cioè di 360° e che un angoloide ha almeno tre facce.

Cominciamo a considerare i poliedri regolari le cui facce siano triangoli equilateri. Gli angoli di questi triangoli sono di 60° e quindi possono esistere angoloidi solo con 3, 4 o 5 facce di 60°. A questi angoloidi corrispondono il *tetraedro regolare*, che abbiamo già considerato, l'*ottaedro regolare* con otto facce e l'*icosaedro regolare* con 20 facce. Il tetraedro regolare ha 4 facce, 4 vertici e 6 spigoli; l'ottaedro regolare ha 8 facce, 6 vertici e 12 spigoli; l'icosaedro regolare ha 20 facce, 12 vertici e 30 spigoli

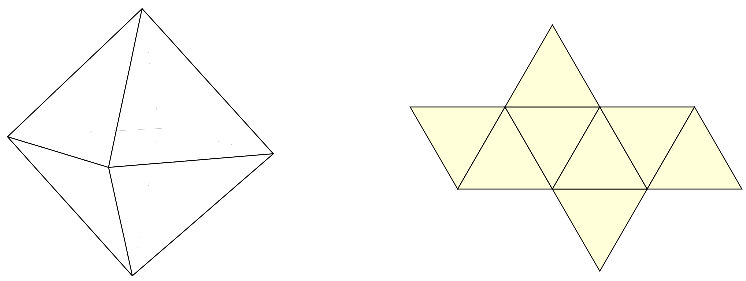


Figura 10.17.: *Ottaedro regolare e sviluppo della superficie totale*

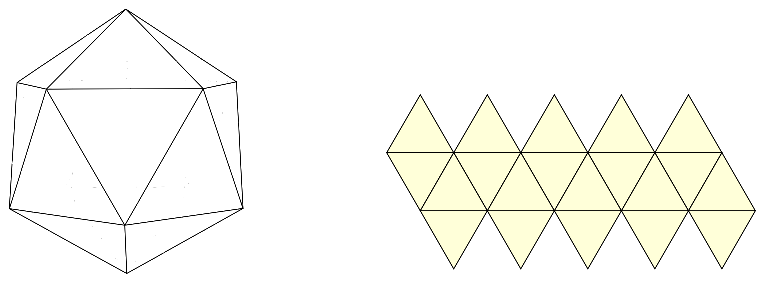


Figura 10.18.: *Icosaedro regolare e sviluppo della superficie totale*

Consideriamo ora il caso in cui le facce sono quadrati. Poiché gli angoli di un quadrato sono di 90°, può esistere solo un angoloide con 3 facce, cui corrisponde il cubo o *esaedro regolare* di cui abbiamo già parlato. L'esaedro regolare ha 6 facce, 8 vertici e 12 spigoli.

Passando poi ai pentagoni, tenendo conto che gli angoli dei pentagoni regolari sono di 108°, se ne deduce che esiste un solo angoloide con 3 facce, cui corrisponde il *dodecaedro regolare*. Il dodecaedro regolare ha 12 facce, 20 vertici e 30 spigoli

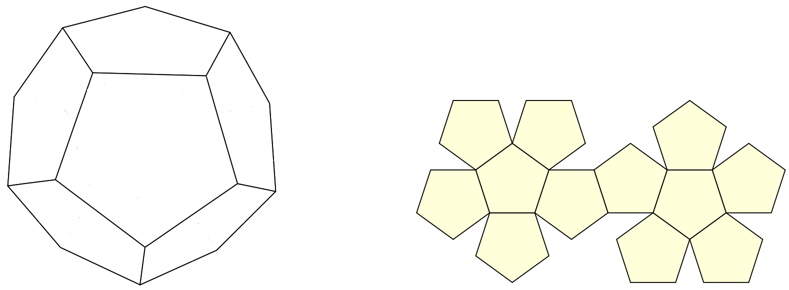


Figura 10.19.: *Dodecaedro regolare e sviluppo della superficie totale*

Per i poligoni regolari con più di 5 lati, il triplo dei loro angoli supera i 360°, per cui non possono esistere poliedri regolari con facce di questo tipo.

Gli unici poliedri regolari sono dunque quelli considerati. Essi sono anche noti con il nome di *solidi platonici*. □

10.4. I corpi rotondi

Definizione 10.20. *Consideriamo un semipiano α di origine* a *e una linea qualsiasi γ giacente in α. La superficie descritta da γ in una rotazione completa attorno ad* a *si chiama* superficie di rotazione. *Se consideriamo, nel semipiano α, una superficie chiusa Σ, la stessa rotazione produce una regione dello spazio detta* solido di rotazione. *La retta* a *si chiama* asse di rotazione *e la linea γ* generatrice *della superficie di rotazione.*

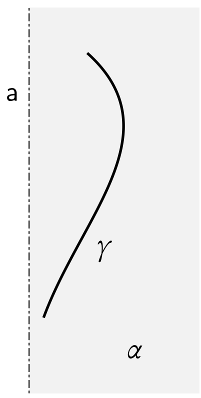
10.4.1. Il cilindro

Definizione 10.21. *Si dice* cilindro indefinito *il solido generato da una striscia in una rotazione completa attorno ad una retta del suo contorno. La superficie generata dalla rotazione della retta non fissa del contorno della striscia si dice* superficie cilindrica indefinita.

La retta fissa del contorno della striscia si dice *asse del cilindro indefinito*, l'altra retta e tutte le rette da essa generate nella rotazione si dicono *generatrici della superficie cilindrica*. L'altezza della striscia si dice *raggio del cilindro indefinito*. La superficie cilindrica è detta *contorno* del cilindro.

Il cilindro indefinito e il suo contorno si possono anche definire come il luogo dei punti dello spazio che hanno distanza minore od uguale (rispettivamente uguale) ad un segmento assegnato r.

Definizione 10.22. *Si dice* cilindro circolare retto *o, semplicemente,* cilindro *l'intersezione di una striscia solida con un cilindro circolare indefinito avente l'asse perpendicolare alle due facce della striscia solida.*



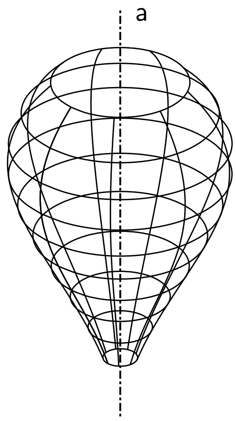
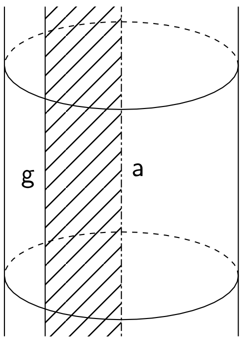


Figura 10.20.: *Una curva piana e la superficie ottenuta per rotazione attorno ad un'asse*

L'altezza della striscia solida si chiama *altezza del cilindro*; i due cerchi intersezione del cilindro indefinito con le facce della striscia si dicono *basi del cilindro*; il raggio del cilindro indefinito è anche *raggio del cilindro*. Sono evidenti il significato di superficie laterale e totale di un cilindro.

Un cilindro si può pensare come il solido generato dalla rotazione completa di un rettangolo attorno ad uno dei suoi lati.

Definizione 10.23. *Un cilindro si dice* equilatero *se la sua altezza è uguale al diametro di base, ovvero se la sua sezione con un piano per l'asse è un quadrato.*



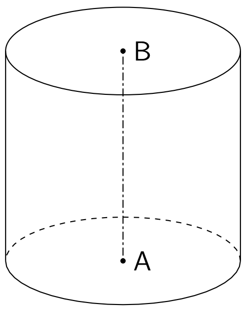


Figura 10.21.: *Cilindro indefinito e cilindro circolare retto*

Teorema 10.24. *La superficie laterale di un cilindro è equivalente ad un rettangolo che ha per lati l'altezza del cilindro e la circonferenza di base rettificata.*

Ha interesse nelle applicazioni anche il *cilindro obliquo*, ottenuto intersecando un cilindro circolare indefinito con due piani di cui uno perpendicolare all'asse e l'altro obliquo, piani che non si intersecano all'interno del cilindro stesso. In questo caso i due segmenti di generatrice più lungo e più corto si chiamano *altezza maggiore* e *altezza minore* del cilindro. Si veda la figura 10.23.

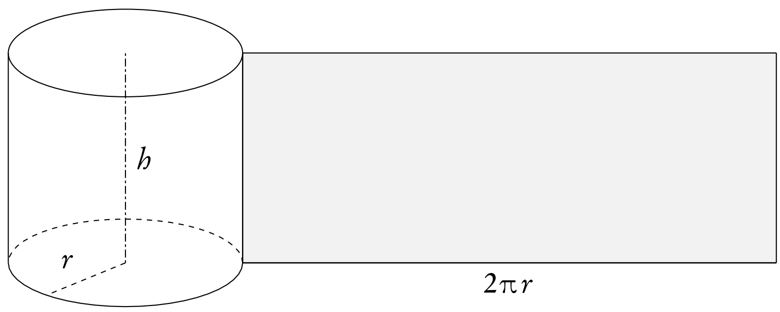


Figura 10.22.: *Cilindro e sviluppo della sua superficie laterale*

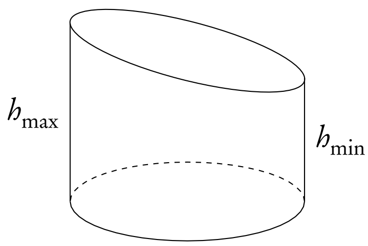


Figura 10.23.: *Cilindro circolare obliquo*

10.4.2. Il cono

Definizione 10.25. *Si dice* cono indefinito *il solido generato da un angolo acuto α nella rotazione completa attorno a uno dei suoi lati. La superficie generata dall'asse che ruota si chiama* superficie conica indefinita.

L'angolo α si dice *semiapertura* del cono. Il lato fisso dell'angolo α si dice *asse del cono*; il lato dell'angolo α che ruota e tutte le sue successive posizioni si dicono *generatrici* della superficie conica indefinita.

Teorema 10.26. *Le sezioni di un cono indefinito con piani perpendicolari all'asse sono cerchi che stanno fra come i quadrati delle rispettive distanze dal vertice.*

Teorema 10.27. *Si dice* cono circolare retto *o, semplicemente,* cono*, l'intersezione di un cono indefinito con un semispazio contenente il vertice del cono e il cui piano origine interseca perpendicolarmente l'asse del cono in un punto diverso dal vertice.*

Il cerchio intersezione del piano origine del semispazio con il cono indefinito si dice *base* del cono, la distanza del vertice dalla base si chiama *altezza*, le intersezioni delle generatrici con il semispazio si chiamano *apotemi*, e sono tutti uguali.

Un cono circolare retto si può anche pensare come il solido ottenuto dalla rotazione di un triangolo rettangolo attorno a uno dei suoi cateti.

Teorema 10.28. *La superficie laterale di un cono è equivalente ad un triangolo avente per base la circonferenza di base rettificata e per altezza l'apotema.*

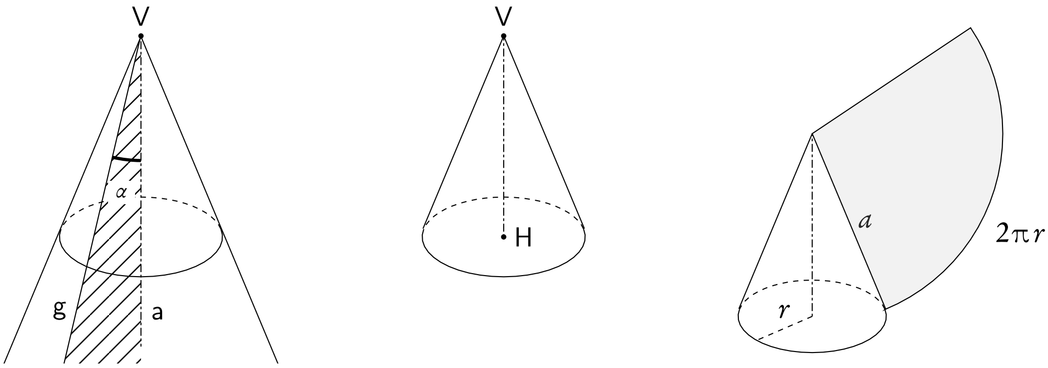


Figura 10.24.: *Cono indefinito, cono circolare retto e sviluppo della sua superficie laterale*

Definizione 10.29. *Dato un cono di vertice* V*, consideriamone la sezione con un piano parallelo alla base. Essa divide il cono in due parti di cui quella contenente il vertice è ancora un cono, l'altra si chiama* tronco di cono a basi parallele *o, semplicemente,* tronco di cono.

I due cerchi sezione si dicono *basi*, la distanza tra esse si dice *altezza*, i segmenti di generatrici si dicono *lati* o *apotemi* del tronco di cono.

Il tronco di cono si può anche pensare come il solido generato da un trapezio rettangolo in una rotazione completa attorno al lato perpendicolare alle basi.

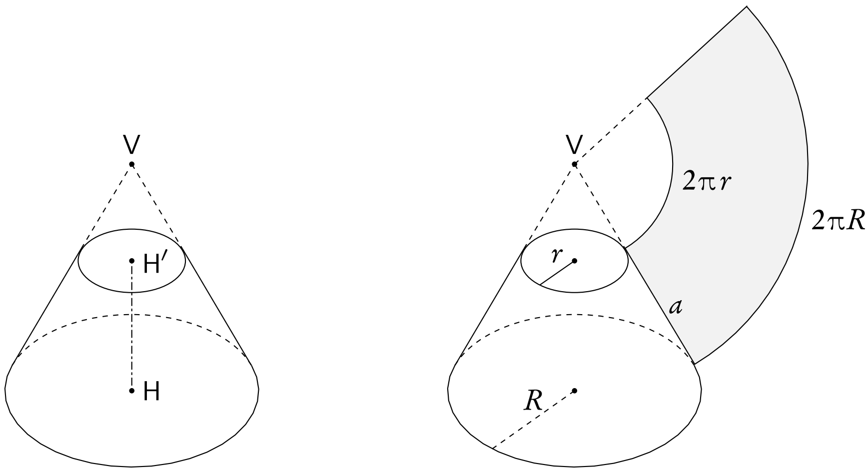


Figura 10.25.: *Tronco di cono e sviluppo della sua superficie laterale*

10.4.3. La sfera

Definizione 10.30. *Si dice* sfera *il solido generato da una rotazione completa di un semicerchio attorno al suo diametro. Si dice* superficie sferica *la superficie generata dalla rotazione completa di una semicirconferenza attorno al suo diametro. La superficie sferica è il contorno della sfera.*

Sia per la sfera che per la superficie sferica si considerano, con ovvio significato, il *centro* e il *raggio*. La superficie sferica e la sferica possono anche essere definite come il luogo dei punti dello spazio aventi da un punto fisso detto centro distanza uguale o, rispettivamente, minore o uguale ad un segmento assegnato detto raggio.

Una retta ha due punti, un punto, nessun punto in comune con una sfera a seconda che la sua distanza dal centro sia minore, uguale o maggiore del raggio; nei tre casi la retta si dice, rispettivamente, *secante*, *tangente*, *esterna*.

Un piano ha in comune con una superficie sferica una circonferenza, un punto, nessun punto, a seconda che la sua distanza dal centro sia minore, uguale o maggiore del raggio; nei tre casi il piano si dice, rispettivamente, *secante*, *tangente*, *esterno*. Se un piano secante passa per il centro, la circonferenza intersezione ha lo stesso centro e raggio della superficie sferica, se non passa per il centro ha raggio minore di quello della superficie sferica.

Una retta o un piano passanti per il centro di una superficie sferica si dicono *diametrali*.

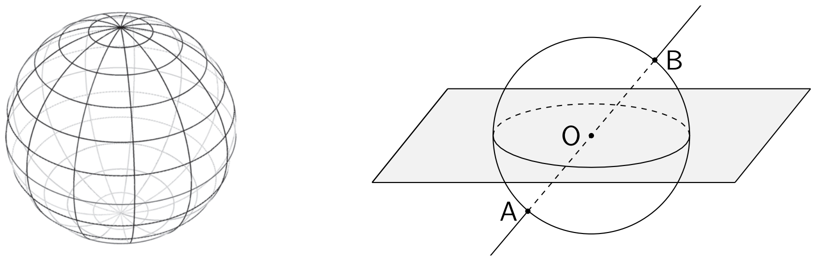


Figura 10.26.: *Sfera, a sinistra, retta e piano diametrali, a destra*

Sia data una superficie sferica e consideriamone un diametro qualunque. I piani passanti per il diametro intersecano la superficie sferica secondo circonferenze massime, che sono dette *meridiani*. I piani perpendicolari al diametro intersecano la superficie secondo circonferenze, dette *paralleli*, di raggio diverso, a seconda della distanza dal centro. Il piano passante per il centro individua la circonferenza massima, detta *equatore*. Il diametro a cui sono perpendicolari i piani sezionanti è detto *asse* e i suoi estremi sono detti *poli*.

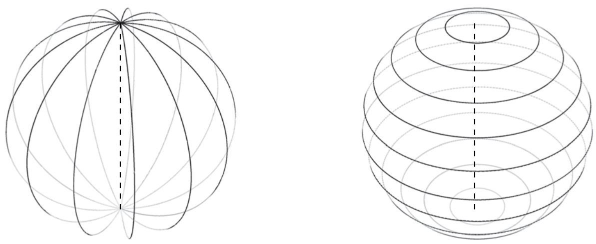


Figura 10.27.: *Meridiani, a sinistra, e paralleli, a destra, su una superficie sferica*

Teorema 10.31. *L'area della superficie sferica è uguale a quattro volte l'area del suo cerchio massimo.*

Detto r il raggio della sfera Σ, si ha dunque

$$\mathcal{A}(\Sigma) = 4\pi r^2. \tag{10.2}$$

Parti della superficie sferica e della sfera

Definizione 10.32. *Data una superficie sferica e un piano α perpendicolare in un punto* C *ad un suo diametro* $\overline{\mathsf{AB}}$, *si dice* calotta sferica *ciascuna delle due parti in cui la superficie viene divisa dal piano α. La corrispondente parte di sfera si chiama invece* segmento sferico a una base

La circonferenza di centro C, intersezione della superficie sferica con il piano α si dice *base della calotta*, mentre il corrispondente cerchio è la *base del segmento sferico*; il segmento $\overline{\mathsf{AC}}$, oppure $\overline{\mathsf{CB}}$, si dice *altezza* della calotta o del segmento sferico.

Definizione 10.33. *Data una superficie sferica e due piani α e β perpendicolari rispettivamente in* C *e* D *ad un suo diametro* $\overline{\mathsf{AB}}$, *si dice* zona sferica *la parte di superficie sferica compresa tra i due piani. La corrispondente parte di sfera si chiama* segmento sferico a due basi.

Sono evidenti il significato di basi e altezza della zona o del segmento.

Teorema 10.34. *L'area $\mathcal{A}$ della superficie di una calotta o di una zona sferica di altezza h, appartenenti a una sfera di raggio r, è data da*

$$\mathcal{A} = 2\pi r h. \tag{10.3}$$

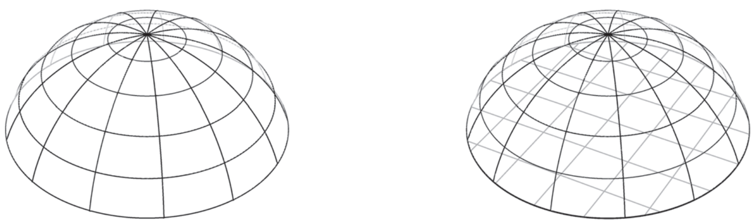


Figura 10.28.: *Calotta sferica e segmento sferico a una base*

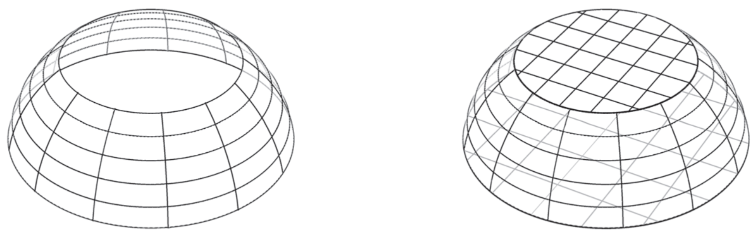


Figura 10.29.: *Zona sferica e segmento sferico a due basi*

Definizione 10.35. *Si dice* fuso sferico *l'intersezione di una superficie sferica con un diedro il cui spigolo passa per il centro della superficie sferica. La corrispondente parte di sfera si chiama* spicchio sferico.

L'angolo diedro si chiama angolo del fuso o dello spicchio. Poiché i fusi appartenenti ad una stessa sfera stanno fra loro come i rispettivi angoli, l'area $\mathcal{A}$ di un fuso si può trovare con una proporzione: detti r il raggio della sfera e α l'angolo del fuso, si ha

$$4\pi r^2 : \mathcal{A} = 360 : \alpha. \tag{10.4}$$

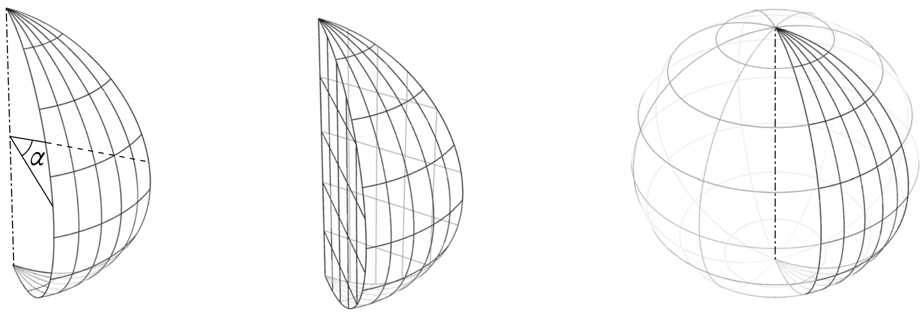


Figura 10.30.: *Fuso sferico, spicchio sferico e la sfera cui appartengono*

Poliedri inscritti e circoscritti

Un poliedro si dice *inscritto* in una sfera se tutti i suoi vertici appartengono alla superficie sferica, si dice invece *circoscritto* se tutte le sue facce sono tangenti alla superficie sferica.

Un poliedro regolare è sempre inscrivibile e circoscrivibile ad una sfera. Le sfere inscritte e circoscritte a un poliedro hanno il medesimo centro, detto *centro del poliedro*, che è anche il punto di incontro di tutte le rette perpendicolari alle facce nei rispettivi centri. Il raggio della sfera inscritta si chiama *apotema* del poliedro, quello della sfera circoscritta si chiama *raggio* del poliedro.

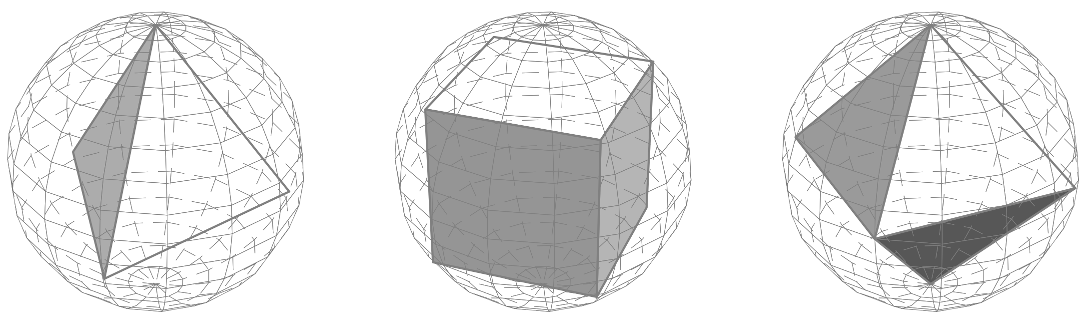


Figura 10.31.: *Tetraedro, esaedro, ottaedro regolari e sfera circoscritta*

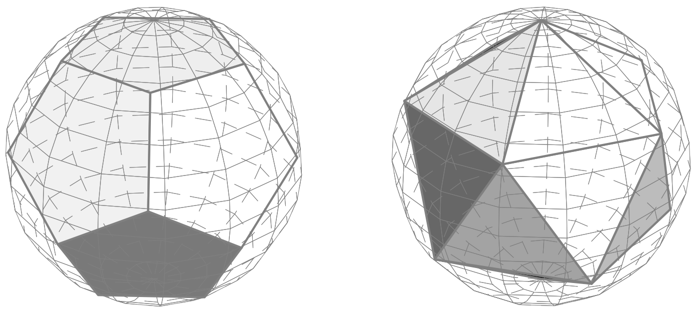


Figura 10.32.: *Dodecaedro e icosaedro regolari e sfera circoscritta*

10.5. Estensione solida e volumi

Analogamente a quanto fatto per le linee e le superfici, anche per le figure solide ammettiamo come primitivo il concetto di *estensione solida*.

Definizione 10.36. *Due solidi aventi uguale estensione si dicono* equivalenti.

Tutto quanto detto a proposito dell'equivalenza di superfici si trasferisce all'equivalenza tra solidi. In analogia al caso delle superfici, le estensioni solide costituiscono una classe di grandezze omogenee e, sempre in analogia con il caso delle superfici, è possibile introdurre la misura delle estensioni solide, una volta scelto il campione che, come è naturale, sarà un cubo costruito sull'unità delle lunghezze. Diremo *volume* la misura dell'estensione solida.

Di seguito riportiamo gli enunciati dei teoremi che permettono il calcolo del volume dei solidi che abbiamo considerato. Quando non precisato, il significato dei simboli è evidente.

Teorema 10.37. *Il volume di un prisma o di un cilindro è uguale al prodotto dell'area di base per la lunghezza dell'altezza. In particolare il volume di un parallelepipedo rettangolo è uguale al prodotto delle sue tre dimensioni:*

$$\mathcal{V} = \mathcal{A}_b h. \tag{10.5}$$

Teorema 10.38. *Il volume di una piramide o di un cono è uguale a un terzo del prodotto dell'area di base per l'altezza:*

$$\mathcal{V} = \frac{1}{3}\mathcal{A}_b h. \tag{10.6}$$

Teorema 10.39. *Il volume di una sfera è uguale al prodotto di* $4\pi/3$ *per il cubo del raggio:*

$$\mathcal{V} = \frac{4}{3}\pi r^3. \tag{10.7}$$

Teorema 10.40. *Il volume di un segmento sferico di altezza* h *e raggi di base* r_1 *(eventualmente nullo nel caso di una sola base) e* r_2 *è dato da*

$$\mathcal{V} = \frac{\pi h}{6}(h^2 + 3r_1^2 + 3r_2^2). \tag{10.8}$$

Per il calcolo del volume di uno spicchio sferico si può usare una proporzione analoga alla (10.4)

$$\frac{4}{3}\pi r^3 : \mathcal{V} = 360 : \alpha. \tag{10.9}$$

10.6. Altre formule di geometria solida

Per concludere questa sommaria introduzione alla stereometria riportiamo alcune altre formule di geometria solida utili in molte applicazioni.

Cominciamo con le formule relative ai poliedri regolari, per i quali usiamo le seguenti convenzioni:

— l: lunghezza di uno spigolo;
— $\mathcal{V}$: volume;
— $\mathcal{A}_\mathrm{T}$: area della superficie totale;
— R: raggio della sfera circoscritta;
— r: raggio della sfera inscritta.

Tetraedro regolare

$$\mathcal{V} = l^3\frac{\sqrt{2}}{12}; \quad \mathcal{A}_\mathrm{T} = l^2\sqrt{3}; \quad R = l\frac{\sqrt{6}}{4}; \quad r = l\frac{\sqrt{6}}{12}. \tag{10.10}$$

Esaedro regolare (cubo)

$$\mathcal{V} = l^3; \quad \mathcal{A}_\mathrm{T} = 6l^2; \quad R = l\frac{\sqrt{3}}{2}; \quad r = \frac{l}{2}. \tag{10.11}$$

Ottaedro regolare

$$\mathcal{V} = l^3\frac{\sqrt{2}}{3}; \quad \mathcal{A}_\mathrm{T} = 2l^2\sqrt{3}; \quad R = l\frac{\sqrt{2}}{2}; \quad r = l\frac{\sqrt{6}}{6}. \tag{10.12}$$

Dodecaedro regolare

$$\mathcal{V} = l^2\frac{15+7\sqrt{5}}{4}; \quad \mathcal{A}_\mathrm{T} = 3l^2\sqrt{5(5+2\sqrt{5})}; \quad R = l\frac{\sqrt{3}(1+\sqrt{5})}{4}; \quad r = l\frac{\sqrt{250+110\sqrt{5}}}{20}. \tag{10.13}$$

Icosaedro regolare

$$\mathcal{V} = 5l^3\frac{3+\sqrt{5}}{12}; \quad \mathcal{A}_\mathrm{T} = 5l^2\sqrt{3}; \quad R = l\frac{\sqrt{10+2\sqrt{5}}}{4}; \quad r = l\frac{\sqrt{3}(3+\sqrt{5}}{12}. \tag{10.14}$$

Concludiamo con altri due solidi.

Cilindro obliquo

Con riferimento alla figura 10.23, indichiamo con $h_{\max}$ e $h_{\min}$ l'altezza maggiore e minore rispettivamente e con r il raggio della base circolare. Si ha

$$\mathcal{V} = \pi r^2 \frac{h_{\max} + h_{\min}}{2}, \tag{10.15}$$

per il volume e

$$\mathcal{A}_{\mathsf{L}} = \pi r (h_{\max} + h_{\min}); \quad \mathcal{A}_{\text{base circolare}} = \pi r^2; \quad \mathcal{A}_{\text{ellisse superiore}} = \pi r \sqrt{r^2 + \frac{(h_{\max} - h_{\min})^2}{4}}, \tag{10.16}$$

per l'area della superficie laterale, della base circolare (formula ovvia!) e della ellisse superiore rispettivamente.

Tronco di cono

Anche se il volume del tronco di cono si può ricavare per differenza dei volumi di due coni, può essere utile la formula seguente.

$$\mathcal{V} = \pi h \frac{r^2 + Rr + R^2}{3}. \tag{10.17}$$

11. Goniometria e trigonometria

Le funzioni elementari fin qui trattate (potenze, esponenziali, logaritmi, ecc.) hanno potuto essere introdotte in maniera abbastanza rigorosa rimanendo nell'ambito delle proprietà dei numeri reali. Il discorso è completamente diverso per le funzioni circolari o trigonometriche[(1)] che costituiscono l'oggetto del presente capitolo: per introdurle in maniera elementare occorre fare riferimento a concetti facilmente intuibili, ma difficili da giustificare in maniera rigorosa, per esempio quello di *verso orario* e *verso antiorario* o quello di verso di percorrenza in una circonferenza.

In realtà le funzioni circolari sono strettamente legate alle funzioni esponenziali, ma questo legame può essere riconosciuto solo lavorando nel campo dei numeri complessi, e questo esula dagli scopi di questo testo.

La trattazione che faremo è dunque una trattazione *ingenua* e anche i grafici che proporremo vanno prese quasi come dogmi, anche se geometricamente ci si può rendere conto del fatto che sono effettivamente come li faremo.

11.1. Angoli e loro misura

Abbiamo già introdotto, nella geometria euclidea piana, il concetto di angolo come una parte di piano individuata da due semirette aventi l'origine in comune, unita con le due semirette: vedi la definizione 9.5 nella pagina 297. In trigonometria è però più conveniente dare una definizione diversa, in modo da consentire anche angoli maggiori di un angolo giro, o angoli negativi. Cominciamo preventivamente a orientare il piano, ovvero il verso delle rotazioni nel piano. Fissato nel piano un sistema di coordinate cartesiane accettiamo come *positivo* il verso che porta il semiasse positivo delle ascisse a sovrapporsi a quello positivo delle ordinate secondo un angolo convesso (retto se il sistema è ortogonale e monometrico). Nelle situazioni usuali, cui faremo sempre riferimento in seguito salvo diverso avviso, il verso positivo è quello *antiorario*. Il verso opposto, cioè quello *orario*, sarà quello *negativo*.

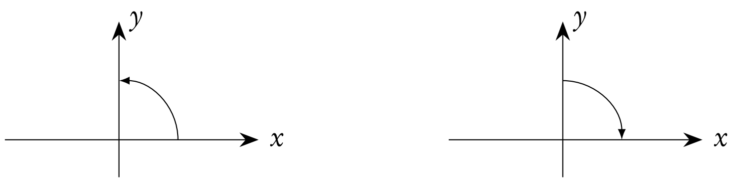


Figura 11.1.: *Orientazione positiva, a sinistra, e negativa, a destra, delle rotazioni nel piano*

[1]Nel seguito i nomi "funzioni circolari" e "funzioni trigonometriche" saranno considerati sinonimi. In realtà alcuni usano i due nomi con significato diverso: precisamente si parla di funzioni circolari quando la variabile è un numero reale, senza riferimento ad angoli o archi, si parla di funzioni trigonometriche quando la variabile è un angolo o un arco.

Consideriamo ora un angolo $\mathsf{A\widehat{V}B}$ e uno dei suoi due lati, per esempio VA. Possiamo pensare all'angolo anche come alla regione di piano descritta dalla semiretta scelta, in una rotazione attorno al vertice, per sovrapporsi all'altro lato, in questo caso VB.

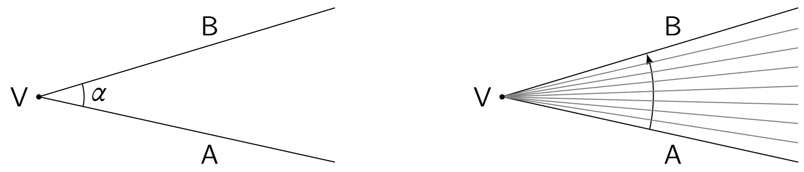


Figura 11.2.: *Angolo e rotazione di uno dei suoi lati*

Nell'ambito della geometria euclidea nulla cambierebbe se scegliessimo il lato VB e lo facessimo ruotare, in verso opposto, fino a sovrapporsi al lato VA. Se però abbiamo orientato il piano, possiamo confrontare il verso di rotazione con quello scelto come positivo e distinguere le due diverse situazioni ottenute dalla rotazione di VA verso VB e di VB verso VA: se il verso positivo è quello antiorario, diremo che la prima situazione corrisponde ad un angolo positivo, la seconda ad un angolo negativo e parleremo di *angolo orientato*. Ciò equivale in sostanza ad attribuire un ordine ai due lati dell'angolo, in modo da distinguere un primo e un secondo lato. Dunque un angolo nel senso della geometria elementare è "supporto" di due angoli orientati in verso opposto.

Una volta sostituito al concetto di angolo come parte di piano quello di angolo come rotazione di una semiretta attorno alla sua origine, possiamo anche pensare che queste rotazioni siano, sia in senso positivo che in senso negativo, maggiori di un angolo giro, esattamente come succede ad un'asta rigida incernierata per un suo estremo: potremo avere dunque anche angoli maggiori di un angolo giro e questo fatto è di grande importanza nello studio della trigonometria. Per distinguere questi angoli da quelli della geometria euclidea si potrebbe usare una diversa nomenclatura, per esempio quella di *angolo generalizzato*. In realtà il mantenimento della stessa nomenclatura non crea alcuna confusione e dunque possiamo concludere con la seguente definizione.

Definizione 11.1. *Data una semiretta di origine* V*, si chiama* angolo *la rotazione della semiretta attorno alla sua origine, fino a sovrapporsi ad una seconda semiretta con la stessa origine. Le due semirette si chiamano* lati *dell'angolo, rispettivamente* primo *e* secondo *lato, la comune origine si chiama* vertice*. Se la rotazione avviene in senso concorde*[2] *all'orientazione scelta sul piano, l'angolo si dice* positivo*, altrimenti si dice* negativo*.*

Anche questi angoli generalizzati sono indicati con le lettere minuscole dell'alfabeto greco.

Osserviamo che due semirette del piano aventi la stessa origine, senza considerare il loro ordine, individuano infiniti angoli positivi e infiniti angoli negativi, alcuni dei quali sono rappresentati nella figura 11.3 dove conviene osservare che nei primi due casi il primo lato è VA, nel terzo è VB.

Si noti ancora che, date in un certo ordine due semirette di comune origine V, la prima semiretta può sovrapporsi alla seconda una prima volta compiendo una rotazione in un senso e successivamente compiendo ancora successive rotazioni complete (cioè di un angolo giro) sia in un verso che nel verso

[2] Si noti che non si può parlare in senso assoluto di angolo positivo o negativo. Il concetto ha senso solo se si confronta il verso di rotazione nell'angolo con quello scelto nel piano.

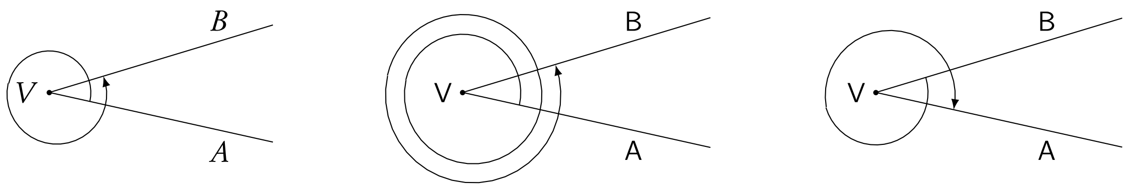


Figura 11.3.: *Diversi angoli con gli stessi lati*

opposto. Si veda la figura 11.4 dove, a differenza della 11.3, la rotazione inizia sempre dalla semiretta individuata come primo lato.

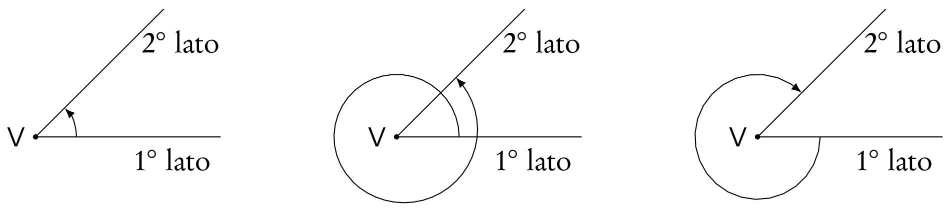


Figura 11.4.: *Diversi angoli individuati da due semirette, prese in un dato ordine*

Con questa nuova definizione di angolo avranno sempre senso la somma e sottrazione di due angoli, senza alcuna limitazione.

Per la misura degli angoli si può continuare ad usare il grado sessagesimale: ora gli angoli potranno avere una qualunque misura positiva o negativa in gradi sessagesimali e alla somma o sottrazione di angoli corrisponderà la somma o sottrazione delle rispettive misure.

Sulla base delle osservazioni fatte, due semirette, date in un certo ordine, individuano una famiglia di angoli aventi come misura un certo valore $\alpha°$, compreso tra $-360°$ e $360°$, a cui si può aggiungere un arbitrario multiplo intero di $360°$. Si veda un esempio nella figura 11.4 con riferimento alla quale, se α ($0° < \alpha < 360°$) è la misura del primo angolo a sinistra, quello al centro ha misura $\alpha° + 360°$, quello a destra $\alpha° - 360°$.

Di solito, e lo faremo anche noi nel seguito, in mancanza di precisazione, parlando dell'angolo individuato da due semirette aventi l'origine in comune faremo sempre riferimento all'angolo positivo, minore di un angolo giro, cioè all'angolo nel senso della geometria euclidea.

11.1.1. La misura in radianti

In quasi tutte le applicazioni della trigonometria (non solo in matematica) la misura degli angoli in gradi sessagesimali non è soddisfacente, e il motivo essenziale, ma non l'unico, di questo fatto è collegato ai discorsi che abbiamo fatto a proposito delle figure 8.2 e seguenti, nella pagina 208. Si preferisce invece dare la definizione seguente.

Definizione 11.2. *Consideriamo un angolo α di vertice* V *e lati* VA *e* VB*, nell'ordine. Preso un punto* P *sul primo lato, nella rotazione che porta questo primo lato a sovrapporsi al secondo, esso descrive un arco di circonferenza di raggio $r = \overline{\mathsf{VP}}$, eventualmente più lungo dell'intera circonferenza, arco che prenderemo con* *il segno positivo o negativo, a seconda che l'angolo sia positivo o negativo. Si chiama* misura in radianti *dell'angolo α, e si indica con α^{r}, o semplicemente con α, il rapporto tra la lunghezza dell'arco descritto da* P *e il raggio della circonferenza:*

$$\alpha^{r} = \frac{arco}{raggio} \ (= \alpha). \tag{11.1}$$

Si noti che, trattandosi di un rapporto tra due grandezze omogenee, il risultato è un "numero puro", cioè privo di dimensioni e non dipende nemmeno dall'unità scelta per la misura delle lunghezze. Per le note proprietà degli archi di circonferenza, questo rapporto è indipendente dal raggio della circonferenza: se si conviene di prendere una circonferenza di raggio unitario, la misura in radianti dell'angolo coinciderà, a meno del segno, con la lunghezza dell'arco. Se si tratta di un angolo nel senso della geometria elementare, la sua misura sarà un numero reale compreso tra 0 e 2π.

Il passaggio dalla misura in gradi alla misura in radianti, o viceversa, di uno stesso angolo potrà essere fatto utilizzando la seguente proporzione:

$$\alpha^{\circ} : \alpha = 180^{\circ} : \pi. \tag{11.2}$$

Per gli angoli più importanti della geometria piana si può costruire la tabella 11.1.

Tabella 11.1.: *Misure in gradi e radianti di alcuni angoli importanti nella geometria piana*

gradi	0°	18°	30°	36°	45°	60°	90°	120°	135°	150°	180°	270°	360°
radianti	0	$\frac{\pi}{10}$	$\frac{\pi}{6}$	$\frac{\pi}{5}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	π	$\frac{3\pi}{2}$	2π

In relazione alla misura in radianti degli angoli e alla definizione delle funzioni trigonometriche, è tradizione introdurre nel piano un sistema cartesiano ortogonale monometrico e tracciare, in questo sistema, la circonferenza di centro l'origine e raggio 1, circonferenza che viene detta *circonferenza goniometrica*. Detto A il punto di coordinate $(1, 0)$, e considerato un angolo qualsiasi individuato dai suoi due lati (presi in un determinato ordine), si trasporta rigidamente l'angolo in modo che il primo lato coincida con il semiasse positivo delle ascisse: il secondo lato intersecherà la circonferenza goniometrica in un punto P. La misura in radianti dell'angolo coincide con la lunghezza con segno dell'arco orientato di circonferenza $\overset{\frown}{AP}$, tenendo eventualmente conto anche del numero di circonferenze complete percorse, in un verso o in quello opposto. In sostanza quello che conta in questa rappresentazione è solo l'intersezione del secondo lato dell'angolo con la circonferenza goniometrica. Nel seguito, rappresentando gli angoli sulla circonferenza goniometrica, evidenzieremo sempre solo il secondo lato, essendo sottinteso che il primo lato coincide con il semiasse positivo delle ascisse.

Con riferimento solo agli angoli positivi e non superiori all'angolo giro si può costruire una figura come la 11.5, che sostituisce efficacemente la tabella 11.1 e, in parte, la integra.

Per la risoluzione degli esercizi, in particolare delle equazioni e disequazioni goniometriche, è molto utile imparare a valutare "a colpo d'occhio" sulla circonferenza goniometrica i multipli degli angoli di $\pi/2$, $\pi/4$, $\pi/3$,$\pi/6$. Si possono tenere presenti, a questo proposito, le figure 11.6 e 11.7.

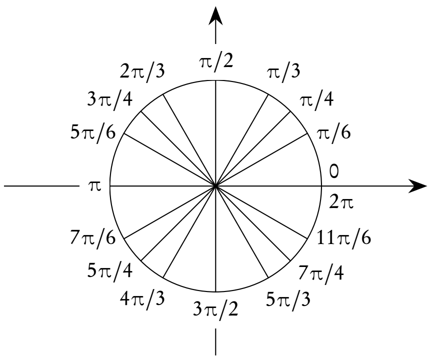


Figura 11.5.: *Circonferenza goniometrica e misure in radianti di alcuni angoli*

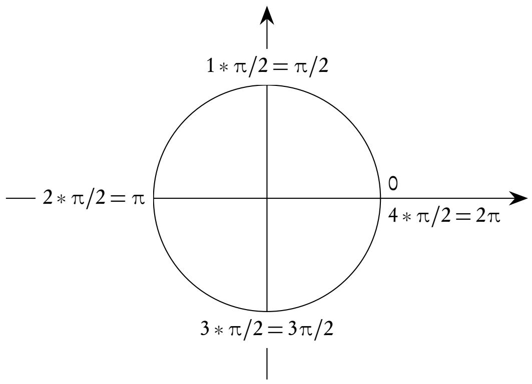


Figura 11.6.: *Visualizzazione, sulla circonferenza goniometrica, di alcuni multipli di* $\pi/2$ *e* $\pi/4$

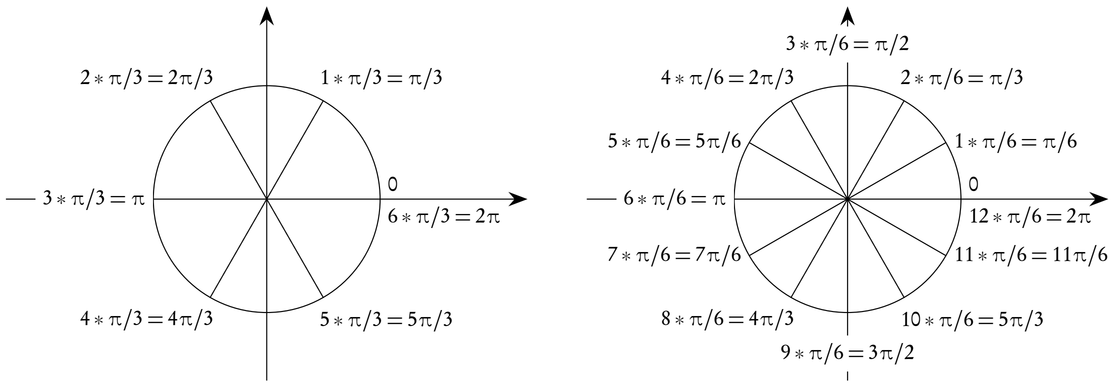


Figura 11.7.: *Visualizzazione, sulla circonferenza goniometrica, di alcuni multipli di* $\pi/3$ *e* $\pi/6$

Osservazione 11.3. Si tenga ben presente che, essendo π un numero irrazionale, $\pi \simeq 3.14159\ldots$, tutti gli angoli considerati nella tabella 11.1 e nelle figure 11.5, 11.6 e 11.7, ad eccezione dell'angolo nullo, hanno una misura in radianti espressa da un numero irrazionale (anzi irrazionale trascendente). È questo un palese svantaggio della misura in radianti rispetto alla misura in gradi sessagesimali, in quanto gli stessi angoli, cioè gli angoli importanti della geometria piana, in gradi sessagesimali hanno una misura espressa addirittura da un numero intero. Tuttavia, come già osservato, la misura in radianti ha una serie di vantaggi, che potranno essere ben apprezzati per esempio nei successivi corsi di analisi, che bilanciano largamente questa difficoltà.

Tenendo conto della proporzione (11.2), si perviene facilmente alla conclusione che l'angolo di 1 radiante corrisponde all'angolo che ha, in gradi, una misura di circa $57,2978°$, ovvero circa $57°\,17'\,45''$. Viceversa la misura in radianti di un angolo di $1°$ è di circa 0.0175.

Anche se non si tratta di angoli di uso comune nella geometria, è utile imparare a valutare, sulla circonferenza goniometrica, la posizione degli angoli che hanno una misura in radianti espressa da un numero intero. Nella figura 11.8 sono riportati, sulla circonferenza goniometrica, gli angoli di 1, 2, 3, 4, 5, 6 radianti. Poiché saremo interessati anche alla circonferenza rettificata, sono riportati sull'asse delle ascisse di un sistema cartesiano ortogonale i multipli interi di mezzo radiante e, per un utile confronto, anche i valori degli angoli multipli di $\pi/2$.

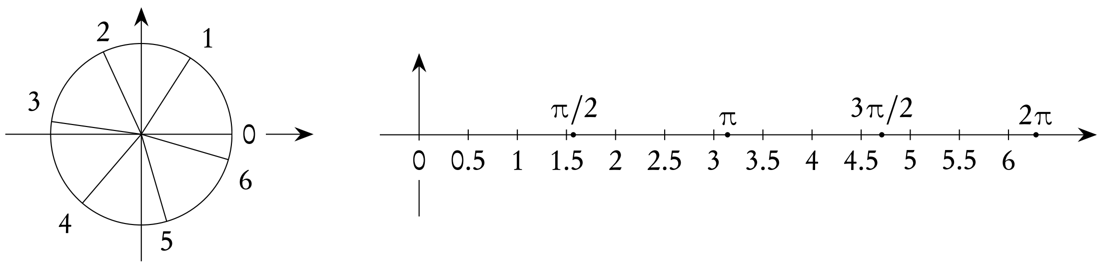


Figura 11.8.: *Multipli di un radiante sulla circonferenza goniometrica e di mezzo radiante sulla sua rettificazione*

Nella definizione e nello studio delle proprietà delle funzioni circolari il concetto di angolo ha un'importanza relativa: quello che conta è il fatto che, con la costruzione che abbiamo effettuato, ad ogni punto della circonferenza goniometrica possono essere associati infiniti numeri reali, tali che la differenza tra due di essi sia un multiplo intero di 2π. In sostanza, data in un piano in cui si sia stabilito un riferimento cartesiano ortogonale monometrico la circonferenza γ di centro l'origine e raggio 1, abbiamo costruito una funzione $\varrho\colon \mathbb{R} \to \gamma$ nel seguente modo: se x è un numero reale $\varrho(x)$ è quel punto P di γ che si ottiene partendo da $\mathsf{A} = (1,0)$ e percorrendo su γ un arco lungo $|x|$, in senso antiorario se $x > 0$, in senso orario se $x < 0$. È chiaro che $\varrho(x) = \varrho(x + 2k\pi)$, $\forall k \in \mathbb{Z}$. Il punto P ottenuto individua il secondo lato di un angolo che ha come primo lato il semiasse positivo delle ascisse, e x è la misura di questo angolo: tuttavia quello che conta nello studio delle funzioni circolari è il valore di x di cui P è immagine tramite questa funzione e in molte applicazioni, per esempio alla fisica, il fatto che x sia anche la misura di un angolo ha poco interesse. Questa funzione si chiama, in maniera molto suggestiva, *avvolgimento della retta attorno a un circolo*.

Nel seguito, dato un numero reale x, avremo sempre a che fare con il punto P ad esso corrispondente sulla circonferenza goniometrica: poiché questo punto individua, nella maniera già detta, un angolo, potremo anche parlare di "angolo x", invece che semplicemente di numero reale x. Addirittura, poiché nell'avvolgimento della retta attorno al circolo unitario questi numeri coincidono con la misura di archi, parleremo spesso di "arco x". Inoltre, poiché tradizionalmente gli angoli sono indicati con le lettere minuscole dell'alfabeto greco, al posto di x useremo spesso α, β, ecc.

11.2. Funzioni periodiche

Anche se questo argomento è di carattere generale e non riguarda solamente la trigonometria, lo trattiamo in questo capitolo, in quanto le funzioni trigonometriche che tra poco definiremo sono le più importanti funzioni periodiche. Abbiamo comunque già incontrato, come esempio di funzione periodica, la funzione $f(x)=x-\lfloor x\rfloor$, nella pagina 127.

Definizione 11.4. *Sia $f: D \subseteq \mathbb{R} \to \mathbb{R}$ una funzione reale di variabile reale. Se esiste un $p \in \mathbb{R}$, strettamente positivo, tale che, per ogni $x \in D$, $x \pm p \in D$ e*

$$f(x+p)=f(x), \tag{11.3}$$

allora p si dice un periodo *per f.*

Essendo $f(x+2p)=f(x+p+p)=f(x+p)=f(x)$, se p è un periodo, anche $2p$ lo è, e quindi anche np, per ogni naturale n. Da ciò segue che il dominio di una funzione che ha un periodo non può essere limitato.

Essendo poi $f(x)=f(x-p+p)=f(x-p)$, la definizione si può anche scrivere $f(x \pm p)=f(x)$.

Geometricamente la definizione data implica che il grafico di una funzione avente un periodo si può tracciare per ripetizione del grafico ottenuto restringendo il dominio a un qualunque intervallo di ampiezza p. Si veda la figura 11.9 per un esempio.

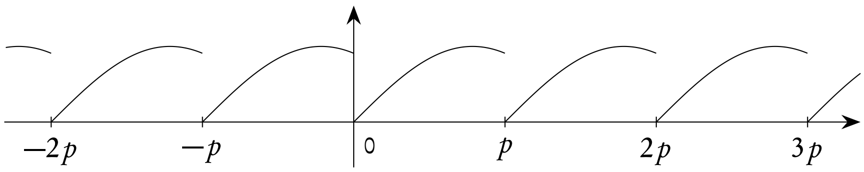


Figura 11.9.: *Un esempio di funzione con periodi*

Esempio 11.1. $f(x)=k$ (funzione costante). Ogni numero reale strettamente positivo è, banalmente, un periodo.

Esempio 11.2. $f(x)=x-\lfloor x\rfloor$, già considerata nella pagina 127. Essa ha come periodo un qualunque numero naturale positivo.

Esempio 11.3. $f(x)=\begin{cases} 1, & \text{se } x \in \mathbb{Q}; \\ 0, & \text{se } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$

Tutti i numeri razionali strettamente positivi sono un periodo per questa funzione. Per provarlo basta osservare che la somma di due razionali è ancora un razionale, mentre la somma di un razionale con un irrazionale è un irrazionale.

Come si vede dagli esempi, e come del resto si deduce dalla definizione, una funzione, se ha un periodo, ne ha infiniti. Consideriamo allora l'insieme di tutti i periodi di una funzione: poiché si tratta di un insieme di reali strettamente positivi, esso può avere un minimo (strettamente positivo) o solo un estremo inferiore (maggiore o uguale a zero). Negli esempi precedenti le funzioni dell'esempio 11.1 e 11.3 non ammettono un minimo periodo, quella dell'esempio 11.2 ha come minimo periodo 1.

Definizione 11.5. *Se l'insieme dei periodi di una funzione ha un minimo, esso si chiama* minimo periodo, *o semplicemente* periodo. *In questo caso la funzione si dice* periodica.

Una funzione che, pur avendo periodi, non ha un minimo periodo, non viene abitualmente considerata periodica. Torneremo ancora sulle funzioni periodiche, dopo l'introduzione delle funzioni trigonometriche.

11.2.1. Prolungamento per periodicità

In molte situazioni è utile prolungare una funzione definita su un sottoinsieme A di $\mathbb{R}$, a un soprainsieme di A, e magari a tutto $\mathbb{R}$. È ovvio che tale prolungamento si può fare in infiniti modi, alcuni dei quali (per esempio il prolungamento per continuità) saranno oggetto dei successivi corsi di analisi. Qui ci occupiamo di una situazione particolare e precisamente il prolungamento per periodicità di una funzione, definita su un intervallo di $\mathbb{R}$, a tutto $\mathbb{R}$. Cominciamo con un esempio.

Esempio 11.4. Sia $f\colon [-1,1[\to \mathbb{R}\ f(x)=x^2$, una funzione. A partire da essa possiamo costruire una funzione definita su tutto $\mathbb{R}$ semplicemente "riproducendo all'infinito" a destra e sinistra di $[-1,1[$, i valori della funzione. Otterremo così una funzione che ha 2 come minimo periodo. Il suo grafico è rappresentato nella figura 11.10.

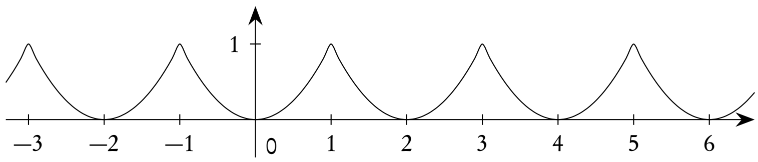


Figura 11.10.: *Un esempio di prolungamento per periodicità*

Si tenga presente che non è detto che l'ampiezza dell'intervallo di definizione di f debba essere il minimo periodo del prolungamento: esso è generalmente solo un periodo. Un esempio banale è costituito da una funzione costante su un intervallo I: se la si prolunga per periodicità su tutto $\mathbb{R}$ si ottiene una funzione costante che, come già osservato, non ha un minimo periodo.

Utilizzando la funzione floor considerata nel paragrafo 4.8.9 nella pagina 126, è possibile scrivere una espressione esplicita per il prolungamento per periodicità di una funzione. Precisamente si ha il seguente risultato.

Teorema 11.6. *Siano $\tau>0$ e a numeri reali. Sia inoltre data una funzione $f\colon [a,a+\tau[\to\mathbb{R}$. Esiste un'unica funzione $g\colon\mathbb{R}\to\mathbb{R}$ che prolunga f e che abbia tra i suoi periodi il numero τ. Tale g è data dalla formula*

$$g(x)=f\left(x-\tau\left\lfloor\frac{x-a}{\tau}\right\rfloor\right),\quad x\in\mathbb{R}. \tag{11.4}$$

11.3. Le funzioni seno e coseno

Definizione 11.7. *Dato un numero reale x, eventualmente pensato come misura di un angolo, consideriamo il punto $\mathsf{P}=(x_\mathsf{P},y_\mathsf{P})$ ad esso corrispondente sulla circonferenza goniometrica. L'ascissa x_P di P si chiama* coseno *del numero reale* x*, l'ordinata* y_P *di* P *si chiama* seno *del numero reale* x *e si scrive*

$$\cos(x) = x_\mathsf{P}, \quad \sin(x) = y_\mathsf{P}. \tag{11.5}$$

Di solito, quando non c'è rischio di ambiguità, si scrive semplicemente $\cos x$ e $\sin x$, tralasciando le parentesi.

La definizione data introduce due nuove funzioni reali di variabile reale, che rientrano nell'ambito delle funzioni elementari. Segue subito dalla definizione che le immagini di queste funzioni soddisfano alle limitazioni seguenti:

$$-1 \le \sin x \le 1, \quad -1 \le \cos x \le 1. \tag{11.6}$$

Inoltre dalle osservazioni fatte sugli angoli generalizzati o sull'avvolgimento della retta attorno al circolo seguono le proprietà seguenti:

$$\sin(x + 2k\pi) = \sin x, \quad \cos(x + 2k\pi) = \cos x. \tag{11.7}$$

Dalla (11.7), e dalla definizione, segue che le funzioni seno e coseno sono periodiche di minimo periodo 2π.

Dalla definizione data e dall'equazione della circonferenza goniometrica nel piano cartesiano discende subito anche la

$$\sin^2 x + \cos^2 x = 1, \tag{11.8}$$

detta *identità fondamentale* della trigonometria.

Osservazione 11.8. In analogia con quanto fatto nell'osservazione 7.6, segnaliamo che le scritture $\sin^2 x$ e $\cos^2 x$ sono un po' ambigue, ma ormai entrate nell'uso comune. Esse devono essere interpretate come $(\sin x)^2$ e $(\cos x)^2$. In particolare occorre tenere ben presente che $\sin x^2 \ne \sin^2 x$ e $\cos x^2 \ne \cos^2 x$.

Tenendo conto del fatto che i numeri reali x di cui vogliamo calcolare il seno e il coseno si possono anche pensare come misure di angoli, e ricordando note proprietà dei triangoli equilateri e rettangoli isosceli, è facile ricavare il seno e il coseno per i valori $\pi/6$, $\pi/4$ e $\pi/3$: si veda la figura 11.11.

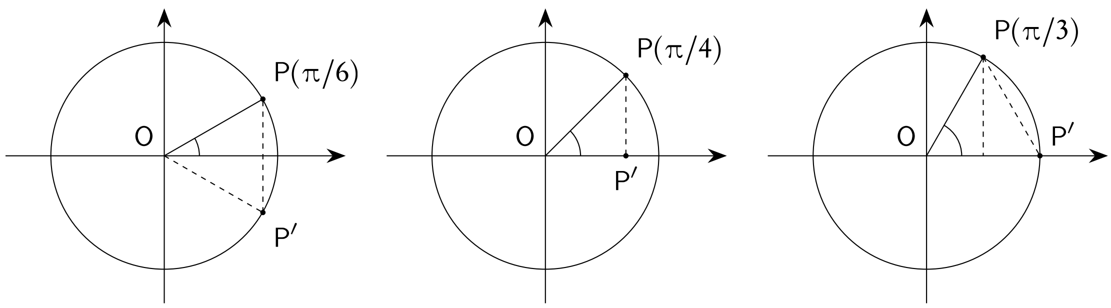


Figura 11.11.: *Determinazione di seno e coseno di* $\pi/6$*,* $\pi/4$ *e* $\pi/3$

Non c'è nessun problema per i valori di seno e coseno per $x = 0$ e $x = \pi/2$. Si perviene facilmente alla costruzione della tabella 11.2.

Tabella 11.2.: *Seno e coseno di alcuni valori particolari*

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin x$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\cos x$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0

Per valori di x maggiori di $\pi/2$ e minori di 2π si possono invocare le simmetrie che si evidenziano dalla figura 11.5. Per valori più grandi di 2π, o negativi, basta invocare la periodicità.

Tenendo conto dei valori della tabella 11.2 si possono riportare sul piano cartesiano alcune coppie di punti $(x, \sin x)$ con lo scopo di tracciare il grafico della funzione seno. Nella figura 11.12 abbiamo evidenziato il procedimento, completando poi il grafico in tutto l'intervallo $[0, \pi/2]$: come già detto, anche se ci si può rendere conto geometricamente della correttezza del grafico, lo prendiamo in questa fase quasi come un dogma di fede.

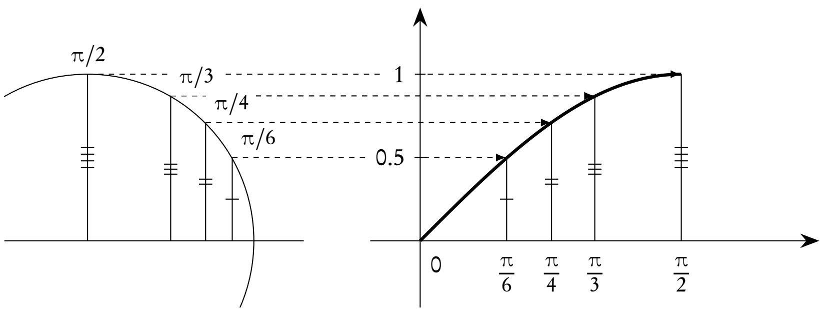


Figura 11.12.: *Costruzione per punti della funzione* $f(x) = \sin x$

Procedimento analogo per la funzione coseno. Tenendo poi conto delle simmetrie si può completare il grafico nel tratto $[0, 2\pi]$ e successivamente prolungarlo per periodicità su tutto l'asse reale. I grafici ottenuti si chiamano *sinusoide* e *cosinusoide* e sono rappresentati nella figura 11.13.

Si noti che il grafico del coseno è ottenuto traslando di $\pi/2$ verso sinistra il grafico del seno.

A proposito dei grafici delle funzioni seno e coseno, e lo stesso discorso vale per le successive funzioni trigonometriche, è opportuno osservare che essi sono tracciati in un sistema cartesiano ortogonale *monometrico*. Come già segnalato a proposito della figura 8.2 nella pagina 208, il grafico ha l'aspetto della figura 11.13 se i valori sull'asse delle ascisse rappresentano misure degli angoli in radianti e non in gradi. Come si vedrà nei successivi corsi di analisi, tutto questo ha come conseguenza che il coefficiente angolare della tangente al grafico di $f(x) = \sin x$, nell'origine, ha coefficiente angolare 1, cioè coincide con la bisettrice del primo e terzo quadrante. Volendo rappresentare invece la funzione $\sin x°$, il coefficiente angolare della tangente al grafico, sempre nell'origine, sarebbe $\pi/180$. Tutto questo risulta evidente se si paragona la figura 11.13 con la 11.14, figura che peraltro abbiamo già considerato nella pagina 208.

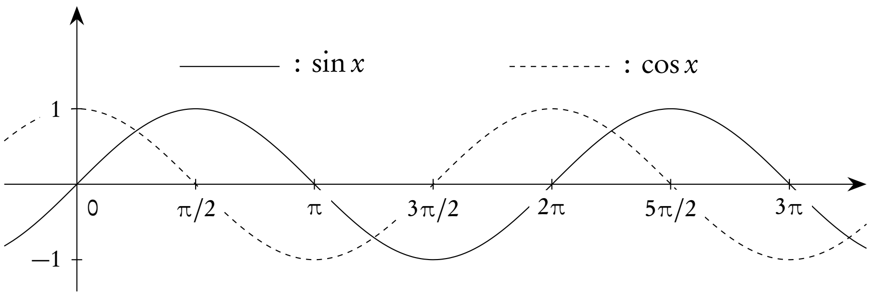


Figura 11.13.: *Grafici delle funzioni seno e coseno (angoli in radianti!)*

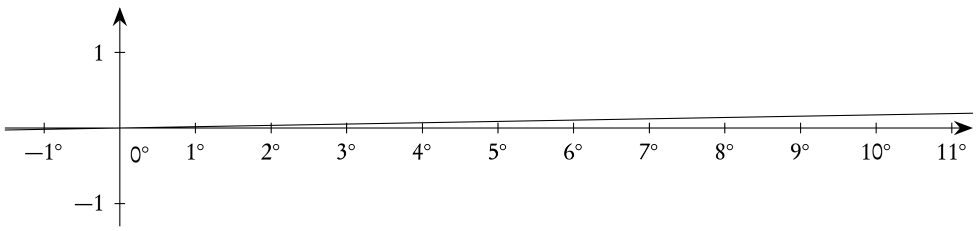


Figura 11.14.: *La funzione $f(x^\circ) = \sin(x^\circ)$, in un sistema monometrico*

11.4. Le funzioni tangente e cotangente

Definizione 11.9. *Si chiama* tangente *di un numero reale il rapporto, quando esiste, del seno e del coseno dello stesso numero; si chiama invece* cotangente *di un numero reale il rapporto, quando esiste, del coseno e del seno dello stesso numero. Si scrive:*

$$\tan x = \frac{\sin x}{\cos x}; \quad \cot x = \frac{\cos x}{\sin x}. \tag{11.9}$$

È chiaro che la tangente risulta definita se $x \neq \pi/2 + k\pi$, $k \in \mathbb{Z}$, mentre la cotangente risulta definita se $x \neq k\pi$, $k \in \mathbb{Z}$.

Si noti che si ha

$$\tan x = \frac{1}{\cot x}; \quad \cot x = \frac{1}{\tan x}, \quad \text{se} \quad x \neq k\frac{\pi}{2},\ k \in \mathbb{Z}. \tag{11.10}$$

Si presti particolare attenzione alle limitazioni necessarie per la validità delle (11.10). In particolare si noti che *non è sempre vero* che $\cot x = 1/\tan x$ o che $\tan x = 1/\cot x$.

Le funzioni tangente e cotangente possono anche essere definite con riferimento alla circonferenza goniometrica, avente centro nel punto $\mathsf{O} = (0,0)$, come per le funzioni seno e coseno. Precisamente, dette t la tangente in $(1,0)$ e s la tangente in $(0,1)$ alla circonferenza goniometrica, e considerato il punto P individuato dal numero reale x sulla circonferenza stessa, siano T e C le intersezioni, quando esistono, della retta OP con t ed s rispettivamente. Si ha

$$y_{\mathsf{T}} = \tan x, \quad x_{\mathsf{C}} = \cot x. \tag{11.11}$$

Si veda in proposito la figura 11.15.

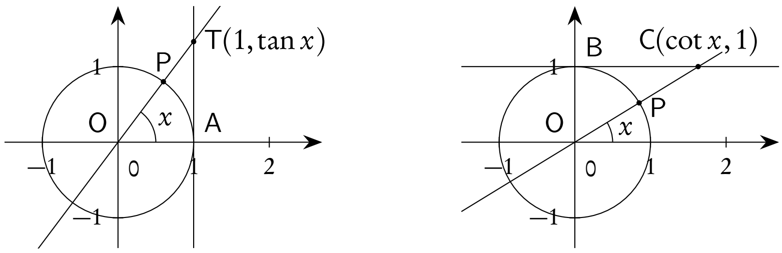


Figura 11.15.: *Tangente e cotangente di un numero reale x*

Per i grafici delle funzioni tangente e cotangente si può procedere in modo simile a quanto già fatto con il seno e il coseno, determinando prima alcuni valori in corrispondenza di archi o angoli notevoli, a partire dalla tabella 11.2. Si ottengono i grafici delle figure 11.16 e 11.17 detti, rispettivamente, *tangentoide* e *cotangentoide*.

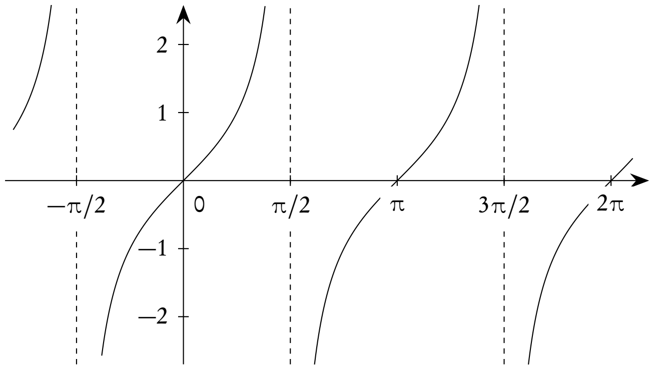


Figura 11.16.: *Grafico della funzione tangente*

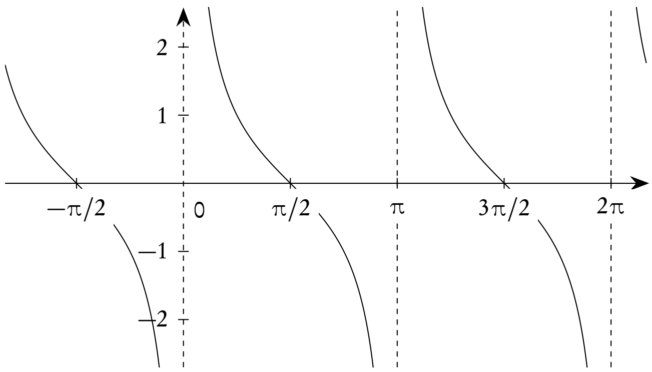


Figura 11.17.: *Grafico della funzione cotangente*

Le rette verticali di equazione

$$x = \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z} \tag{11.12}$$

sono *asintoti* per la tangentoide.

Le rette verticali di equazione

$$x = k\pi, \quad k \in \mathbb{Z} \tag{11.13}$$

sono *asintoti* per la cotangentoide.

Sia la funzione tangente che la funzione cotangente sono periodiche, con minimo periodo π.

11.5. Le funzioni secante e cosecante

In alcune circostanze hanno interesse le funzioni reciproche delle funzioni seno e coseno. Riportiamo qui, solo per completezza, le definizioni e i grafici.

Definizione 11.10. *Si chiama* secante *del numero reale x il reciproco del* $\cos x$, *se* $\cos x \neq 0$, *e si scrive*

$$\sec x = \frac{1}{\cos x}, \quad x \neq \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}. \tag{11.14}$$

Si chiama cosecante *del numero reale x il reciproco di* $\sin x$, *se* $\sin x \neq 0$, *e si scrive*

$$\csc x = \frac{1}{\sin x}, \quad x \neq k\pi,\ k \in \mathbb{Z}. \tag{11.15}$$

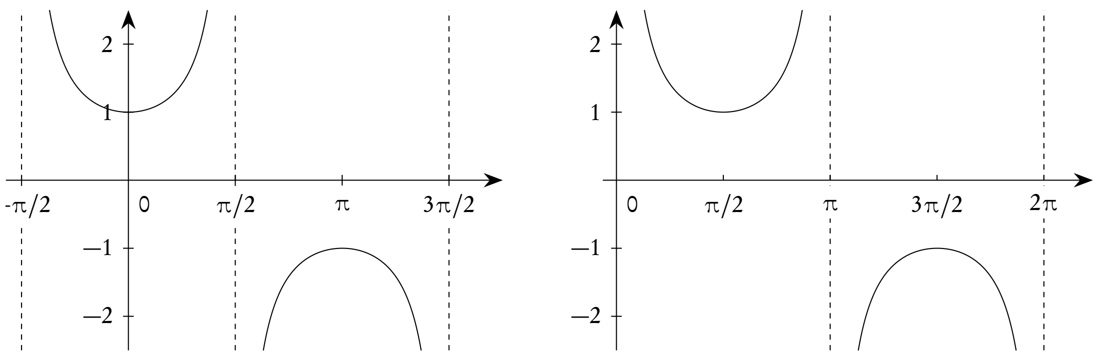


Figura 11.18.: *Grafici delle funzioni secante, a sinistra, e cosecante, a destra*

Le rette di equazione

$$x = \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z} \quad \text{e} \quad x = k\pi, \quad k \in \mathbb{Z}, \tag{11.16}$$

sono asintoti, rispettivamente per la funzione secante e la funzione cosecante.

Si noti, in analogia con quanto succede per la sinusoide e la cosinusoide, che il grafico della cosecante è ottenuto traslando di $\pi/2$ verso destra il grafico della secante.

Le funzioni secante e cosecante possono essere definite anche con riferimento alla circonferenza goniometrica. Precisamente, considerato il punto P individuato dal numero reale x sulla circonferenza, si consideri la retta tangente in P alla circonferenza stessa. Siano poi M e N, quando esistono, le intersezioni di questa tangente con l'asse delle ascisse e delle ordinate, rispettivamente. Si ha

$$x_{\mathsf{M}} = \sec x, \quad y_{\mathsf{N}} = \csc x. \tag{11.17}$$

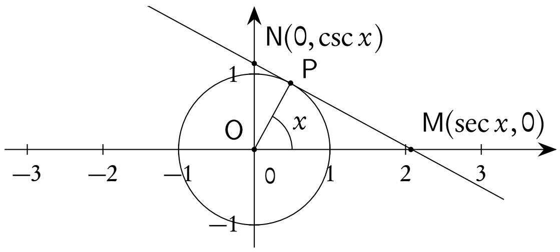


Figura 11.19.: *Secante e cosecante di un numero reale x*

11.6. Relazioni tra le funzioni trigonometriche

Tenendo conto della relazione fondamentale (11.8) e delle definizioni di tangente e cotangente si possono ricavare una serie di formule che permettono di esprimere tutte le funzioni trigonometriche in funzione di una sola di esse, a parte un'indeterminazione di segno che potrà essere risolta tenendo conto del quadrante cui appartiene il punto P associato al numero x sulla circonferenza goniometrica.

Si ha, per esempio,

$$\sin^2 x + \cos^2 x = 1 \quad \Rightarrow \quad \tan^2 x + 1 = \frac{1}{\cos^2 x} \quad \Rightarrow \quad \cos x = \pm\frac{1}{\sqrt{1+\tan^2 x}} \qquad \left(x \neq \frac{\pi}{2} + k\pi\right),$$

che permette di calcolare il coseno, una volta nota la tangente.

Procedendo in maniera analoga si possono ricavare le formule seguenti, dove non abbiamo esplicitato le condizioni per la validità. Esse esprimono tutte le funzioni mediante, rispettivamente, il seno, il coseno, la tangente, la cotangente. Non abbiamo riportato le formule coinvolgenti le funzioni secante e cosecante, perché di uso molto limitato.

$$\cos x = \pm\sqrt{1-\sin^2 x}; \quad \tan x = \pm\frac{\sin x}{\sqrt{1-\sin^2 x}}; \quad \cot x = \pm\frac{\sqrt{1-\sin^2 x}}{\sin x}. \tag{11.18}$$

$$\sin x = \pm\sqrt{1-\cos^2 x}; \quad \tan x = \pm\frac{\sqrt{1-\cos^2 x}}{\cos x}; \quad \cot x = \pm\frac{\cos x}{\sqrt{1-\cos^2 x}}. \tag{11.19}$$

$$\sin x = \pm\frac{\tan x}{\sqrt{1+\tan^2 x}}; \quad \cos x = \pm\frac{1}{\sqrt{1+\tan^2 x}}; \quad \cot x = \frac{1}{\tan x}. \tag{11.20}$$

$$\sin x = \pm\frac{1}{\sqrt{1+\cot^2 x}}; \quad \cos x = \pm\frac{\cot x}{\sqrt{1+\cot^2 x}}; \quad \tan x = \frac{1}{\cot x}. \tag{11.21}$$

Segnaliamo che non è opportuno memorizzare queste formule: all'occorrenza basterà applicare il procedimento usato nell'esempio precedente, opportunamente adattato al caso in esame.

Esempio 11.5. Sapendo che $\cos x = -2/3$ e che $\pi < x < 3\pi/2$, per calcolare $\sin x$ e $\tan x$ si osserva che $\sin x$ è negativo, mentre $\tan x$ è positivo. Si ottiene

$$\sin x = -\sqrt{1-\cos^2 x} = -\frac{\sqrt{5}}{3}; \quad \tan x = \frac{\sin x}{\cos x} = \frac{\sqrt{5}}{2}.$$

11.7. Formule trigonometriche

Le funzioni trigonometriche soddisfano alcune identità molto importanti che precisano quali legami intercorrono tra esse e le operazioni elementari, in particolare la somma, sui reali. Tranne che per le formule di addizione e sottrazione, proponiamo lo schema della dimostrazione per ottenere le successive, che sono tutte conseguenza delle formule di addizione e sottrazione: in effetti i passaggi per ottenerle sono molto semplici ed è più facile memorizzare questi passaggi che non le formule stesse.

Formule di addizione e sottrazione

Sono di gran lunga le più importanti ed anzi, come già detto, da esse conseguono tutte le altre. Queste formule, assieme all'identità fondamentale e a poche altre condizioni caratterizzano le funzioni trigonometriche, nel senso che queste ultime potrebbero essere introdotte quasi solo con la richiesta che valgano la relazione fondamentale e le formule di addizione e sottrazione.

$$\cos(\alpha\pm\beta)=\cos\alpha\cos\beta\mp\sin\alpha\sin\beta \quad \text{(formule di addizione e sottrazione per il coseno);} \tag{11.22}$$

$$\sin(\alpha\pm\beta)=\sin\alpha\cos\beta\pm\cos\alpha\sin\beta \quad \text{(formule di addizione e sottrazione per il seno);} \tag{11.23}$$

$$\tan(\alpha\pm\beta)=\frac{\tan\alpha\pm\tan\beta}{1\mp\tan\alpha\tan\beta} \quad \text{(formule di addizione e sottrazione per la tangente).} \tag{11.24}$$

Nelle formule per la tangente bisogna tenere conto delle condizioni $\alpha\neq \pi/2+k\pi$, $\beta\neq\pi/2+h\pi$, $\alpha\pm\beta\neq\pi/2+l\pi$, $h,k,l\in\mathbb{Z}$.

Come conseguenza di queste formule, ma anche con semplici considerazioni di simmetria, si possono ricavare le seguenti relazioni, di uso frequente e che coinvolgono alcune coppie di archi cosiddetti "associati".

$$\cos(\pi-x)=-\cos x \qquad \sin(\pi-x)=\sin x \tag{11.25}$$

$$\cos(\pi+x)=-\cos x \qquad \sin(\pi+x)=-\sin x \tag{11.26}$$

$$\cos\left(\frac{\pi}{2}-x\right)=\sin x \qquad \sin\left(\frac{\pi}{2}-x\right)=\cos x \tag{11.27}$$

$$\cos\left(\frac{\pi}{2}+x\right)=-\sin x \qquad \sin\left(\frac{\pi}{2}+x\right)=\cos x \tag{11.28}$$

Formule di duplicazione e triplicazione

Sostituendo, nelle formule di addizione, β con α, oppure α e β con $\alpha/2$, si ottengono le formule di duplicazione, che scriviamo in diverse forme equivalenti, tutte importanti nelle applicazioni.
allowdisplaybreaks

$$\sin 2\alpha=2\sin\alpha\cos\alpha, \quad \sin\alpha=2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}; \tag{11.29}$$

$$\cos 2\alpha=\cos^2\alpha-\sin^2\alpha=1-2\sin^2\alpha=2\cos^2\alpha-1, \quad \cos\alpha=\cos^2\frac{\alpha}{2}-\sin^2\frac{\alpha}{2}; \tag{11.30}$$

$$\tan 2\alpha=\frac{2\tan\alpha}{1-\tan^2\alpha}, \quad \alpha\neq\frac{\pi}{2}+k\pi. \tag{11.31}$$

Procedendo con la stessa tecnica si possono ricavare anche le formule per calcolare le funzioni trigonometriche di $n\alpha$, una volta note quelle di α. Riportiamo, perché spesso utilizzate, quelle di triplicazione per il seno e il coseno.

$$\sin 3\alpha = 3\sin\alpha - 4\sin^3\alpha; \tag{11.32}$$

$$\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha. \tag{11.33}$$

Formule di bisezione

Dalle formule di duplicazione si possono ottenere, con facili calcoli, le formule cosiddette di bisezione, che permettono di calcolare le funzioni circolari di $\alpha/2$, una volta note quelle di α. Si può partire dall'ultima delle (11.30) e dalla identità fondamentale scritta mediante $\alpha/2$.

$$\cos\alpha = \cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2} = 1 - 2\sin^2\frac{\alpha}{2}, \quad \text{e} \quad \cos\alpha = \cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2} = 2\cos^2\frac{\alpha}{2} - 1.$$

Dalla prima si può ricavare $\sin\alpha/2$, dalla seconda $\cos\alpha/2$.

$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}; \tag{11.34}$$

$$\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}}; \tag{11.35}$$

$$\tan\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}} = \frac{\sin\alpha}{1+\cos\alpha} = \frac{1-\cos\alpha}{\sin\alpha}, \quad \alpha \neq k\pi. \tag{11.36}$$

Il segno davanti ai radicali va scelto tenendo del quadrante in cui cade il punto P associato, sulla circonferenza goniometrica, al numero $\alpha/2$. Le prime formule per la tangente sono ottenute facendo il rapporto tra quelle del seno e quelle del coseno, le altre moltiplicando numeratore e denominatore per $1-\cos\alpha$ o per $1+\cos\alpha$ e osservando che si ha concordanza di segno tra $\tan\alpha/2$ e $\sin\alpha$, mentre $1-\cos\alpha$ e $1+\cos\alpha$ non sono mai negativi.

Applicando ripetutamente queste formule si possono ricavare anche le formule coinvolgenti $\alpha/2^n$. Si noti che, invece, non esistono formule generali, coinvolgenti solo espressioni razionali o radicali, per la trisezione, in accordo con il fatto che la trisezione di un un angolo generico non si può fare con riga e compasso.

Formule parametriche

Dall'ultima delle (11.29), tenendo conto dell'identità fondamentale espressa mediante $\alpha/2$ e infine dividendo numeratore e denominatore per $\cos^2\alpha/2$, si ottiene

$$\sin\alpha = 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} = \frac{2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}}{\cos^2\frac{\alpha}{2}+\sin^2\frac{\alpha}{2}} = \frac{2\tan\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}, \quad \alpha \neq \pi + 2k\pi. \tag{11.37}$$

Procedendo in maniera analoga con l'ultima delle (11.30) si ottiene invece

$$\cos\alpha = \frac{\cos^2\frac{\alpha}{2}-\sin^2\frac{\alpha}{2}}{\cos^2\frac{\alpha}{2}+\sin^2\frac{\alpha}{2}} = \frac{1-\tan^2\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}, \quad \alpha \neq \pi + 2k\pi. \tag{11.38}$$

Queste formule esprimono, in forma razionale, il seno e il coseno di un angolo in funzione della tangente dell'angolo metà e per questo motivo sono anche dette *formule razionali* per il seno e il coseno.

Posto

$$t = \tan\frac{\alpha}{2},$$

le formule razionali si scrivono nella forma

$$\sin\alpha = \frac{2t}{1+t^2}, \qquad \cos\alpha = \frac{1-t^2}{1+t^2}. \tag{11.39}$$

Nella forma espressa dalle (11.39) esse sono dette *formule parametriche* per il seno e il coseno, in quanto esprimono il seno e il coseno in funzione di un parametro.

Se ricordiamo che $\cos\alpha$ e $\sin\alpha$ sono l'ascissa e l'ordinata di un punto della circonferenza goniometrica, e teniamo conto della condizione $\alpha \neq \pi + 2k\pi$, concludiamo che le equazioni

$$\begin{cases} x = \dfrac{1-t^2}{1+t^2} \\ y = \dfrac{2t}{1+t^2} \end{cases} \tag{11.40}$$

sono le equazioni parametriche della circonferenza unitaria di centro l'origine, con l'esclusione del punto $(-1,0)$. Per un utile controllo costruiamo la tabella seguente.

t	-10	-5	-2	-1	0	1	2	5	10
x	-0.980	-0.923	-0.600	0	1	0	-0.600	-0.923	-0.980
y	-0.198	-0.385	-0.800	-1	0	1	0.800	0.385	0.198

La figura 11.20 mostra la circonferenza con evidenziati i punti della tabella.

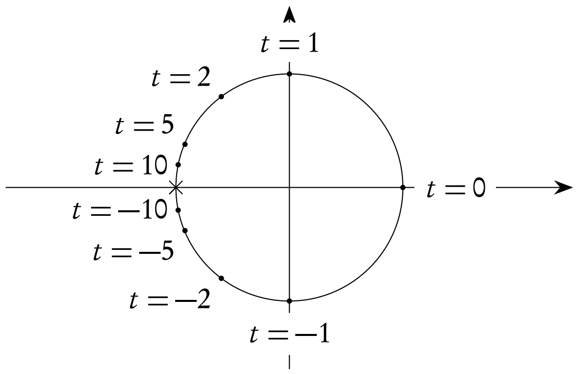


Figura 11.20.: *Circonferenza unitaria in una speciale rappresentazione parametrica*

Si osservi come, a partire dal punto $(1,0)$, corrispondente al valore $t=0$ del parametro, e facendo crescere il parametro, il punto corrispondente varia sulla circonferenza, muovendosi in senso antiorario verso il punto $(-1, 0)$, che non viene mai raggiunto; facendo invece decrescere il parametro, il punto corrispondente varia sulla circonferenza, muovendosi in senso orario verso il punto $(-1, 0)$, che non viene mai raggiunto. Se immaginiamo che t rappresenti il tempo, possiamo pensare che tale movimento diventi "sempre più lento" al crescere, o al decrescere, di t, a partire dal valore 0.

Formule di prostaferesi

Riconsideriamo le formule di addizione e sottrazione per il seno e il coseno, scrivendole separatamente.

$$\begin{cases} \sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \\ \sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta \end{cases}, \qquad \begin{cases} \cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta \end{cases}.$$

Se sommiamo e poi sottraiamo le prime due e le seconde due otteniamo

$$\begin{cases} \sin(\alpha+\beta) + \sin(\alpha-\beta) = 2\sin\alpha\cos\beta \\ \sin(\alpha+\beta) - \sin(\alpha-\beta) = 2\cos\alpha\sin\beta \end{cases}, \qquad \begin{cases} \cos(\alpha+\beta) + \cos(\alpha-\beta) = 2\cos\alpha\cos\beta \\ \cos(\alpha+\beta) - \cos(\alpha-\beta) = -2\sin\alpha\sin\beta \end{cases}. \tag{11.41}$$

Posto $\alpha+\beta = p$ e $\alpha-\beta = q$ si trova

$$\alpha = \frac{p+q}{2}, \quad \beta = \frac{p+q}{2}.$$

Tenendo conto di queste uguaglianze, le (11.41) si scrivono nella forma

$$\begin{cases} \sin p + \sin q = 2\sin\dfrac{p+q}{2}\cos\dfrac{p-q}{2} \\[2ex] \sin p - \sin q = 2\cos\dfrac{p+q}{2}\sin\dfrac{p-q}{2} \end{cases}, \qquad \begin{cases} \cos p + \cos q = 2\cos\dfrac{p+q}{2}\cos\dfrac{p-q}{2} \\[2ex] \cos p - \cos q = -2\sin\dfrac{p+q}{2}\sin\dfrac{p-q}{2} \end{cases}. \tag{11.42}$$

Le (11.42) trasformano una somma o differenza di seni o coseni in un prodotto e sono particolarmente utili, per esempio, nella risoluzione di certe disequazioni. Si chiamano formule di prostaferesi, dal greco *prosthesis* e *aphairesis*, ovvero *somma* e *sottrazione*.

Esistono alcune filastrocche utili per memorizzare queste formule: riteniamo molto più efficiente memorizzare il procedimento riportato per ottenerle, molto semplice e rapido.

Formule di Werner

Riprendiamo in esame le (11.41): lette da destra a sinistra esse permettono di trasformare un prodotto tra due delle funzioni seno e coseno in una somma. Poiché le prime due esprimono sempre il prodotto tra il seno di un angolo e il coseno di un altro, rimangono le tre formule seguenti:

$$\begin{cases} \sin\alpha\cos\beta = \dfrac{1}{2}\left(\sin(\alpha+\beta) + \sin(\alpha-\beta)\right) \\[2ex] \cos\alpha\cos\beta = \dfrac{1}{2}\left(\cos(\alpha+\beta) + \cos(\alpha-\beta)\right) \\[2ex] \sin\alpha\sin\beta = -\dfrac{1}{2}\left(\cos(\alpha+\beta) - \cos(\alpha-\beta)\right) \end{cases} \tag{11.43}$$

Queste formule consentono di trasformare un prodotto di funzioni trigonometriche in una somma e sono particolarmente utili, per esempio, nel calcolo di certi integrali. È appena il caso il caso di accennare al fatto che non vale la pena memorizzare queste formule: il procedimento per ricavarle è praticamente immediato.

11.8. Angoli notevoli

Abbiamo già considerato, vedi la tabella 11.2 nella pagina 366, i valori delle funzioni seno e coseno per alcuni angoli, o archi, particolari. Vogliamo ora completare questa tabella con altri valori notevoli, ricordando che è sufficiente limitarsi agli archi compresi tra 0 e $\pi/2$, per poi eventualmente procedere per simmetrie e periodicità.

Consideriamo innanzitutto il valore $\pi/10$. Osserviamo che l'angolo al centro che sottende il lato del decagono regolare inscritto in una circonferenza misura $2\pi/10$, ovvero $\pi/5$. Inoltre questo lato, essendo la sezione aurea del raggio, misura

$$l_{10} = \frac{\sqrt{5}-1}{2}\,r. \tag{11.44}$$

Considerato allora sulla circonferenza trigonometrica il punto P corrispondente al valore $\alpha = \pi/10$ e il suo simmetrico Q rispetto all'asse x, il segmento $\overline{\mathsf{PQ}}$ è il lato del decagono regolare inscritto. Il valore di $\sin\alpha$ è la metà di questo lato: si veda la figura 11.21. Il coseno si può poi trovare mediante l'identità fondamentale. Si ottiene:

$$\sin\left(\frac{\pi}{10}\right) = \frac{\sqrt{5}-1}{4}, \quad \cos\left(\frac{\pi}{10}\right) = \frac{\sqrt{10+2\sqrt{5}}}{4}. \tag{11.45}$$

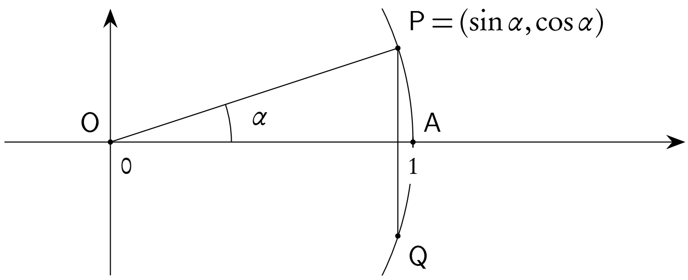


Figura 11.21.: *Determinazione di* $\sin\pi/10$

A partire dalla tabella 11.2 e dai valori appena ottenuti per il valore $\pi/10$, usando le formule di duplicazione e bisezione si possono calcolare le funzioni trigonometriche di altri valori notevoli. Riportiamo quelli più importanti nella tabella 11.3, dove abbiamo anche riportato, per un utile confronto, i valori in gradi corrispondenti alle misure in radianti.

11.9. L'inversione delle funzioni trigonometriche

Nessuna delle funzioni trigonometriche, come funzioni di $\mathbb{R}$ in $\mathbb{R}$, è iniettiva; le funzioni seno, coseno secante e cosecante non sono nemmeno suriettive. Tuttavia è possibile considerare opportune restrizioni sul dominio e sul codominio, in modo da ottenere funzioni biunivoche e quindi invertibili. Non ci occuperemo delle inverse delle funzioni secante e cosecante, che sono di limitato uso.

La scelta di quale restrizione considerare, in particolare sul dominio, è largamente arbitraria: quella che noi faremo è quella stabilita dalla consuetudine e ha una serie di vantaggi, tra cui quello di operare

Tabella 11.3.: *Valori notevoli delle funzioni trigonometriche*

gradi	radianti	seno	coseno	tangente	cotangente
$0°$	0	0	1	0	non definita
$15°$	$\frac{\pi}{12}$	$\frac{\sqrt{6}-\sqrt{2}}{4}$	$\frac{\sqrt{6}+\sqrt{2}}{4}$	$2-\sqrt{3}$	$2+\sqrt{3}$
$18°$	$\frac{\pi}{10}$	$\frac{\sqrt{5}-1}{4}$	$\frac{\sqrt{10+2\sqrt{5}}}{4}$	$\sqrt{\frac{5-2\sqrt{5}}{5}}$	$\sqrt{5+2\sqrt{5}}$
$22° 30'$	$\frac{\pi}{8}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$	$\frac{\sqrt{2+\sqrt{2}}}{2}$	$\sqrt{2}-1$	$\sqrt{2}+1$
$30°$	$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{3}$	$\sqrt{3}$
$36°$	$\frac{\pi}{5}$	$\frac{\sqrt{10-2\sqrt{5}}}{4}$	$\frac{\sqrt{5}+1}{4}$	$\sqrt{5-2\sqrt{5}}$	$\sqrt{\frac{5+2\sqrt{5}}{5}}$
$45°$	$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1	1
$54°$	$\frac{3}{10}\pi$	$\frac{\sqrt{5}+1}{4}$	$\frac{\sqrt{10-2\sqrt{5}}}{4}$	$\sqrt{\frac{5+2\sqrt{5}}{5}}$	$\sqrt{5-2\sqrt{5}}$
$60°$	$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$	$\frac{\sqrt{3}}{3}$
$67° 30'$	$\frac{3}{8}\pi$	$\frac{\sqrt{2+\sqrt{2}}}{2}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$	$\sqrt{2}+1$	$\sqrt{2}-1$
$72°$	$\frac{2}{5}\pi$	$\frac{\sqrt{10+2\sqrt{5}}}{4}$	$\frac{\sqrt{5}-1}{4}$	$\sqrt{5+2\sqrt{5}}$	$\sqrt{\frac{5-2\sqrt{5}}{5}}$
$75°$	$\frac{5}{12}\pi$	$\frac{\sqrt{6}+\sqrt{2}}{4}$	$\frac{\sqrt{6}-\sqrt{2}}{4}$	$2+\sqrt{3}$	$2-\sqrt{3}$
$90°$	$\frac{\pi}{2}$	1	0	non definita	0
$180°$	π	0	-1	0	non definita
$270°$	$\frac{3}{2}\pi$	-1	0	non definita	0
$360°$	2π	0	1	0	non definita

con archi non troppo grandi. Purtroppo non è possibile fare la stessa scelta per tutte le quattro funzioni cui siamo interessati. Le quattro situazioni sono le seguenti:

$$\text{sin}\colon \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]; \quad \text{tan}\colon \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[; \quad \text{cos}: [0, \pi]; \quad \text{cot}\colon]0, \pi[. \tag{11.46}$$

11.9.1. la funzione arcseno

Per ottenere una funzione invertibile a partire dalla funzione seno si considera la restrizione all'intervallo $[-\pi/2, \pi/2]$ sul dominio e $[-1, 1]$ sul codominio:

$$\sin_{|[-\pi/2,\pi/2]}\colon \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]. \tag{11.47}$$

Definizione 11.11. *L'inversa della restrizione* (11.47) *della funzione seno si chiama funzione* arcseno *e si denota con* arcsin *o con* asin *o con* invsin *o ancora con* $\sin^{-1}$*:*

$$\arcsin\colon [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \tag{11.48}$$

L'ultima delle notazioni indicate è comune soprattutto sulle calcolatrici tascabili; come è noto si tratta della notazione standard per l'inversa di una funzione: in questo caso, oltre al rischio di confusione con il simbolo del reciproco, la notazione ci pare impropria, in quanto la funzione in esame *non* è l'inversa della funzione seno, ma solo di una sua restrizione. Discorsi simili si applicano alle altre funzioni trigonometriche inverse che considereremo.

La costruzione del grafico della funzione arcseno è immediata: come già noto basta considerare il simmetrico, rispetto alla bisettrice del primo e terzo quadrante, del grafico della restrizione della funzione seno.

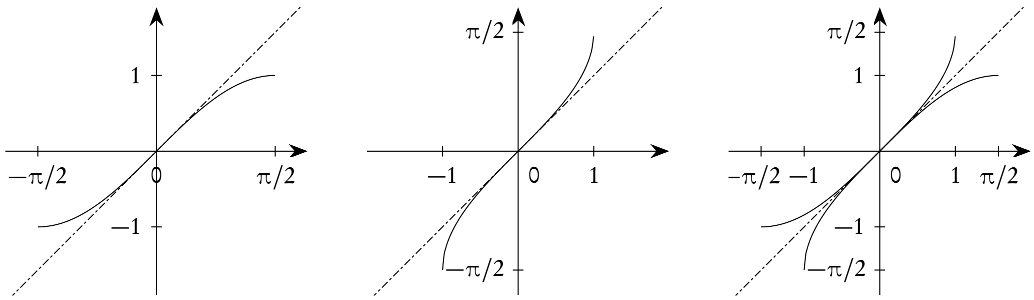


Figura 11.22.: *Restrizione della funzione seno, funzione arcseno e confronto tra le due funzioni*

La funzione arcseno, a partire dal numero reale x dell'intervallo $[-1, 1]$, fornisce l'unico numero reale, cioè l'unico arco, dell'intervallo $[-\pi/2, \pi/2]$ il cui seno è x. Si noti che è fondamentale precisare che l'arco fornito da $\arcsin x$ sta in $[-\pi/2, \pi/2]$: di archi che abbiano un seno x compreso tra -1 e 1 ce ne sono infiniti.

È utile osservare che, avendo la funzione seno come tangente nell'origine proprio la bisettrice del primo e terzo quadrante[3], ciò succede anche per la funzione arcseno: le due funzioni sono quindi tangenti

[3]Archi misurati in radianti!

nell'origine e anzi questo è il loro unico punto comune. È come dire che l'equazione $\sin x = \arcsin x$ ha come unica soluzione $x = 0$.

Per rendersi ancora meglio conto del significato della funzione arcseno, si può utilizzare la circonferenza goniometrica: la funzione seno a partire dall'arco x (che può essere un qualunque numero reale) fornisce l'ordinata del punto P corrispondente a x sulla circonferenza goniometrica; la funzione arcseno, a partire da un numero x compreso tra -1 e 1, considerato come ordinata di un punto sull'asse delle y, fornisce quell'unico punto P della circonferenza goniometrica, situato nel primo o quarto quadrante, che corrisponde all'arco, appartenente a $[-\pi/2, \pi/2]$, il cui seno vale x.

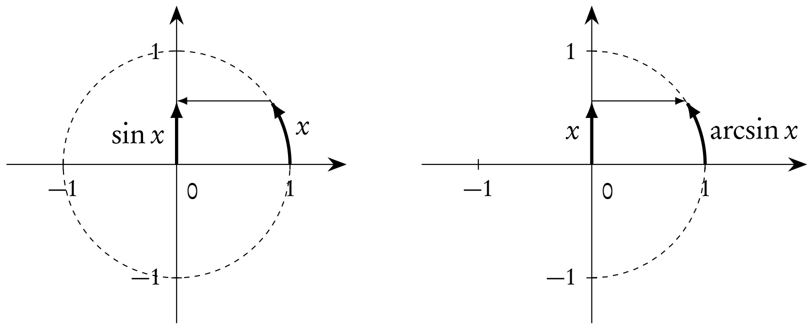


Figura 11.23.: *Le funzioni seno e arcseno viste sulla circonferenza goniometrica*

Se le funzioni seno e arcseno fossero una l'inversa dell'altra non ci sarebbe molto da dire relativamente alla funzione composta $\sin(\arcsin x)$: basterebbe solo osservare che, in accordo con la regola generale sulle funzioni inverse si otterrebbe l'identità sul dominio di arcseno, cioè sull'intervallo $[-1, 1]$. Poiché però l'arcseno è l'inversa di una restrizione della funzione seno, una precisazione si rende opportuna. Dato un reale x dell'intervallo $[-1, 1]$, la funzione arcseno produce un reale dell'intervallo $[-\pi/2, \pi/2]$, e questo è esattamente l'intervallo a cui abbiamo ristretto la funzione seno per poterla invertire. Se ne deduce che la funzione in oggetto è proprio l'identità sull'intervallo $[-1, 1]$:

$$\sin(\arcsin x) = x, \ \forall x \in [-1, 1]. \tag{11.49}$$

Ben diverso è il caso della funzione $\arcsin(\sin x)$, cioè della composizione in ordine inverso. Intanto questa ha come dominio tutto $\mathbb{R}$, ma non può essere l'identità su $\mathbb{R}$, perché arcseno ha come immagine l'intervallo $[-\pi/2, \pi/2]$. Dunque il fatto che le funzioni seno e arcseno non siano una l'inversa dell'altra ha, in questa composizione, importanti conseguenze. Per capire come vanno le cose cominciamo con l'osservare che la funzione è periodica di periodo 2π, e quindi basterà limitare l'indagine a un intervallo ampio 2π: sceglieremo, per motivi di convenienza, l'intervallo $[-\pi/2, 3\pi/2]$. Per quanto riguarda la prima metà di questo intervallo, e cioè $[-\pi/2, \pi/2]$, non ci sono problemi: per le proprietà delle funzioni inverse si avrà semplicemente l'identità di $[-\pi/2, \pi/2]$. Per rendersene ben conto si può esaminare la parte sinistra della figura 11.24: partendo da x si ottiene $\sin x$; successivamente partendo da $\sin x$ si ottiene $\arcsin(\sin x)$ che coincide con il valore di x da cui si era partiti. Se invece $x \in [\pi/2, 3\pi/2]$, $\sin x$ è l'ordinata del corrispondente punto sulla circonferenza goniometrica, mentre $\arcsin(\sin x)$ corrisponde al punto della circonferenza goniometrica avente quell'ordinata, ma appartenente all'intervallo $[-\pi/2, \pi/2]$: $\arcsin(\sin x)$ sarà dunque, per motivi di simmetria, $\pi - x$. Si veda la parte destra della figura 11.24.

Il grafico di questa funzione è rappresentato nella figura 11.25.

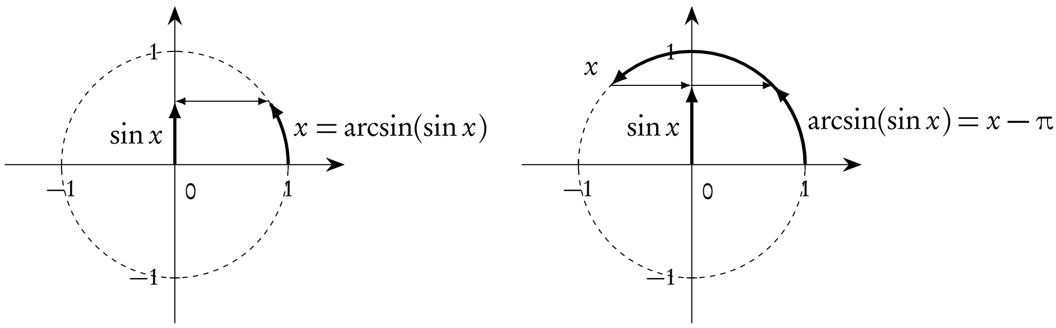


Figura 11.24.: *Calcolo di* $\arcsin(\sin x)$, $x \in [-\pi/2, \pi/2]$ *a sinistra, e* $x \in [\pi/2, 3\pi/2]$ *a destra*

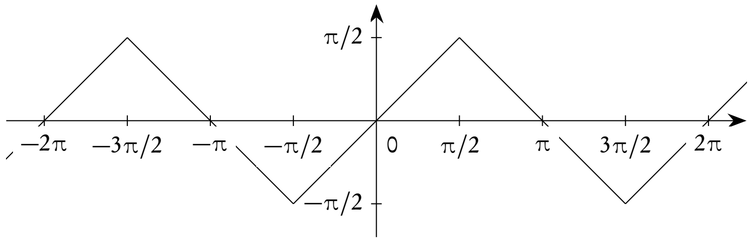


Figura 11.25.: *La funzione* $\arcsin(\sin x)$

Osservazione 11.12. I problemi derivanti dal fatto che la funzione arcseno non è l'inversa della funzione seno, ma solo di una sua restrizione, sono simili a quelli trattati nell'esempio relativo alla funzioni radice e elevamento al quadrato, discusso nella pagina 117.

11.9.2. La funzione arccoseno

Per ottenere una funzione invertibile a partire dalla funzione coseno si considera la restrizione all'intervallo $[0, \pi]$ sul dominio e $[-1, 1]$ sul codominio:

$$\cos_{|[0,\pi]} \colon [0, \pi] \to [-1, 1]. \tag{11.50}$$

Definizione 11.13. *L'inversa della restrizione* (11.50) *della funzione coseno si chiama funzione* arccoseno *e si denota con* arccos *o con* acos *o con* invcos *o ancora con* $\cos^{-1}$*:*

$$\arccos \colon [-1, 1] \to [0, \pi]. \tag{11.51}$$

La costruzione del grafico della funzione arccoseno è immediata: come già noto basta considerare il simmetrico, rispetto alla bisettrice del primo e terzo quadrante, del grafico della restrizione della funzione seno.

La funzione arccoseno, a partire dal numero reale x dell'intervallo $[-1, 1]$, fornisce l'unico numero reale, cioè l'unico arco, dell'intervallo $[0, \pi]$ il cui coseno è x. Si noti che è fondamentale precisare che l'arco fornito da $\arccos x$ sta in $[0, \pi]$: di archi che abbiano un coseno x compreso tra -1 e 1 ce ne sono infiniti.

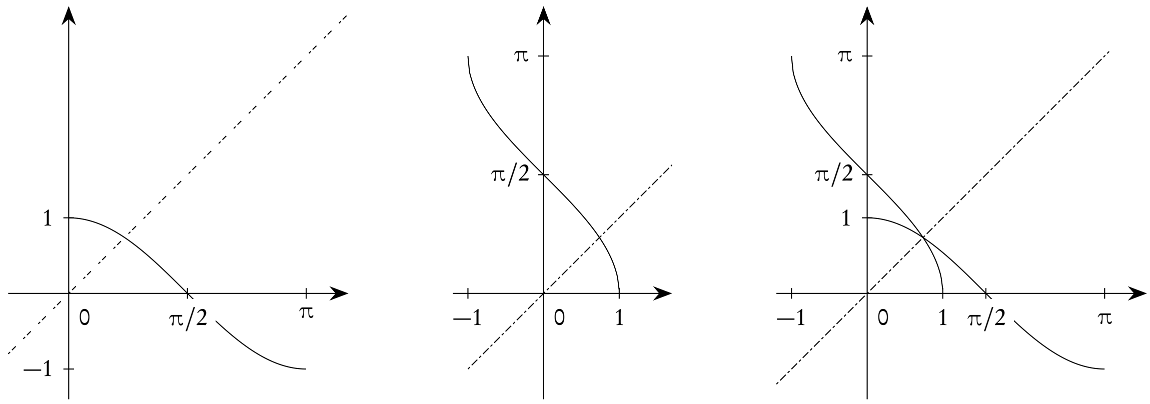


Figura 11.26.: *Restrizione della funzione coseno, funzione arccoseno e confronto tra le due funzioni*

È utile osservare che, poiché la funzione coseno interseca la bisettrice $y = x$ in un solo punto, ciò succede anche per la funzione arccoseno: le due funzioni hanno quindi questo punto come unico punto comune. È come dire che l'equazione $\cos x = \arccos x$ ha un'unica soluzione, il cui valore approssimato è 0.739085.

Per rendersi ancora meglio conto del significato della funzione arccoseno, si può utilizzare la circonferenza goniometrica: la funzione coseno a partire dall'arco x (che può essere un qualunque numero reale) fornisce l'ascissa del punto P corrispondente a x sulla circonferenza goniometrica; la funzione arccoseno, a partire da un numero x compreso tra -1 e 1, considerato come ascissa di un punto sull'asse delle x, fornisce quell'unico punto P della circonferenza goniometrica, situato nel primo o secondo quadrante, che corrisponde all'arco, appartenente a $[0, \pi]$, il cui coseno vale x.

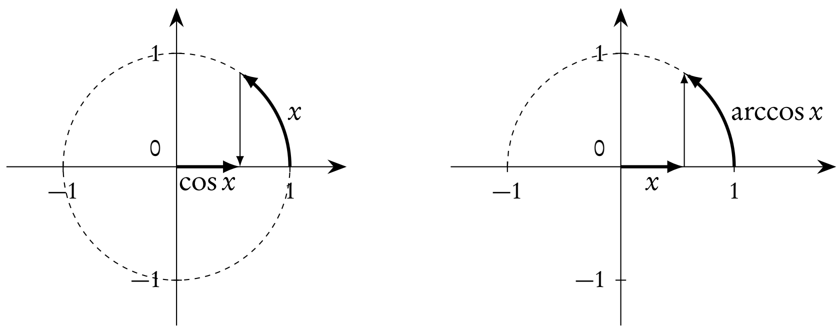


Figura 11.27.: *Le funzioni coseno e arccoseno viste sulla circonferenza goniometrica*

Se le funzioni coseno e arccoseno fossero una l'inversa dell'altra non ci sarebbe molto da dire relativamente alla funzione composta $\cos(\arccos x)$: basterebbe solo osservare che, in accordo con la regola generale sulle funzioni inverse si otterrebbe l'identità sul dominio di arccoseno, cioè sull'intervallo $[-1, 1]$. Poiché però l'arccoseno è l'inversa di una restrizione della funzione coseno, una precisazione si rende opportuna. Dato un reale x dell'intervallo $[-1, 1]$, la funzione arccoseno produce un reale dell'intervallo $[0, \pi]$, e questo è esattamente l'intervallo a cui abbiamo ristretto la funzione coseno per poterla invertire. Se ne deduce che la funzione in oggetto è proprio l'identità sull'intervallo $[-1, 1]$:

(11.52) $$\cos(\arccos x) = x,\ \forall x \in [-1, 1].$$

Ben diverso è il caso della funzione $\arccos(\cos x)$, cioè della composizione in ordine inverso. Intanto questa ha come dominio tutto $\mathbb{R}$, ma non può essere l'identità su $\mathbb{R}$, perché arccoseno ha come immagine l'intervallo $[0, \pi]$. Dunque il fatto che le funzioni coseno e arccoseno non siano una l'inversa dell'altra ha, in questa composizione, importanti conseguenze. Per capire come vanno le cose cominciamo con l'osservare che la funzione è periodica di periodo 2π, e quindi basterà limitare l'indagine a un intervallo ampio 2π: sceglieremo, per motivi di convenienza, l'intervallo $[0, 2\pi]$. Per quanto riguarda la prima metà di questo intervallo, e cioè $[0, \pi]$, non ci sono problemi: per le proprietà delle funzioni inverse si avrà semplicemente l'identità di $[0, \pi]$. Per rendersene ben conto si può esaminare la parte sinistra della figura 11.28: partendo da x si ottiene $\cos x$; successivamente partendo da $\cos x$ si ottiene $\arccos(\cos x)$ che coincide con il valore di x da cui si era partiti. Se invece $x \in [\pi, 2\pi]$, $\cos x$ è l'ascissa del corrispondente punto sulla circonferenza goniometrica, mentre $\arccos(\cos x)$ corrisponde al punto della circonferenza goniometrica avente quell'ascissa, ma appartenente all'intervallo $[0, \pi]$: $\arccos(\cos x)$ sarà dunque, per motivi di simmetria, $2\pi - x$. Si veda la parte destra della figura 11.28.

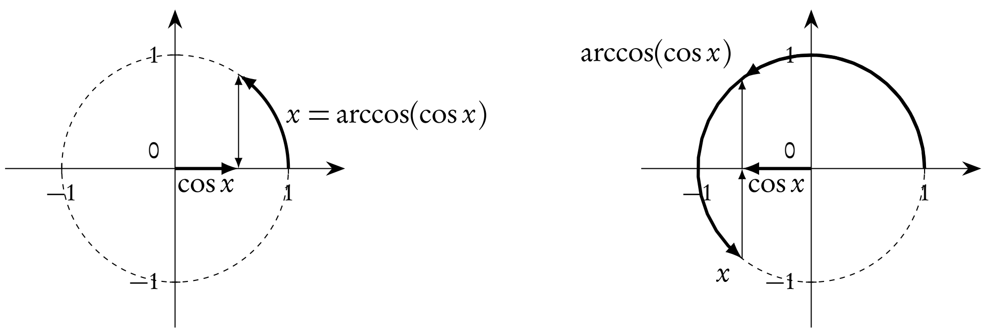


Figura 11.28.: *Calcolo di* $\arccos(\cos x)$, $x \in [0, \pi]$ *a sinistra, e* $x \in [\pi, 2\pi]$ *a destra*

Il grafico di questa funzione è rappresentato nella figura 11.29.

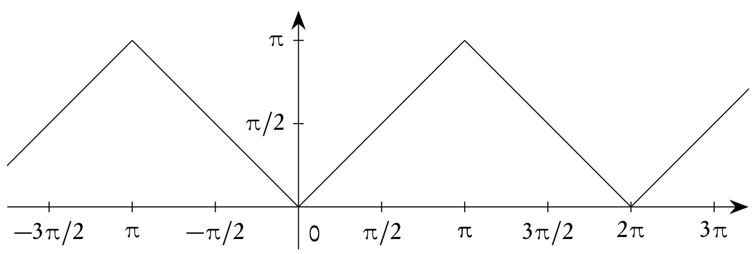


Figura 11.29.: *La funzione* $\arccos(\cos x)$

11.9.3. La funzione arctangente

Per ottenere una funzione invertibile a partire dalla funzione tangente si considera la restrizione all'intervallo $]-\pi/2, \pi/2[$ sul dominio, mentre non è necessaria alcuna restrizione sul codominio, in quanto l'immagine di questa restrizione della tangente è sempre l'insieme $\mathbb{R}$:

$$\tan_{|\,]-\pi/2,\pi/2[}\colon \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \mathbb{R}. \tag{11.53}$$

Definizione 11.14. *L'inversa della restrizione* (11.53) *della funzione tangente si chiama funzione* arctangente *e si denota con* arctan *o con* atan *o con* invtan *o ancora con* $\tan^{-1}$*:*

$$\arctan\colon \mathbb{R} \to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[. \tag{11.54}$$

La costruzione del grafico della funzione arctangente è immediata: come già noto basta considerare il simmetrico, rispetto alla bisettrice del primo e terzo quadrante, del grafico della restrizione della funzione tangente.

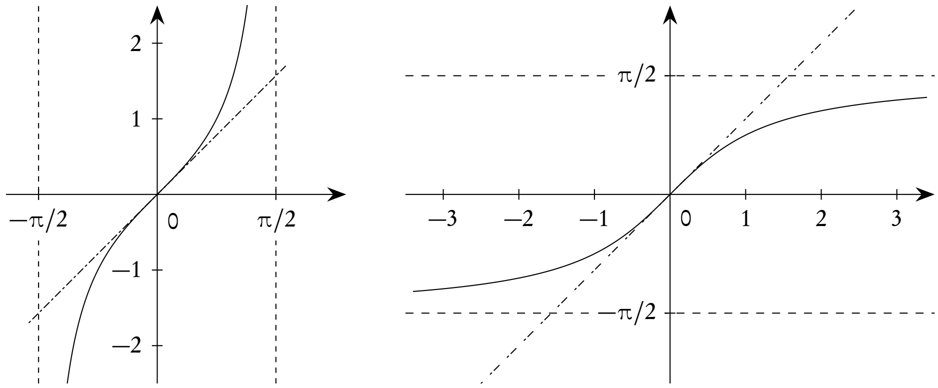


Figura 11.30.: *Restrizione della funzione tangente e funzione arctangente*

La funzione arctangente, a partire dal numero reale x, fornisce l'unico numero reale, cioè l'unico arco, dell'intervallo $]-\pi/2, \pi/2[$ la cui tangente è x. Si noti che è fondamentale precisare che l'arco fornito da $\arctan x$ sta in $]-\pi/2, \pi/2[$: di archi che abbiano per tangente x ce ne sono infiniti.

È utile osservare che, avendo la funzione tangente come retta tangente nell'origine proprio la bisettrice del primo e terzo quadrante, ciò succede anche per la funzione arctangente: le due funzioni sono quindi tangenti nell'origine e anzi questo è il loro unico punto comune, in $]-\pi/2, \pi/2[$. È come dire che l'equazione $\tan x = \arctan x$ ha, sempre in $]-\pi/2, \pi/2[$, come unica soluzione $x = 0$. Altri punti comuni si trovano sia a sinistra che a destra di questo intervallo: basta ricordare che la funzione tangente si ripete per periodicità, assumendo sempre tutti in valori reali in ogni intervallo del tipo $]-\pi/2 + k\pi, \pi/2 + k\pi[$. Si veda la figura 11.31.

Inoltre la restrizione della funzione tangente ha le rette $x = \pm\pi/2$ come asintoti verticali, quindi la funzione arctangente ha due asintoti orizzontali $y = \pm\pi/2$.

Per rendersi ancora meglio conto del significato della funzione arctangente, in analogia a quanto fatto per il caso del seno e del coseno, si può utilizzare la circonferenza goniometrica: la funzione tangente a partire dall'arco x (che può essere un qualunque numero reale) fornisce l'ordinata del punto T corrispondente a x sulla retta $x = 1$; la funzione arctangente, a partire da un numero x qualunque, considerato come ordinata di un punto sulla retta $x = 1$, fornisce quell'unico punto P della circonferenza

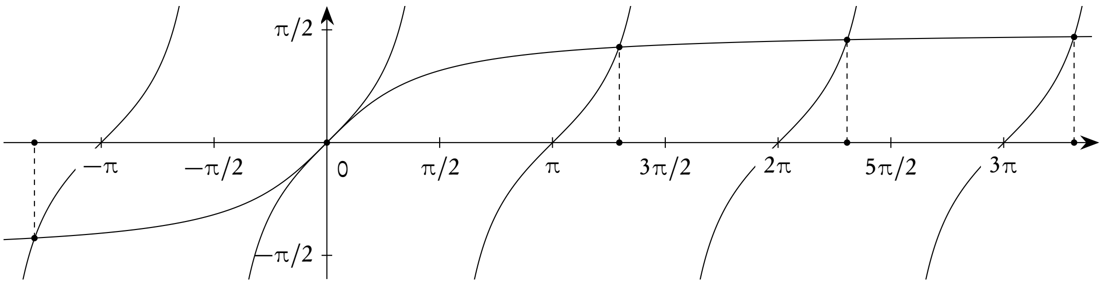


Figura 11.31.: *Intersezioni tra la funzione tangente e la funzione arctangente*

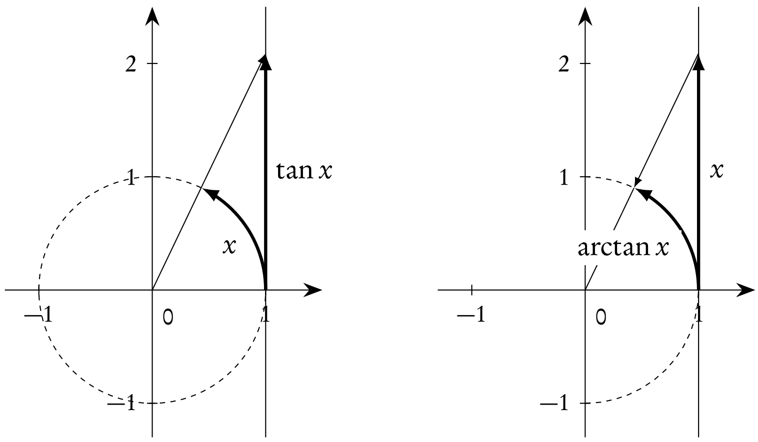


Figura 11.32.: *Le funzioni tangente e arctangente viste sulla circonferenza goniometrica*

goniometrica, situato nel primo o quarto quadrante, che corrisponde all'arco, appartenente a $]-\pi/2, \pi/2[$, la cui tangente vale x.

Per quanto riguarda la funzione $\tan(\arctan x)$, si conclude facilmente che essa è proprio l'identità su tutto $\mathbb{R}$:

$$\tan(\arctan x) = x, \ \forall x \in \mathbb{R}. \tag{11.55}$$

Anche per la funzione $\arctan(\tan x)$ le cose vanno in maniera più semplice che non per le corrispondenti funzioni coinvolgenti il seno e il coseno. Si tratta infatti di una funzione periodica di periodo π e, nel tratto $]-\pi/2, \pi/2[$, essa è la funzione identica, in quanto l'arctangente è proprio l'inversa della tangente ristretta a questo intervallo. Come la tangente, questa funzione non sarà definita nei punti $\pi/2 + k\pi$, $k \in \mathbb{Z}$. Il suo grafico è riportato nella figura 11.33.

La funzione arctangente è molto importante nelle applicazioni: come si vedrà nei successivi corsi di analisi essa gioca un ruolo cruciale nel calcolo degli integrali indefiniti.

11.9.4. La funzione arccotangente

Mentre per le restrizioni delle funzioni seno, coseno e tangente le convenzioni che abbiamo indicato sono universalmente accettate, non altrettanto succede per la funzione cotangente. In effetti, rispettando

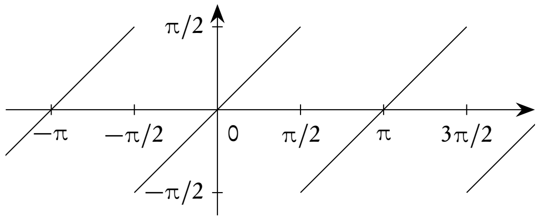


Figura 11.33.: *La funzione* arctan(tan x)

l'obbligo di avere una funzione iniettiva e l'opportunità di non lavorare con archi troppo grandi, ci sono due possibilità, ognuna delle quali presenta qualche vantaggio. Si può scegliere l'intervallo $]-\pi/2, \pi/2]$, privato dell'origine dove la cotangente non è definita, oppure l'intervallo $]0, \pi[$. Nel primo caso si ottiene una funzione dispari (ovvero simmetrica rispetto all'origine), con l'eccezione del punto 0, dove la funzione vale $\pi/2$, ma con lo svantaggio di avere una discontinuità(4) (un salto) proprio in corrispondenza dell'origine. C'è anche un ulteriore vantaggio con la prima scelta, vantaggio cui faremo un cenno successivamente, a proposito della funzione $\arctan(1/x)$. Nel secondo caso si ottiene una funzione continua, ma senza particolari simmetrie. La seconda scelta presenta anche l'ulteriore vantaggio di assumere, per la funzione cotangente, una convenzione quasi identica(5) a quella scelta per la funzione coseno. La quasi totalità dei testi adotta la seconda convenzione(6), i software di calcolo simbolico si dividono, ma con preferenza per la prima: Mathcad e Maple scelgono la seconda, Mathematica e Matlab la prima. Noi ci atterremo alla seconda convenzione, presente in tutti i testi che abbiamo consultato.

Per ottenere una funzione invertibile a partire dalla funzione cotangente considereremo dunque la restrizione all'intervallo $]0, \pi[$ sul dominio, mentre non è necessaria alcuna restrizione sul codominio, in quanto l'immagine di questa restrizione della cotangente è sempre l'insieme $\mathbb{R}$:

$$\cot_{|\,]0,\pi[}\colon\ ]0, \pi[\ \to \mathbb{R}. \tag{11.56}$$

Definizione 11.15. *L'inversa della restrizione* (11.56) *della funzione cotangente si chiama funzione* arccotangente *e si denota con* arccot *o con* acot *o con* invcot *o ancora con* $\cot^{-1}$*:*

$$\operatorname{arccot}\colon \mathbb{R} \to\ ]0, \pi[\,. \tag{11.57}$$

La costruzione del grafico della funzione arccotangente è immediata: come già noto basta considerare il simmetrico, rispetto alla bisettrice del primo e terzo quadrante, del grafico della restrizione della funzione cotangente.

La funzione arccotangente, a partire dal numero reale x, fornisce l'unico numero reale, cioè l'unico arco, dell'intervallo $]0, \pi[$ la cui cotangente è x. Si noti che è fondamentale precisare che l'arco fornito da $\operatorname{arccot} x$ sta in $]0, \pi[$: di archi che abbiano una cotangente x qualunque ce ne sono infiniti.

È utile osservare che, poiché la funzione cotangente interseca la bisettrice $y = x$ in un solo punto, ciò succede anche per la funzione arccotangente: le due funzioni hanno quindi questo punto come

4 In questo contesto ci si può limitare a considerare continua una funzione il cui grafico non presenti "strappi": nei successivi corsi di analisi il concetto sarà meglio precisato.

5 L'unica differenza sta nel fatto che per il coseno si sceglie $[0, \pi]$, per la cotangente $]0, \pi[$, cioè lo stesso intervallo privato degli estremi.

6 Se possibile i matematici vogliono evitare funzioni discontinue!

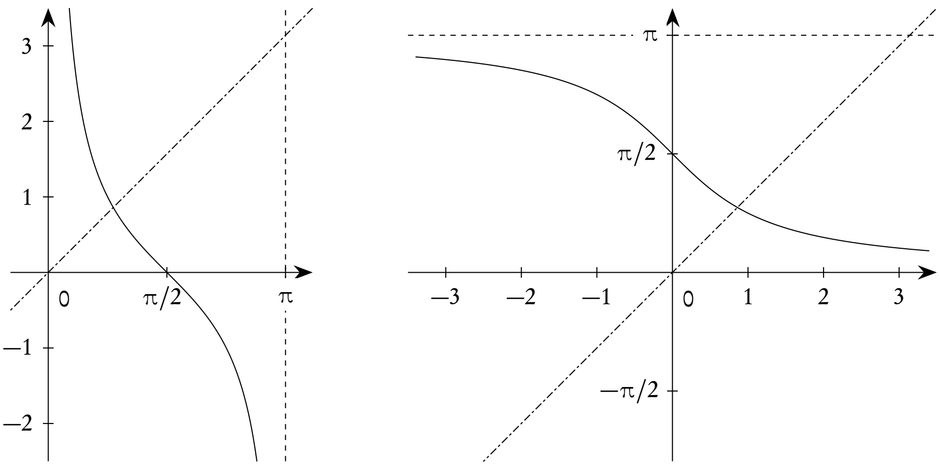


Figura 11.34.: *Restrizione della funzione cotangente e funzione arccotangente*

unico punto comune, nell'intervallo $]0, \pi[$. È come dire che l'equazione $\cot x = \operatorname{arccot} x$ ha, sempre nell'intervallo $]0, \pi[$ un'unica soluzione, il cui valore approssimato è 0.860334. Altre intersezioni si trovano sia a sinistra che a destra di questo intervallo: basta ricordare che la funzione cotangente si ripete per periodicità, assumendo sempre tutti i valori reali in ogni intervallo del tipo $]k\pi, \pi + h\pi[$. Si veda la figura 11.35.

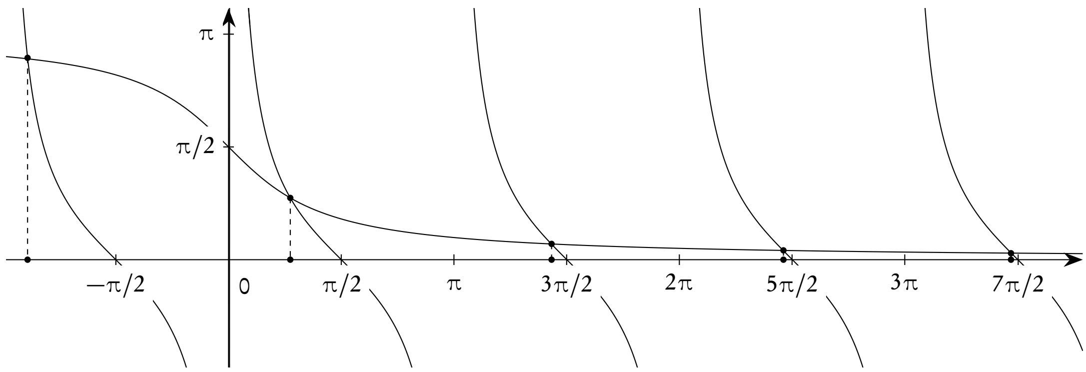


Figura 11.35.: *Intersezioni tra la funzione cotangente e la funzione arccotangente*

Per rendersi ancora meglio conto del significato della funzione arccotangente, in analogia a quanto fatto per il caso del seno, del coseno e della tangente, si può utilizzare la circonferenza goniometrica: la funzione cotangente a partire dall'arco x (che può essere un qualunque numero reale) fornisce l'ascissa del punto C corrispondente a x sulla retta $y = 1$; la funzione arccotangente, a partire da un numero x qualunque, considerato come ascissa di un punto sulla retta $y = 1$, fornisce quell'unico punto P della circonferenza goniometrica, situato nel primo o secondo quadrante, che corrisponde all'arco, appartenente a $][0, \pi[$, la cui cotangente vale x.

Per quanto riguarda la funzione $\cot(\operatorname{arccot} x)$, si conclude facilmente che essa è proprio l'identità su

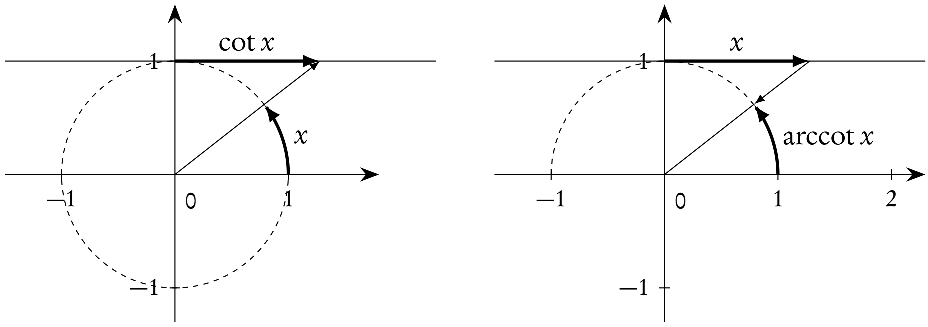


Figura 11.36.: *Le funzioni cotangente e arccotangente viste sulla circonferenza goniometrica*

tutto $\mathbb{R}$:

$$\cot(\operatorname{arccot} x) = x, \ \forall x \in \mathbb{R}. \tag{11.58}$$

Anche per la funzione $\operatorname{arccot}(\cot x)$ le cose vanno in maniera più semplice che non per le corrispondenti funzioni coinvolgenti il seno e il coseno. Si tratta infatti di una funzione periodica di periodo π e, nel tratto $]0, \pi[$, essa è la funzione identica, in quanto l'arccotangente è proprio l'inversa della cotangente ristretta a questo intervallo. Come la cotangente, questa funzione non sarà definita nei punti $k\pi$, $k \in \mathbb{Z}$. Il suo grafico è riportato nella figura 11.37.

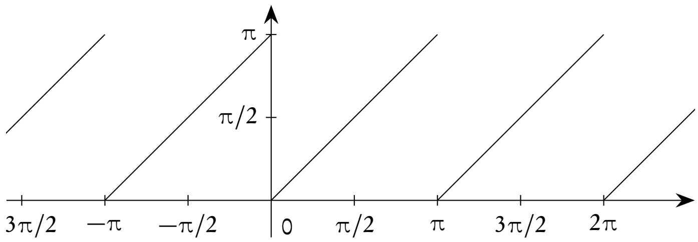


Figura 11.37.: *La funzione* $\operatorname{arccot}(\cot x)$

Per ragioni di completezza proponiamo anche, nella figura 11.38, i grafici relativi alla prima delle scelte indicate per invertire la funzione cotangente, scelta effettuata, come già osservato, dai due più diffusi e importanti software di calcolo simbolico, Mathematica e Matlab: prestare sempre la massima attenzione nell'uso di qualunque software. Se si vuole utilizzare in questi software la funzione arccotangente seguendo la definizione classica, sarà sufficiente utilizzare la successiva formula (11.60) per ridefinirla mediante la funzione arctangente.

11.9.5. Qualche relazione importante

Abbiamo già trattato le relazioni che intercorrono tra le funzioni trigonometriche e le inverse delle loro restrizioni. Ci occupiamo ora di qualche altra relazione importante, che dimostreremo come utile esercizio.

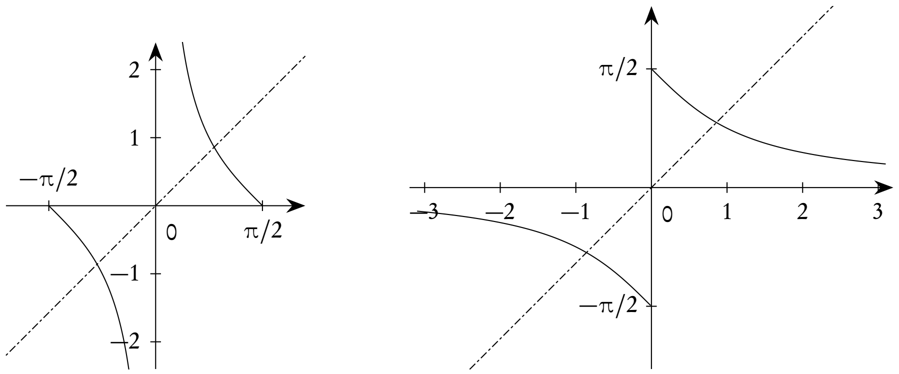


Figura 11.38.: *Definizione alternativa della funzione arccotangente (Mathematica, Maple)*

$$\arccos x + \arcsin x = \frac{\pi}{2}, \quad x \in [-1, 1]. \tag{11.59}$$

Con riferimento alla figura 11.39 distinguiamo i due casi $x > 0$ (figura di sinistra) e $x < 0$ (figura di destra). Per provare la (11.59) è sufficiente osservare che, in entrambi i casi, i triangoli OBE e OFD sono uguali, e tenere conto che, nel secondo caso, $\arcsin x$ è negativo.

La (11.59) permette di esprimere una delle due funzioni per mezzo dell'altra: in alcuni software di calcolo in effetti è definita solo la funzione arcsin, mentre arccos si ottiene proprio per mezzo della relazione appena provata.

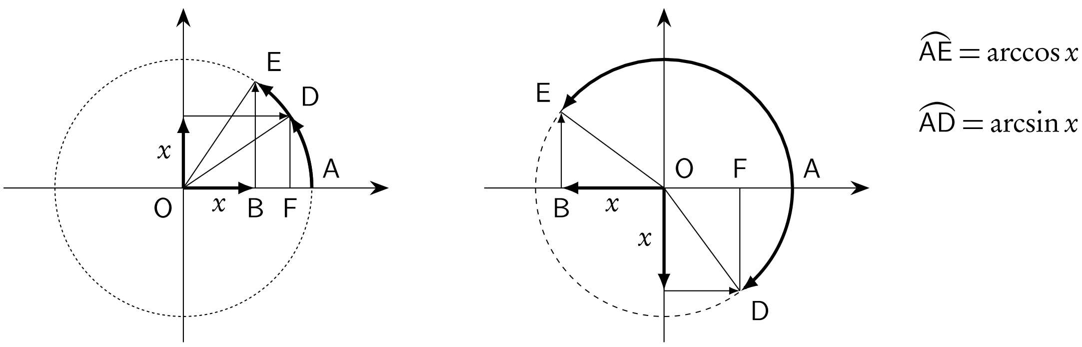


Figura 11.39.: $\arcsin x + \arccos x = \pi/2$

Una relazione identica alla (11.59) vale anche per le funzioni arctan e arccot:

$$\arctan x + \operatorname{arccot} x = \frac{\pi}{2}, \quad x \in \mathbb{R}. \tag{11.60}$$

La dimostrazione della (11.60) si fa in maniera quasi identica a quella proposta per la precedente (11.59), riportando x una volta sulla retta $x = 1$ e una volta sulla $y = 1$ a partire, rispettivamente, dai punti $(1, 0)$ e $(0, 1)$.

La (11.60) permette di esprimere una delle due funzioni per mezzo dell'altra: in alcuni software di calcolo in effetti è definita solo la funzione arctan, mentre arccot si ottiene proprio per mezzo della relazione indicata.

$$\cos(\arcsin x) = \sqrt{1-x^2}, \quad \sin(\arccos x) = \sqrt{1-x^2}, \quad x \in [-1,1]. \tag{11.61}$$

Le (11.61) sono una conseguenza quasi immediata delle definizioni: nella prima, per esempio, si chiede di calcolare il coseno di un arco il cui seno è x; ebbene, se x è il seno di un arco, allora il suo coseno è $\pm\sqrt{1-x^2}$. Il fatto che si debba prendere solo il segno "+" davanti al radicale è dovuto al fatto che $\arcsin x$ appartiene al primo o al quarto quadrante, dove il coseno è positivo. Stesso ragionamento per la seconda: se x è il coseno di un angolo, allora il suo seno è $\pm\sqrt{1-x^2}$. Il fatto che si debba prendere solo il segno "+" davanti al radicale è dovuto al fatto che $\arccos x$ appartiene al primo o al secondo quadrante, dove il seno è positivo.

$$\sin(\arctan x) = \frac{x}{\sqrt{1+x^2}}, \quad x \in \mathbb{R}. \tag{11.62}$$

Anche in questo caso si tratta di una conseguenza quasi immediata delle definizioni: si chiede di trovare il seno di un arco la cui tangente è x. Basta ricordare la prima delle (11.20) per concludere con la (11.62), tenendo anche conto che c'è concordanza di segno, perché x e $\sin(\arctan x)$ sono o entrambi positivi o entrambi negativi.

$$\cos(\arctan x) = \frac{1}{\sqrt{1+x^2}}, \quad x \in \mathbb{R}. \tag{11.63}$$

Utilizzando la seconda delle (11.20), la (11.63) si prova subito, tenendo anche conto che, essendo $\arctan x$ nel primo o quarto quadrante, il coseno è sempre positivo.

$$\arctan x + \arctan\frac{1}{x} = \begin{cases} \dfrac{\pi}{2}, & \text{se } x > 0; \\ -\dfrac{\pi}{2}, & \text{se } x < 0. \end{cases} \tag{11.64}$$

Poniamo $\alpha = \arctan x$, $\beta = \arctan {}^1\!/x$ e cominciamo a considerare il caso $x > 0$. Allora sia α che β sono archi del primo quadrante e, tenendo conto delle (11.20), si trova

$$\sin\alpha = \frac{x}{\sqrt{1+x^2}}, \quad \sin\beta = \frac{\dfrac{1}{x}}{\sqrt{1+\left(\dfrac{1}{x}\right)^2}} = \frac{1}{\sqrt{1+x^2}} = \cos\alpha.$$

Questo basta per concludere che α e β sono complementari, ovvero $\alpha+\beta = \pi/2$. Se $x<0$ il discorso è quasi identico, salvo il fatto che, in questo caso, sia α che β sono negativi e quindi la loro somma è $-\pi/2$.

A proposito della funzione $\arctan {}^1\!/x$ è utile osservare che il suo grafico, a parte il caso $x=0$ dove non è definita, coincide con il grafico della funzione inversa della restrizione della cotangente, secondo la scelta effettuata da Mathematica e Matlab: è anche questo, forse, uno dei motivi di questa scelta, che, lo ribadiamo, è diversa da quella da noi fatta.

11.10. Ricerca del periodo per funzioni elementari

Nelle applicazioni è molto importante saper riconoscere le funzioni periodiche da quelle non periodiche ed eventualmente determinarne il periodo. Purtroppo la cosa non è affatto semplice, anche limitandosi alle funzioni trigonometriche e a quelle da esse ottenute mediante somme, prodotti, quozienti o composizioni. Per renderci conto della difficoltà del problema, è opportuno considerare alcuni esempi.

Esempio 11.6. Le funzioni $f(x) = \sin x$ e $g(x) = 1 - \sin x$ sono entrambe periodiche di periodo 2π. La loro somma è però la funzione costante $h(x) = 1$ che, come già osservato, non ha un minimo periodo e quindi non viene considerata periodica. Un esempio simile si può costruire con le funzioni $f(x) = \sin^2 x$ e $g(x) = \cos^2 x$, la cui somma è 1.

Esempio 11.7. Le due funzioni $f(x) = \sin x$ e $g(x) = \sin(\pi x)$ hanno periodo, rispettivamente, 2π e 2. La loro somma non è nemmeno una funzione periodica. Discorso analogo vale per il prodotto o il quoziente di queste due funzioni.

Esempio 11.8. Le due funzioni $f(x) = \sin x$ e $g(x) = \cos x$ sono entrambe periodiche di periodo 2π. Il loro quoziente è invece la funzione $h(x) = \tan x$ che è periodica di periodo π. Discorso analogo sarebbe vero per il prodotto.

Esempio 11.9. Le due funzioni $f(x) = \sin x$ e $g(x) = 1 + \cos x$ sono entrambe periodiche di periodo 2π. Il loro quoziente è la funzione $h(x) = \sin x/(1 + \cos x)$ che è ancora periodica di periodo 2π.

Esempio 11.10. La funzione $f(x) = \sin x$ è periodica di periodo 2π. Il suo valore assoluto, $g(x) = |\sin x|$, è periodica di periodo π.

Esempio 11.11. La funzione $f(x) = \sin x + 2$ è periodica di periodo 2π. Il suo valore assoluto, $g(x) = |\sin x + 2|$, è ancora periodica di periodo 2π.

Riportiamo alcune poche indicazioni, valide in particolare per le funzioni trigonometriche, a cui attenersi nella ricerca del periodo: al di fuori di queste, non esistono regole generali.

1. Se una funzione $f(x)$ è periodica di periodo T, allora la funzione $f(kx)$, con k reale diverso da zero, è periodica di periodo $T/|k|$.
2. Se si hanno due funzioni periodiche con *diverso* periodo T_1 e T_2, e se esistono multipli interi comuni dei due periodi, allora le funzioni somma, prodotto, quoziente, hanno periodo uguale al minimo comune multiplo dei periodi.
3. Se si hanno due funzioni periodiche con *lo stesso* periodo T, allora le funzioni somma, prodotto, quoziente, hanno periodi minori o uguali al periodo comune T.

Le funzioni trigonometriche giocano un ruolo molto importante nell'insieme delle funzioni periodiche. Questa osservazione è molto più profonda di quanto non possa sembrare a prima vista: in sostanza le funzioni seno e coseno sono i modelli di base di tutte le funzioni periodiche, purché non troppo patologiche. Questo fatto è contenuto nell'importantissimo teorema di sviluppabilità di una funzione in serie di Fourier, di cui riportiamo qui solo l'enunciato in una forma molto schematica.

Teorema 11.16. *Data una funzione sufficientemente regolare e periodica di periodo 2π[7], vale, con qualche limitazione, la seguente uguaglianza*

$$f(x) = k + \sum_{n=0}^{+\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big). \tag{11.65}$$

Questo vuol dire, in un senso che sarà precisato nei successivi corsi di analisi, che tutte le funzioni periodiche possono essere pensate, entro certi limiti, come combinazioni delle funzioni seno e coseno. Tra le applicazioni famose di questo teorema citiamo la scomposizione di un suono nei suoi armonici, scomposizione che sta alla base dell'acustica musicale.

11.11. Risoluzione di triangoli

Tra le applicazioni più importanti delle funzioni trigonometriche che abbiamo introdotto ci sono i teoremi relativi ai triangoli. Useremo le seguenti convenzioni sui triangoli, in particolare sui triangoli rettangoli: Se A, B e C sono i vertici, a, b e c sono i lati opposti, α, β e γ sono gli angoli interni che hanno quei vertici. Se il triangolo è rettangolo, A è il vertice dell'angolo retto, quindi a è l'ipotenusa e $\alpha = \pi/2$. Tutti gli angoli sono considerati nel senso della geometria euclidea piana (dunque sempre positivi e non generalizzati).

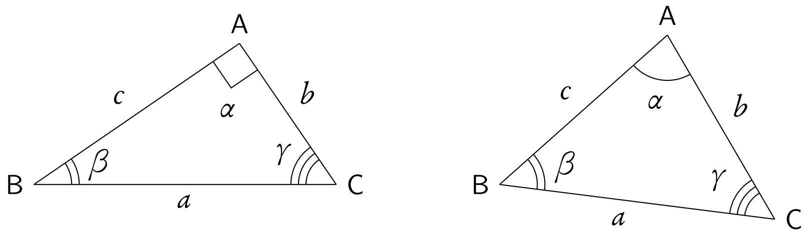


Figura 11.40.: *Convenzioni di nomenclatura sui triangoli*

Teorema 11.17 (Teorema sui triangoli rettangoli)**.** *In un triangolo rettangolo ciascun cateto è uguale al prodotto dell'ipotenusa per il seno dell'angolo opposto, oppure all'ipotenusa per il coseno dell'angolo acuto adiacente, oppure all'altro cateto per la tangente dell'angolo opposto.*

In formule il teorema 11.17 si esprime come segue:

$$b = a\sin\beta = a\cos\gamma = c\tan\beta; \tag{11.66}$$

$$c = a\sin\gamma = a\cos\beta = b\tan\gamma. \tag{11.67}$$

Il contenuto sostanziale del teorema 11.17, unitamente al teorema di Pitagora, è il seguente: noti, in un triangolo rettangolo, due elementi[8] diversi dall'angolo retto e che non siano i due angoli acuti, si possono trovare tutti gli altri elementi. In effetti:

[7] Questa limitazione può in realtà essere rimossa, ma non intendiamo approfondire qui l'argomento.
[8] Ricordiamo che si chiamano "elementi" di un triangolo i tre lati e i tre angoli.

— noti due lati si può trovare il terzo con il teorema di Pitagora; successivamente si possono trovare gli angoli acuti (o meglio il loro seno e coseno) mediante le formule (11.66) o (11.67).
— noti un lato e un angolo si può trovare un altro lato mediante le formule (11.66) o (11.67) e successivamente l'altro lato ancora mediante le stesse formule o con il teorema di Pitagora.

Esempio 11.12. Se $b = 3$ e $c = 4$, con Pitagora si trova $a = 5$ e, successivamente, $\sin\beta = 3/5$, $\sin\gamma = 4/5$.

Osservazione 11.18. Si noti che, in generale, quello che conta è trovare il seno oppure il coseno di un angolo, oltre a valutare in quale quadrante si trova l'angolo stesso. Solo alla conclusione di un problema potrà essere necessario dare una valutazione approssimata dell'angolo stesso. Nell'esempio precedente, individuato il seno di β, se ne possono trovare il coseno o la tangente mediante le formule trigonometriche (sapendo che β è acuto) ed eventualmente usare questi valori per altre parti del problema: solo alla fine, se richiesto, si potrà scrivere, $\beta \simeq 0.6435$, oppure $\beta \simeq 36.8699°$.

Sia ora data un corda $\overline{\mathsf{AB}} = a$ in una circonferenza di centro O e raggio r. La corda individua sulla circonferenza due archi, entrambi di estremi A e B, un arco maggiore e uno minore, a meno che la corda stessa non sia un diametro nel qual caso i due archi sono entrambi uguali alla semicirconferenza. Se indichiamo con α uno degli angoli alla circonferenza (tutti tra di loro uguali!) che insistono sull'arco minore e con β uno degli angoli alla circonferenza (ancora tutti tra di loro uguali) che insistono sull'arco maggiore, si deve avere $\alpha + \beta = \pi$, in quanto la somma dei corrispondenti angoli al centro, $2\alpha + 2\beta$, è 2π. Gli angoli α e β, che insistono sulla stessa corda, ma non sullo stesso arco, sono dunque supplementari e hanno lo stesso seno. È appena il caso di ricordare, perché sarà utile in seguito, che tra gli angoli α come tra gli angoli β ci sono sia angoli con entrambi i lati secanti che angoli (esattamente due di ciascun tipo) con un lato secante e uno tangente.

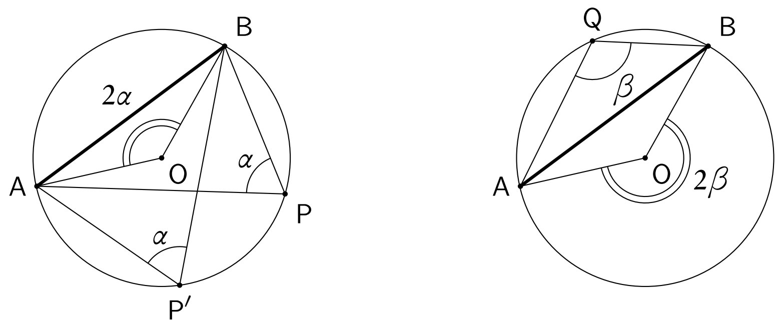


Figura 11.41.: *Angoli alla circonferenza che insistono sulla stessa corda*

Tenendo conto che tra gli angoli come α ce ne sono due che hanno un lato passante per il centro, e dunque sono angoli acuti di un triangolo rettangolo di ipotenusa uguale al diametro, si può concludere con il seguente teorema.

Teorema 11.19 (Teorema della corda)**.** *Data, in una circonferenza di raggio r, una corda a, e detti α e β gli angoli che insistono sui due archi individuati dalla corda, si ha*

$$a = 2r\sin\alpha = 2r\sin\beta. \tag{11.68}$$

Tenendo conto del fatto che ogni triangolo si può inscrivere in una circonferenza, di cui i lati sono corde, si ottiene subito il seguente teorema.

Teorema 11.20 (Teorema dei seni o di Eulero). *In ogni triangolo il rapporto tra un lato e il seno dell'angolo opposto è costante e uguale al diametro 2R della circonferenza circoscritta.*

$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma} = 2R. \tag{11.69}$$

Questo teorema consente di risolvere un triangolo, cioè di determinarne tutti gli elementi quando siano dati due angoli e un lato oppure due lati e un angolo, diverso da quello compreso tra i due lati. Vediamo in dettaglio le due situazioni.

Risoluzione di un triangolo dati due angoli e un lato

Dati, per esempio, gli angoli α e β e il lato c, si trova subito $\gamma = \pi - \alpha - \beta$ e quindi $c/\sin\gamma$. Dal teorema dei seni si ha, allora

$$\frac{a}{\sin\alpha} = \frac{c}{\sin\gamma},$$

da cui si trova a, e infine

$$\frac{b}{\sin\beta} = \frac{c}{\sin\gamma},$$

da cui si ricava b.

Risoluzione di un triangolo dati due lati e un angolo non compreso tra essi

Dati, per esempio, c, a, e l'angolo α, sia H il piede della perpendicolare condotta da B alla semiretta AC. Nel triangolo rettangolo ABH è nota l'ipotenusa $\overline{\mathrm{AB}} = c$ e l'angolo α. Si può dunque trovare il cateto $\overline{\mathrm{BH}} = h$. È immediato che, essendo BHC un triangolo rettangolo, eventualmente degenere, con ipotenusa $\overline{\mathrm{BC}}$, deve essere $a \geq h$. Se $a = h$ esiste un solo triangolo rettangolo ABC con $\mathrm{C} \equiv \mathrm{H}$. Se poi $a \geq c$, esiste un solo triangolo. Se, invece $h < a < c$ esistono due triangoli. In questi ultimi due casi l'angolo γ (o meglio il suo seno), si può trovare usando teorema dei seni:

$$\frac{a}{\sin\alpha} = \frac{c}{\sin\gamma} \quad \Rightarrow \quad \sin\gamma = \frac{c}{a}\sin\alpha.$$

La figura 11.42 illustra le tre situazioni.

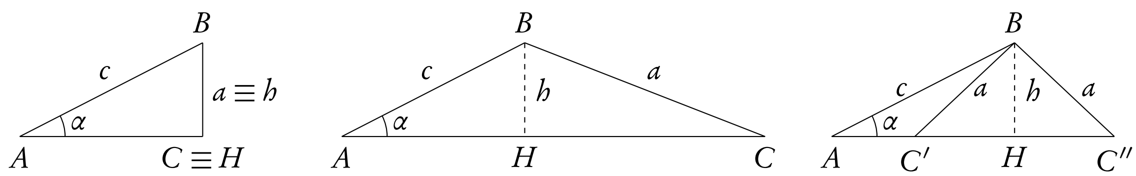


Figura 11.42.: *Risoluzione di un triangolo dati due lati e un angolo non compreso*

Teorema 11.21 (Teorema del coseno o di Carnot). *In ogni triangolo si ha*

$$a^2 = b^2 + c^2 - 2bc\cos\alpha. \tag{11.70}$$

Formule analoghe per b^2 e c^2.

Anche se insito nelle notazioni usate, conviene ribadire che α è l'angolo compreso tra i due lati b e c. Questo teorema può essere considerato un'estensione del teorema di Pitagora, in quanto se α è retto, b e c sono i cateti e a è l'ipotenusa e la (11.70) si riduce proprio al teorema di Pitagora ($\cos\alpha = 0$).

Questo teorema consente di risolvere un triangolo quando sono dati due lati e l'angolo compreso, oppure quando sono dati i tre lati. Per il primo caso è sufficiente osservare che dalla (11.70), noti b, c e l'angolo α fra essi compreso si trova subito il terzo lato a; per il secondo caso, sempre dalla (11.70), noti a, b e c si ricava α e poi si può usare il teorema dei seni per trovare uno dei restanti angoli.

Citiamo anche tre altri teoremi sui triangoli qualunque, anche se il loro uso è molto limitato e comunque non è indispensabile: i teoremi dei seni e del coseno sono sufficienti a trattare tutti i casi.

Teorema 11.22 (Teorema delle proiezioni)**.** *In ogni triangolo valgono le seguenti formule.*

$$a = b\cos\gamma + c\cos\beta\,; \quad b = c\cos\alpha + a\cos\gamma\,; \quad c = a\cos\beta + b\cos\alpha\,. \tag{11.71}$$

Il nome di questo teorema è legato al fatto che esso può essere formulato a parole dicendo che in un triangolo un lato è la somma delle proiezioni con segno degli altri due lati.

Teorema 11.23 (Teorema di Nepero o delle tangenti)**.** *In un triangolo qualunque valgono le seguenti formule.*

$$\frac{a-b}{a+b} = \frac{\tan\frac{\alpha-\beta}{2}}{\tan\frac{\alpha+\beta}{2}}\,; \qquad \frac{b-c}{b+c} = \frac{\tan\frac{\beta-\gamma}{2}}{\tan\frac{\beta+\gamma}{2}}\,; \qquad \frac{c-a}{c+a} = \frac{\tan\frac{\gamma-\alpha}{2}}{\tan\frac{\gamma+\alpha}{2}}\,. \tag{11.72}$$

Questo teorema può essere usato, in alternativa al teorema del coseno, per risolvere un triangolo nel caso siano dati due lati e l'angolo compreso. Per esempio, se sono dati a, b e γ, è noto anche $\alpha+\beta = \pi-\gamma$. Dalla prima delle (11.72) si può allora ricavare la tangente della semidifferenza di α e β e quindi $\alpha-\beta$. Noti $\alpha+\beta$ e $\alpha-\beta$ si possono facilmente ricavare sia α che β.

Teorema 11.24 (Teorema di Briggs)**.** *In un triangolo qualunque, indicato con p il semiperimetro, valgono le seguenti formule.*

$$\sin\frac{\alpha}{2} = \sqrt{\frac{(p-b)(p-c)}{bc}}\,; \quad \sin\frac{\beta}{2} = \sqrt{\frac{(p-c)(p-a)}{ca}}\,; \quad \sin\frac{\gamma}{2} = \sqrt{\frac{(p-a)(p-b)}{ab}}\,. \tag{11.73}$$

$$\cos\frac{\alpha}{2} = \sqrt{\frac{p(p-a)}{bc}}\,; \quad \cos\frac{\beta}{2} = \sqrt{\frac{p(p-b)}{ca}}\,; \quad \cos\frac{\gamma}{2} = \sqrt{\frac{p(p-c)}{ab}}\,. \tag{11.74}$$

Questo teorema può essere usato, ancora in alternativa al teorema del coseno, per risolvere un triangolo in cui siano dati i tre lati.

Una quasi immediata applicazione dei teoremi sui triangoli rettangoli porta al seguente teorema concernente l'area di un triangolo.

Teorema 11.25 (Area di un triangolo)**.** *L'area di un triangolo qualunque è data dal semiprodotto di due lati per il seno dell'angolo compreso. In formule*

$$\mathcal{A} = \frac{1}{2}ab\sin\gamma = \frac{1}{2}bc\sin\alpha = \frac{1}{2}ac\sin\beta\,. \tag{11.75}$$

Conseguenza di questo teorema è poi il seguente per il calcolo dell'area di un quadrilatero convesso.

Teorema 11.26 (Area di un quadrilatero convesso)**.** *L'area di un quadrilatero convesso è uguale al semiprodotto delle diagonali per il seno di uno degli angoli da esse individuati.*

Concludiamo questo excursus sulle applicazioni della trigonometria ai triangoli e poligoni con il teorema di Tolomeo.

Teorema 11.27 (Teorema di Tolomeo)**.** *In un quadrilatero convesso inscrivibile in una circonferenza, il prodotto delle misure delle diagonali è uguale alla somma dei prodotti delle misure dei lati opposti.*

11.12. Equazioni e disequazioni goniometriche

Si chiamano *goniometriche* quelle equazioni o disequazioni in cui l'incognita figura come argomento di una funzione trigonometrica. Anche per queste disequazioni, come già per quelle irrazionali e per quelle logaritmiche ed esponenziali, non esistono metodi generali di soluzione e spesso si possono applicare solo metodi numerici, come si può vedere dal seguente esempio.

Esempio 11.13. Si debba risolvere la disequazione

$$\frac{1}{x} + \sin x > 0.$$

Si può esaminare la figura 11.43, dove abbiamo tracciato il grafico della funzione a primo membro. La funzione è dispari, per cui basterà risolverla per $x > 0$. Si può dedurre dal grafico che l'insieme delle soluzioni è

$$]0, x_A[\ \cup\ ]x_B, x_C[\ \cup\ ]x_D, x_E[\ \cup \dots$$

I valori x_A, x_B, x_C, ..., possono essere determinati solo per via numerica, anche se si può notare che si avvicinano sempre di più a multipli di π, in quanto al crescere di x la quantità $1/x$ diventa via via più piccola. Si trova

$$x_A \simeq 3.42; \quad x_B \simeq 6.11; \quad x_C \simeq 9.53; \quad x_D \simeq 12.47; \quad x_E \simeq 15.79,$$

mentre per π e i suoi multipli si ha

$$\pi \simeq 3.14; \quad 2\pi \simeq 6.28; \quad 3\pi \simeq 9.42; \quad 4\pi \simeq 12.56; \quad 5\pi \simeq 15.71.$$

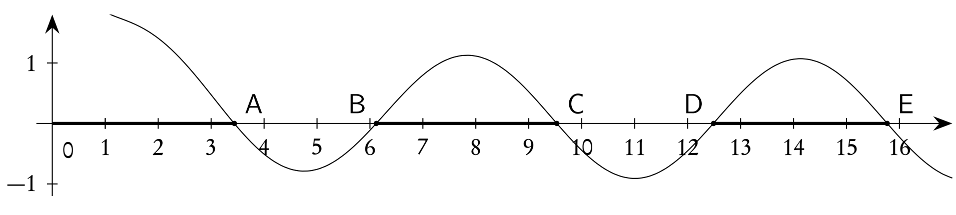


Figura 11.43.: *Grafico della funzione* $x + \sin 1/x$

Nel seguito tratteremo alcuni tipi di equazioni o disequazioni che possono essere risolti con strategie semplici, in particolare le equazioni o disequazioni elementari a cui le altre possono venire ricondotte.

Le tecniche risolutive delle equazioni e disequazioni sono sostanzialmente identiche: ci occuperemo esplicitamente quasi sempre di disequazioni, da cui si possono dedurre le strategie che servono per risolvere le equazioni. Come al solito, ragioneremo su alcuni esempi per chiarire i metodi.

11.12.1. Disequazioni elementari

Detto a un numero reale, si chiamano *elementari* le disequazioni di uno dei tipi

$$\sin x \lesseqgtr a; \quad \cos x \lesseqgtr a; \quad \tan x \lesseqgtr a; \quad \cot x \lesseqgtr a. \tag{11.76}$$

Anche se non tratteremo esplicitamente equazioni o disequazioni coinvolgenti le funzioni secante e cosecante, le tecniche proposte sono identiche e, in ogni caso, queste disequazioni possono essere ricondotte a equazioni e disequazioni coinvolgenti il seno e il coseno.

È consigliabile servirsi di una tecnica grafica per risolvere questo tipo di disequazioni, tecnica che può prevedere o l'uso della circonferenza goniometrica o quello dei grafici delle funzioni goniometriche. Negli esempi che seguono li utilizzeremo entrambi.

Esempio 11.14. Risolvere la disequazione

$$\sin x > \frac{1}{2}.$$

Posto $Y = \sin x$ e $X = \cos x$, la disequazione data è equivalente al sistema

$$\begin{cases} Y > \dfrac{1}{2} \\ X^2 + Y^2 = 1 \end{cases}.$$

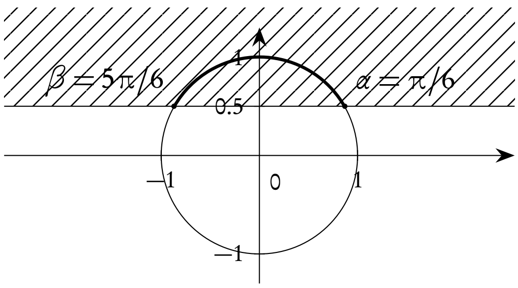


Figura 11.44.: *La disequazione* $\sin x > 1/2$

Utilizzando la figura 11.44, e tenendo conto della periodicità, si conclude subito che l'insieme di soluzioni è dato da

$$\frac{\pi}{6} + 2k\pi < x < \frac{5\pi}{6} + 2k\pi, \quad k \in \mathbb{Z}.$$

L'uso del grafico della funzione seno porta allo stesso risultato, ma è molto più significativo, perché rende immediatamente conto della periodicità. Tracciato il grafico della funzione $f(x) = \sin x$, si riporta il valore 1/2 sull'asse delle ordinate. Si tratta di trovare i valori di x sull'asse delle ascisse il cui seno supera 1/2, ovvero in corrispondenza ai quali il grafico "sta sopra" la retta $y = 1/2$. Il grafico rende evidente che i valori richiesti sono quelli compresi tra α e β e che poi la situazione si ripete identicamente sia a sinistra che a destra ad intervalli regolari di 2π. In questo caso, poiché il valore sull'asse delle ordinate è 1/2, i valori sull'asse delle ascisse possono essere espressi in termini di multipli e sottomultipli di π, ma di solito non si è così fortunati e si fa ricorso alle funzioni trigonometriche inverse e a determinazioni numeriche approssimate dei valori cercati. Si veda la figura 11.45.

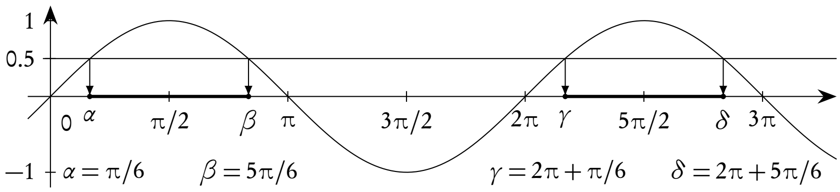


Figura 11.45.: *Ancora la disequazione* $\sin x > 1/2$

Esempio 11.15. Risolvere la disequazione

$$\cos x < \frac{1}{3}.$$

Posto $Y = \sin x$ e $X = \cos x$, la disequazione data è equivalente al sistema

$$\begin{cases} X < \dfrac{1}{3} \\ X^2 + Y^2 = 1 \end{cases}.$$

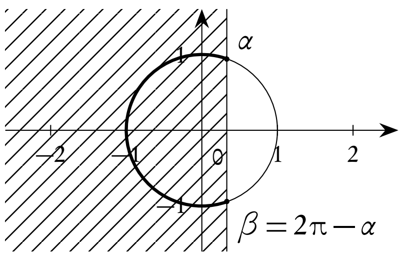


Figura 11.46.: *La disequazione* $\cos x < 1/3$

Osservato che

$$\alpha = \arccos\left(\frac{1}{3}\right),$$

utilizzando la figura 11.46, e tenendo conto della periodicità, si conclude subito che l'insieme di soluzioni è dato da

$$\alpha + 2k\pi < x < 2\pi - \alpha + 2k\pi, \quad k \in \mathbb{Z}, \quad \alpha = \arccos\left(\frac{1}{3}\right) \simeq 1.2309.$$

L'uso del grafico della funzione coseno porta allo stesso risultato, ma è molto più significativo, perché rende immediatamente conto della periodicità. Tracciato il grafico della funzione $f(x) = \cos x$, si riporta il valore $1/3$ sull'asse delle ordinate. Si tratta di trovare i valori di x sull'asse delle ascisse il cui coseno non supera $1/3$, ovvero in corrispondenza ai quali il grafico "sta sotto" la retta $y = 1/3$. Il grafico mostra subito che i valori richiesti sono quelli compresi tra α e β e che poi la situazione si ripete identicamente sia a sinistra che a destra ad intervalli regolari di 2π. In questo caso, poiché il valore sull'asse delle ordinate è $1/3$, i valori sull'asse delle ascisse non possono essere espressi in termini di multipli e sottomultipli di π, e bisogna fare ricorso alle funzioni trigonometriche inverse. In problemi come questo è opportuno denominare l'angolo per esempio con α e non "portarsi dietro" la scrittura $\arccos(1/3)$ che è troppo pesante. Qualunque calcolatrice tascabile fornisce facilmente il valore approssimato $\arccos(1/3) \simeq 1.2309$. Si veda la figura 11.47.

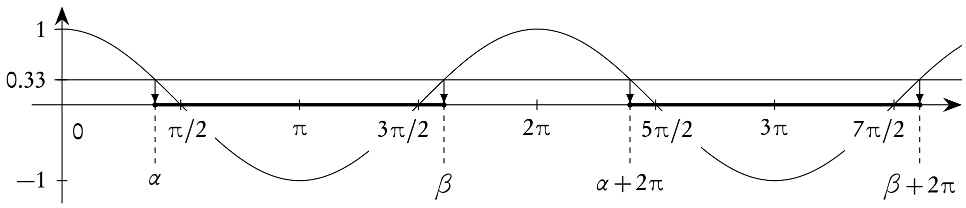


Figura 11.47.: *Ancora la disequazione* $\cos x < 1/3$

Esempio 11.16. Risolvere la disequazione

$$\tan x > \sqrt{3}.$$

Nelle disequazioni elementari coinvolgenti la funzione tangente o cotangente si deve ricordare che la periodicità è ora di π anziché 2π. Usando la circonferenza goniometrica ci si può limitare solo a una semicirconferenza: conviene scegliere la semicirconferenza giacente nel 1° e 4° quadrante per la tangente, nel 1° e 2° quadrante per la cotangente.

Posto, come nei precedenti casi, $Y = \sin x$ e $X = \cos x$ la disequazione data è equivalente al sistema

$$\begin{cases} \dfrac{Y}{X} > \sqrt{3} \\ X^2 + Y^2 = 1 \end{cases}.$$

Se teniamo conto del fatto che ci limitiamo a considerare solo il primo e quarto quadrante, avremo $X > 0$. Nel sistema precedente si può dunque semplificare la disequazione, moltiplicando ambo i membri per X.

$$\begin{cases} Y > X\sqrt{3} \\ X^2 + Y^2 = 1 \end{cases}.$$

A questo punto possiamo procedere come nel caso del seno e del coseno, considerando, lo ribadiamo, solo il primo e quarto quadrante.

Utilizzando la figura 11.48, e tenendo conto della periodicità, si conclude subito che l'insieme di soluzioni è dato da

$$\frac{\pi}{3} + k\pi < x < \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}.$$

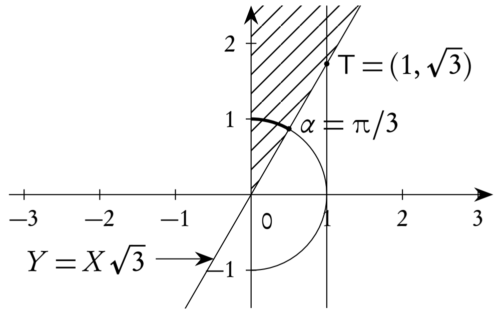


Figura 11.48.: *La disequazione* $\tan x > \sqrt{3}$

L'uso del grafico della funzione tangente porta allo stesso risultato, ma è molto più significativo, perché rende immediatamente conto della periodicità. Tracciato il grafico della funzione $f(x) = \tan x$, si riporta il valore $\sqrt{3}$ sull'asse delle ordinate. Si tratta di trovare i valori di x sull'asse delle ascisse la cui tangente supera $\sqrt{3}$, ovvero in corrispondenza ai quali il grafico "sta sopra" la retta $y = \sqrt{3}$. L'esame della figura 11.49 consente di concludere facilmente.

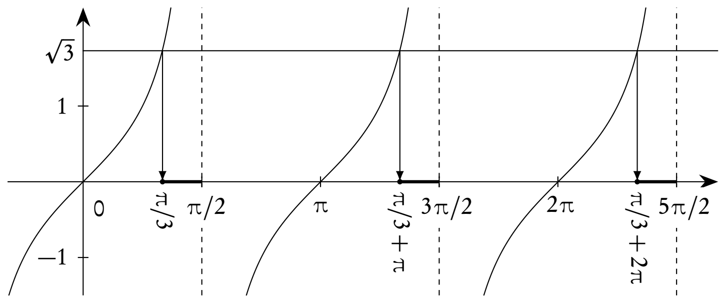


Figura 11.49.: *Ancora la disequazione* $\tan x > \sqrt{3}$

Esempio 11.17. Risolvere la disequazione

$$\cot x > -2.$$

Posto ancora $Y = \sin x$ e $X = \cos x$, la disequazione data è equivalente al sistema

$$\begin{cases} \dfrac{X}{Y} > -2 \\ X^2 + Y^2 = 1 \end{cases}.$$

Se teniamo conto del fatto che ci limitiamo a considerare solo il primo e secondo quadrante, avremo $Y > 0$. Nel sistema precedente si può dunque semplificare la disequazione, moltiplicando ambo i membri per Y.

$$\begin{cases} X > -2Y \\ X^2 + Y^2 = 1 \end{cases}.$$

A questo punto possiamo procedere come nel caso del seno e del coseno, considerando, lo ribadiamo, solo il primo e secondo quadrante.

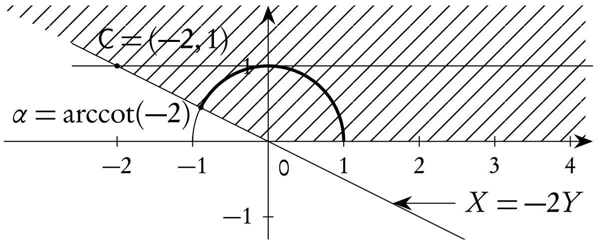


Figura 11.50.: *La disequazione* $\cot x > -2$

Utilizzando la figura 11.50, e tenendo conto della periodicità, si conclude subito che l'insieme di soluzioni è dato da

$$k\pi < x < \alpha + k\pi, \quad \alpha = \operatorname{arccot}(-2) \simeq 2.67795, \quad k \in \mathbb{Z}.$$

L'uso del grafico della funzione cotangente porta allo stesso risultato, ma, al solito, è molto più significativo, perché rende immediatamente conto della periodicità. Tracciato il grafico della funzione $f(x) = \cot x$, si riporta il valore -2 sull'asse delle ordinate. Si tratta di trovare i valori di x sull'asse delle ascisse la cui cotangente supera -2, ovvero in corrispondenza ai quali il grafico "sta sopra" la retta $y = -2$. L'esame della figura 11.51 consente di concludere facilmente.

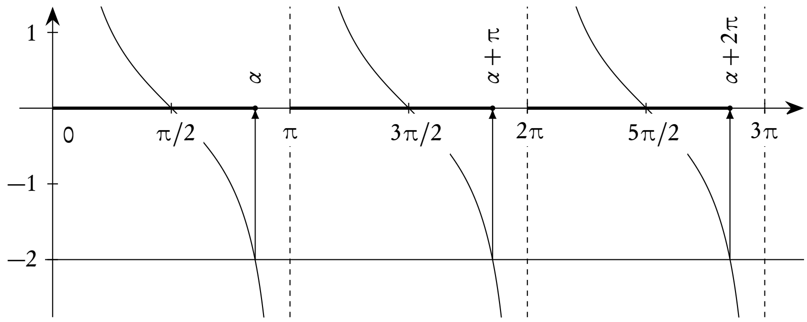


Figura 11.51.: *Ancora la disequazione* $\cot x > -2$

11.12.2. Disequazioni lineari in seno e coseno

Si dicono *disequazioni lineari in seno e coseno* le disequazioni del tipo

$$a \sin x + b \cos x + c \lesseqgtr 0. \tag{11.77}$$

Queste disequazioni possono essere risolte con diverse strategie. La prima prevede di procedere come nelle disequazioni elementari, ponendo $X = \cos x$, $Y = \sin x$.

Esempio 11.18. Risolvere la disequazione

$$\sin x + \sqrt{3} \cos x \leq \sqrt{3}.$$

Con le posizioni indicate, la disequazione diventa equivalente al sistema

$$\begin{cases} Y+\sqrt{3}X \leq \sqrt{3} \\ X^2+Y^2=1 \end{cases}.$$

Conviene trovare preventivamente gli (eventuali) punti di intersezione della retta origine del semipiano individuato dalla disequazione del sistema con la circonferenza. Si trova facilmente

$$A=(1,0), \quad B=\left(\frac{\frac{1}{2},\sqrt{3}}{2}\right).$$

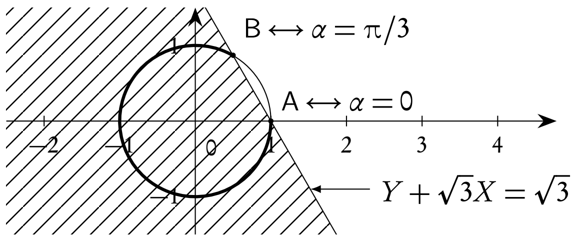


Figura 11.52.: *La disequazione* $\sin x+\sqrt{3}\cos x \leq \sqrt{3}$

Dall'esame della figura 11.52 si conclude facilmente che le soluzioni della disequazione sono:

$$\frac{\pi}{3}+2k\pi \leq x \leq 2\pi+2k\pi, \quad k \in \mathbb{Z}.$$

Un'altra strategia molto efficiente e utile anche in altre circostanze per trattare le disequazioni del tipo (11.77) consiste nel trasformarle in una disequazione elementare. Si può procedere come di seguito indicato.

Poiché $a^2+b^2 \neq 0$, si possono dividere ambo i membri della (11.77) per $\sqrt{a^2+b^2}$:

$$\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{b}{\sqrt{a^2+b^2}}\cos x+\frac{c}{\sqrt{a^2+b^2}} \lesseqgtr 0.$$

Poiché

$$\left(\frac{a}{\sqrt{a^2+b^2}}\right)^2+\left(\frac{b}{\sqrt{a^2+b^2}}\right)^2=1,$$

esiste sicuramente α tale che

$$\cos\alpha=\frac{a}{\sqrt{a^2+b^2}} \quad \wedge \quad \sin\alpha=\frac{b}{\sqrt{a^2+b^2}}.$$

La (11.77) può allora essere messa nella forma

$$\cos\alpha\sin x+\sin\alpha\cos x+\frac{c}{\sqrt{a^2+b^2}} \lesseqgtr 0,$$

ovvero

$$\sin(x+\alpha)+\frac{c}{\sqrt{a^2+b^2}} \lesseqgtr 0,$$

che diventa una disequazione elementare con la posizione $x+\alpha=t$.

Riprendendo in esame l'esempio precedente, la disequazione data si può riscrivere nella forma

$$\frac{1}{2}\sin x+\frac{\sqrt{3}}{2}\cos x \le \frac{\sqrt{3}}{2} \quad \Rightarrow \quad \cos\frac{\pi}{3}\sin x+\sin\frac{\pi}{3}\cos x \le \frac{\sqrt{3}}{2} \quad \Rightarrow \quad \sin\left(x+\frac{\pi}{3}\right) \le \frac{\sqrt{3}}{2},$$

ovvero

$$\sin t \le \frac{\sqrt{3}}{2}.$$

Quest'ultima risulta verificata[9] per

$$\frac{2}{3}\pi+2k\pi \le t \le \frac{7}{3}\pi+2k\pi, \quad k\in\mathbb{Z},$$

che porge le stesse soluzioni di prima, tenendo conto della sostituzione inversa da effettuare.

Si potrebbero anche usare le formule parametriche, trasformando una disequazione del tipo (11.77) in un disequazione elementare in funzione di $\tan x/2$. Tuttavia le formule parametriche non solo valide per tutti i valori di x, e bisognerebbe valutare a parte che cosa succede per i valori di x esclusi: per questo sconsigliamo, in generale, questo modo di procedere.

11.12.3. Disequazioni con una sola funzione trigonometrica

Sono le disequazioni del tipo $f(\sin x) \lesseqgtr 0$, o le analoghe che si ottengono sostituendo la funzione seno con un'altra funzione trigonometrica. Con la sostituzione $\sin x = t$, o le analoghe per le altre funzioni trigonometriche, si riducono ad una disequazione del tipo $f(t) \lesseqgtr 0$: se si è in grado di risolvere quest'ultima, con la sostituzione inversa si è ricondotti a una o più disequazioni elementari. La situazione più comune è quella in cui $f(x)$ è un polinomio di secondo grado $f(x)=ax^2+bx+c$, a volte di grado superiore, ma scomponibile in fattori. Un esempio chiarirà il metodo.

Esempio 11.19. Risolvere la disequazione

$$2\sin^2 x+\sin x-1>0.$$

Posto $\sin x=t$ si ottiene una disequazione di secondo grado verificata per $-1<t<1/2$. La disequazione proposta è dunque equivalente al sistema

$$\begin{cases} \sin x > -1 \\ \sin x < \dfrac{1}{2} \end{cases}.$$

[9] Volendo rimanere all'interno dell'intervallo $[0,2\pi]$, si sarebbero potute scrivere le soluzioni come

$$0\le t\le \frac{2}{3}\pi \vee \frac{2}{3}\pi \le t < 2\pi,$$

naturalmente tenendo poi conto della periodicità. Abbiamo preferito, invece, scrivere l'insieme delle soluzioni come un unico intervallo, per questioni di semplicità. È naturalmente solo questione di scelta, nulla cambia nella sostanza.

Con la abituale sostituzione $Y = \sin$, $X = \cos x$, si ottiene il sistema

$$\begin{cases} Y > -1 \\ Y < \dfrac{1}{2} \\ X^2 + Y^2 = 1 \end{cases}$$

che può essere risolto facendo uso della circonferenza goniometrica oppure del grafico della funzione seno. Nella figura 11.53 sono evidenziati ambo i metodi.

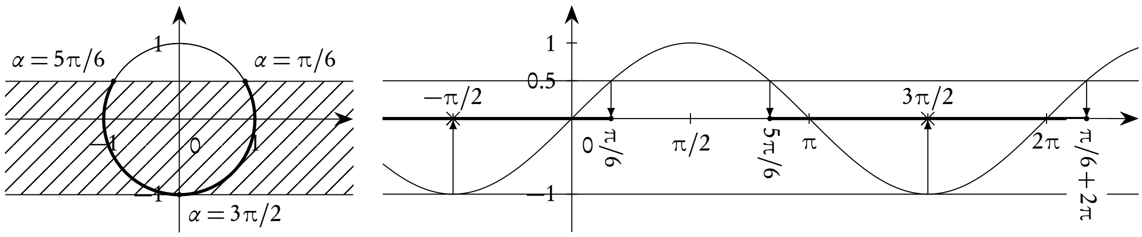


Figura 11.53.: *La disequazione* $2\sin^2 x + \sin x - 1 > 0$

Si conclude con le soluzioni seguenti:

$$-\frac{\pi}{2} + 2k\pi < x < \frac{\pi}{6} + 2k\pi \quad \vee \quad \frac{5\pi}{6} + 2k\pi < x < \frac{3\pi}{2} + 2k\pi.$$

11.12.4. Disequazioni omogenee di secondo grado in seno e coseno

Sono le disequazioni del tipo

$$a\sin^2 x + b\sin x\cos x + c\cos^2 x + d = 0. \tag{11.78}$$

Si possono risolvere trasformandole in equazioni lineari in seno e coseno, in funzione di $2x$, con le formule di bisezione e quelle di duplicazione del seno:

$$\sin^2 x = \frac{1 - \cos 2x}{2}, \quad \cos^2 x = \frac{1 + \cos 2x}{2}, \quad \sin x\cos x = \frac{1}{2}\sin 2x. \tag{11.79}$$

Esempio 11.20. Risolvere la disequazione

$$4\sin^2 x - 2\sqrt{3}\sin x\cos x - 2\cos^2 x - 1 > 0.$$

Utilizzando le (11.79) si ottiene

$$\sqrt{3}\cos 2x + \sin 2x < 0, \quad \text{ovvero, posto } 2x = t, \quad \sqrt{3}\cos t + \sin t < 0.$$

La disequazione in t ha le soluzioni

$$\frac{2}{3}\pi + 2k\pi < t < \frac{5}{3}\pi + 2k\pi, \quad k \in \mathbb{Z}.$$

Eseguendo la sostituzione inversa $t=2x$ e dividendo per 2, si ottiene infine

$$\frac{\pi}{3}+k\pi<t<\frac{5}{6}\pi+k\pi, \quad k\in\mathbb{Z}.$$

Si presti attenzione a dividere per 2 solo dopo aver scritto per esteso le soluzioni in t, compresa la periodicità.

È possibile anche seguire una diversa strategia, che comunque sconsigliamo[10] per i motivi che appariranno chiari ripetendo la risoluzione dell'esempio appena proposto. Poiché $\cos^2 x+\sin^2 x=1$, nulla cambia nella disequazione data se moltiplichiamo l'addendo -1 (il "termine noto") per $\cos^2 x+\sin^2 x$. Semplificando e riordinando si ottiene

$$3\sin^2 x-2\sqrt{3}\sin x\cos x-3\cos^2 x>0.$$

Dividendo per $\cos^2 x$, dopo aver osservato che i valori dove $\cos x=0$ sono soluzioni, si ottiene una disequazione di secondo grado nella funzione tangente.

$$3\tan^2 x-2\sqrt{3}\tan x-3>0,$$

che ha l'insieme di soluzioni già trovato, esclusi però i valori $\pi/2+k\pi$, dove la tangente non è definita. Questi valori vanno comunque accettati perché in corrispondenza ad essi $\cos x=0$ e, come già osservato, dove $\cos x=0$ la disequazione è verificata.

11.12.5. Disequazioni simmetriche in seno e coseno

Sono le disequazioni del tipo

$$a(\sin x\pm\cos x)+b\sin x\cos x+c \lesseqqgtr 0. \tag{11.80}$$

In realtà si chiamano più propriamente *simmetriche* quelle in cui compare il segno "+", *semisimmetriche* quelle in cui compare il segno "−". Il nome è dovuto al fatto che, nel primo caso, nulla cambia se sostituiamo il seno con il coseno.

Queste disequazioni si possono risolvere con la sostituzione $x=t+\pi/4$ che le trasforma in una disequazione di secondo grado in seno e coseno.

Esempio 11.21. Risolvere la disequazione

$$\sin x-\cos x-2\sqrt{2}\sin x\cos x<0.$$

Con la sostituzione indicata, dopo semplificazione, si previene alla disequazione

$$2\sin^2 t+\sin t-1<0,$$

che ha come soluzioni

$$\frac{5}{6}\pi+2k\pi<t<\frac{3}{2}\pi+2k\pi \quad \vee \quad \frac{3}{2}\pi+2k\pi<t<\frac{13}{6}\pi+2k\pi.$$

[10]Ne facciamo un cenno principalmente perché è proposta nella maggior parte dei testi scolastici in uso.

Effettuando la sostituzione inversa $t = x - \pi/4$ si ottiene infine

$$\frac{13}{12}\pi + 2k\pi < x < \frac{7}{4}\pi + 2k\pi \quad \vee \quad \frac{7}{4}\pi + 2k\pi < x < \frac{29}{12}\pi + 2k\pi.$$

Si noti che, sia nello scrivere le soluzioni in t che quelle in x, abbiamo preferito non limitarci all'intervallo $[0, 2\pi]$, per poter scrivere l'insieme di soluzioni in maniera più compatta.

11.12.6. Le altre disequazioni

Come già accennato, non esistono tecniche risolutive generali per le disequazioni trigonometriche. Per risolvere quelle che non rientrano nei modelli considerati si possono utilizzare i metodi appresi per le funzioni algebriche, opportunamente adattati, o usare le formule trigonometriche per cercare di ridurle ai modelli proposti. Proponiamo un esempio per chiarire il concetto.

Esempio 11.22. Risolvere la disequazione

$$\frac{\cos 2x + \sin x - 1}{(\sin x + \cos x)\big(1 - \sqrt{2}\sin x\big)} \geq 0\,.$$

Si deve preventivamente trovare il dominio e successivamente il segno di ognuno dei fattori. Per ottenere il risultato finale conviene utilizzare il solito grafico del tipo "+/−", con l'avvertenza che, trattandosi di funzioni periodiche, è sufficiente considerare un intervallo ampio quanto il periodo. Tra le operazioni preliminari ci sarà dunque anche quella di determinare il periodo, cosa non sempre agevole come mostra il caso della funzione tangente che è periodica di periodo π nonostante sia il rapporto di due funzioni di periodo 2π. Nei casi semplici, trovato il periodo delle singole funzioni, basterà prendere, se c'è, il minimo comune multiplo dei periodi.

Nel caso in esame per trovare il segno del numeratore bisognerà risolvere la disequazione $\cos 2x + \sin x - 1 > 0$: essa può essere ricondotta a una disequazione di secondo grado in seno usando le formule di duplicazione del coseno. Per trovare il segno dei due fattori del denominatore bisognerà risolvere le disequazioni $\sin x + \cos x > 0$ e $1 - \sqrt{2}\sin x > 0$, che sono di tipo elementare. Si può poi osservare che la funzione ha 2π tra i periodi (anzi 2π è il minimo periodo, ma questo fatto non è essenziale per il problema che stiamo trattando): basterà dunque considerare un grafico limitato all'intervallo $[0, 2\pi]$. Il grafico finale è il seguente.

+/−	0		$\frac{\pi}{6}$		$\frac{\pi}{4}$		$\frac{3\pi}{4}$		$\frac{5\pi}{6}$		π		$\frac{7\pi}{4}$		2π
$\cos 2x + \sin x - 1$	0	+	0	−	−	−	−	−	0	+	0	−	−	−	0
$\sin x + \cos x$	+	+	+	+	+	+	0	−	−	−	−	−	0	+	+
$1 - \sqrt{2}\sin x$	+	+	+	+	0	−	0	+	+	+	+	+	+	+	+
Complessivo	0	+	0	−	×	+	×	+	0	−	0	+	×	−	0

Le soluzioni, limitatamente all'intervallo $[0, 2\pi[$, sono dunque

$$0 \leq x \leq \frac{\pi}{6} \quad \vee \quad \frac{\pi}{4} < x < \frac{3\pi}{4} \quad \vee \quad \frac{3\pi}{4} < x \leq \frac{5\pi}{6} \quad \vee \quad \pi \leq x < \frac{7\pi}{4}.$$

11.12.7. Disequazioni con funzioni trigonometriche inverse

Le disequazioni coinvolgenti le funzioni trigonometriche inverse sono in genere abbastanza complesse e ci limiteremo qui a considerare solo alcuni casi molto semplici. Consigliamo di utilizzare sempre i grafici delle funzioni in oggetto. Proponiamo alcuni esempi.

Esempio 11.23. Risolvere la disequazione

$$\arcsin x < \frac{\pi}{3}.$$

Tracciato il grafico della funzione $\arcsin x$ e riportato il valore $\pi/3$ sull'asse delle ordinate, la figura 11.54 consente di concludere facilmente. Per determinare il valore $\sqrt{3}/2$ sull'asse delle ascisse basta ricordare che la funzione arcseno è l'inversa, nel tratto $[-\pi/2, \pi/2]$, della funzione seno, per cui sull'asse delle ordinate compaiono gli archi e sulle delle ascisse il seno. Si può anche osservare che la funzione arcseno è strettamente crescente, per cui si può prendere l'arcseno di ambo i membri della disequazione data, senza cambiare il verso. Tenendo conto che, in $[-1, 1]$, $\sin(\arcsin x) = x$, si ottiene

$$\sin(\arcsin x) < \sin\left(\frac{\pi}{3}\right) \quad \Rightarrow \quad x < \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2} \quad \Rightarrow \quad -1 \leq x < \frac{\sqrt{3}}{2},$$

ove per concludere abbiamo tenuto conto che la funzione $\arcsin x$ è definita in $[-1, 1]$.

Esempio 11.24. Risolvere la disequazione

$$\arccos x > \frac{3}{4}.$$

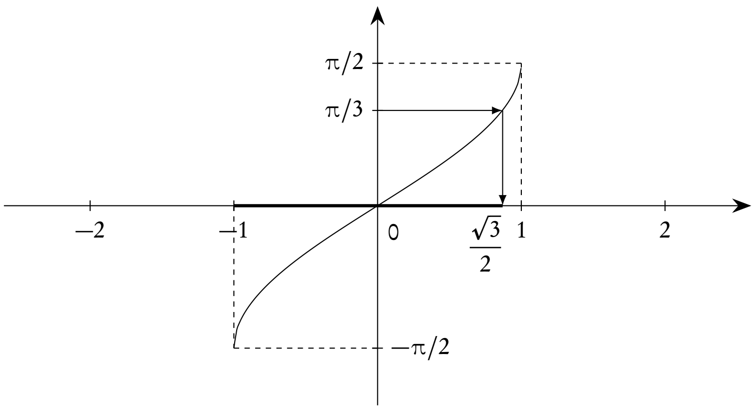


Figura 11.54.: *La disequazione* $\arcsin x < \pi/3$ *(Sistema non monometrico!)*

Tracciato il grafico della funzione $\arccos x$ e riportato il valore $3/4$ sull'asse delle ordinate, la figura 11.55 consente di concludere facilmente. Per determinare il valore $\cos(3/4)$ sull'asse delle ascisse basta ricordare che la funzione arccoseno è l'inversa, nel tratto $[0, \pi]$, della funzione coseno, per cui sull'asse delle ordinate compaiono gli archi e sulle delle ascisse il coseno. Si può anche osservare che la funzione arccoseno è strettamente decrescente, per cui si può prendere l'arccoseno di ambo i membri della disequazione data, cambiando però il verso. Tenendo conto che, in $[-1, 1]$, $\cos(\arccos x) = x$, si ottiene

$$\cos(\arccos x) < \cos\left(\frac{3}{4}\right) \quad \Rightarrow \quad x < \cos\left(\frac{3}{4}\right) \simeq 0.73 \quad \Rightarrow \quad -1 \le x < \cos\left(\frac{3}{4}\right),$$

ove per concludere abbiamo tenuto conto che $\arccos x$ è definita in $[-1, 1]$.

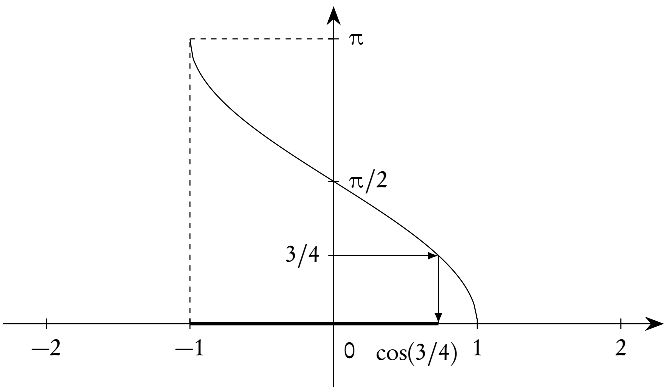


Figura 11.55.: *La disequazione* $\arccos x > 3/4$ *(Sistema non monometrico!)*

Esempio 11.25. Risolvere la disequazione

$$\arcsin^2 x - 3\arcsin x + 2 > 0.$$

Con la sostituzione $\arcsin x = t$ la disequazione si riduce a una disequazione di secondo grado e, successivamente, al sistema

$$\arcsin x < 1 \quad \wedge \quad \arcsin x > 2.$$

La prima è verificata per $-1 \leq x < \sin 1 \simeq 0.84$, che sono anche le soluzioni del sistema, visto che la seconda non ha soluzioni ($2 > \pi/2$).

11.13. Esercizi

N.B. In tutto questo paragrafo, se non diversamente specificato, $k, h \in \mathbb{Z}$. Utilizzeremo inoltre, salvo diverse specificazioni, le convenzioni di nomenclatura sui triangoli indicate nella figura 11.40 nella pagina 390.

Esercizio 11.1. *Per quali valori di $a \in \mathbb{R}$ la quantità*

$$\frac{a-1}{a+2}$$

può essere il seno di un angolo?

Risoluzione. È sufficiente che

$$-1 \leq \frac{a-1}{a+2} \leq 1 \quad \Rightarrow \quad \frac{2a+1}{a+2} \geq 0 \wedge \frac{2}{a+2} \geq 0 \quad \Rightarrow \quad a \geq -\frac{1}{2}. \qquad \square$$

Esercizio 11.2. *Dimostrare che*

$$\sin^4 x = \frac{1}{8}\cos 4x - \frac{1}{2}\cos 2x + \frac{3}{8}.$$

Risoluzione. Si ha

$$\sin^4 x = \left(\sin^2 x\right)^2 = \left(\frac{1-\cos 2x}{2}\right)^2 = \frac{1-2\cos 2x + \cos^2 2x}{4} =$$
$$= \frac{1}{4}\left(1 - 2\cos 2x + \frac{1+\cos 4x}{2}\right) = \frac{1}{8}\cos 4x - \frac{1}{2}\cos 2x + \frac{3}{8}.$$

Una formula di questo tipo può essere utile per trasformare una potenza in una somma. □

Esercizio 11.3. *Dimostrare che*

$$\frac{\sin x + \sin 3x}{\cos x + \cos 3x} = \tan 2x,$$

determinando anche le condizioni per la validità dell'uguaglianza.

Risoluzione. Usando le formule di triplicazione[(11)] si trova

$$\frac{\sin x + \sin 3x}{\cos x + \cos 3x} = \frac{\sin x + 3\sin x - 4\sin^3 x}{\cos x + 4\cos^3 x - 3\cos x} = \frac{4\sin x(1-\sin^2 x)}{2\cos x(2\cos^2 x - 1)} = \frac{\sin 2x}{\cos 2x} = \tan 2x.$$

Per quanto riguarda le condizione di validità deve essere

$$\cos x + \cos 3x = 2\cos x \cos 2x \neq 0 \quad \Rightarrow \quad x \neq \frac{\pi}{2} + k\pi \quad \wedge \quad x \neq \frac{\pi}{4} + k\frac{\pi}{2}.$$

Si noti che per il secondo membro della formula data sarebbe stata sufficiente la seconda delle condizioni; la prima derivata dal fatto che, nel corso della dimostrazione, si è semplificato per $\cos x$, semplificazione che richiede, appunto, che $\cos x \neq 0$. □

[11]In alternativa si potevano usare le formule di prostaferesi.

Esercizio 11.4. *Trovare il massimo, il minimo e le ascisse dei punti di massimo e minimo di* $f(x) = 3\cos x + 4\sin x$ *in* $[0, 2\pi]$.

Risoluzione. La funzione data (lineare in seno e coseno) può essere trasformata in una funzione contenente solo il seno o solo il coseno, utilizzando le formule di addizione, procedura standard per risolvere equazioni e disequazioni lineari in seno e coseno. Si può raccogliere a fattor comune $\sqrt{3^2+4^2} = 5$. Si ha

$$3\cos x + 4\sin x = 5\left(\frac{3}{5}\cos x + \frac{4}{5}\sin x\right).$$

Detto α l'angolo, compreso tra 0 e 2π, che ha $\sin\alpha = {}^3/_5$ e $\cos\alpha = {}^4/_5$, la funzione f può essere riscritta come

$$f(x) = 5\sin(\alpha + x).$$

Qualunque sia α, $\sin(\alpha + x)$ ha massimo 1 e minimo -1. Dunque la funzione f avrà massimo 5 e minimo -5. Il massimo e il minimo vengono raggiunti, rispettivamente, quando $\alpha + x = {}^\pi/_2 + 2k\pi$ e $\alpha + x = -{}^\pi/_2 + 2k\pi$. Se teniamo conto che, avendo seno e coseno entrambi positivi, α sta nel primo quadrante, concludiamo che i valori di $x \in [0, 2\pi]$ che rendono massima e minima la funzione sono, rispettivamente,

$$x_1 = \frac{\pi}{2} - \alpha \simeq 0.93 \quad \text{e} \quad x_2 = \frac{3\pi}{2} - \alpha \simeq 4.07.$$

Nella figura 11.56 abbiamo tracciato, per un utile confronto, il grafico della funzione f.

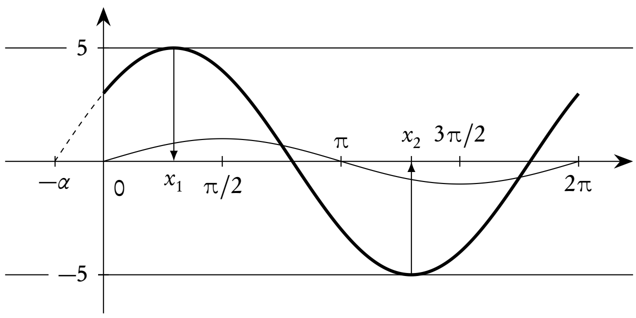


Figura 11.56.: *La funzione* $f(x) = 3\cos x + 4\sin x$ *in confronto con la funzione* $\sin x$ *(Sistema non monometrico!)*

La figura 11.56 rende evidente che le funzioni lineari in seno e coseno sono ottenute per traslazione e dilatazione a partire dalle funzioni trigonometriche elementari. □

Esercizio 11.5. *Per quali* $x \in [0, 2\pi[$ *vale l'uguaglianza*

$$\sqrt{\frac{1-\cos x}{2}}\sqrt{\frac{1+\cos x}{2}} = \sin\frac{x}{2}\cos\frac{x}{2}?$$

Risoluzione. Se $x \in [0, 2\pi[$, allora ${}^x/_2 \in [0, \pi[$ e quindi, per le formule di bisezione,

$$\sin\frac{x}{2} \geq 0 \quad \Rightarrow \quad \sin\frac{x}{2} = \sqrt{\frac{1-\cos x}{2}}.$$

Se $x \in [0, \pi]$, allora $x/2 \in [0, \pi/2]$, se $x \in]\pi, 2\pi[$, allora $x/2 \in]\pi/2, \pi[$, dunque

$$\cos\frac{x}{2} = \sqrt{\frac{1+\cos x}{2}} \quad \text{se} \quad x \in [0, \pi].$$

L'uguaglianza proposta vale dunque in $[0, \pi]$. □

Esercizio 11.6. *Risolvere l'equazione*

$$\cos^2(x+1) + \sin^2(x-1) = 1.$$

Risoluzione. L'equazione si può riscrivere come $\sin^2(x-1) = 1 - \cos^2(x+1)$, ovvero $\sin^2(x-1) = \sin^2(x+1)$, o ancora $\sin(x-1) = \pm\sin(x+1)$. Si tratta allora di vedere quando due angoli hanno seno uguale od opposto. Ci sono quattro possibilità:

- $(x-1) = (x+1) + 2k\pi$;
- $(x-1) = \pi - (x+1) + 2k\pi$;
- $(x-1) = -(x+1) + 2k\pi$;
- $(x-1) = \pi + (x+1) + 2k\pi$.

La prima e la quarta non sono mai verificate, la seconda e la terza comportano, rispettivamente

$$x = \frac{\pi}{2} + k\pi, \quad x = k\pi.$$

Si può scrivere, brevemente,

$$x = k\frac{\pi}{2}.$$ □

Esercizio 11.7. *Calcolare*

$$\sin\left(\frac{7\pi}{12}\right)\sin\left(\frac{\pi}{12}\right).$$

Risoluzione. Usando le formule di Werner si ha

$$\sin\left(\frac{7\pi}{12}\right)\sin\left(\frac{\pi}{12}\right) = \frac{1}{2}\left(\cos\left(\frac{6\pi}{12}\right) - \cos\left(\frac{8\pi}{12}\right)\right) = \frac{1}{4}.$$

□

Esercizio 11.8. *Dimostrare che, per ogni $x \in \mathbb{R}$, la quantità*

$$a = \sqrt{(\sin x + 3) + 2\sqrt{2}\sqrt{1+\sin x}} + \sqrt{(\sin x + 3) - 2\sqrt{2}\sqrt{1+\sin x}}$$

è positiva e costante e calcolarne il valore.

Risoluzione. Poiché la quantità a deve essere costante su tutto $\mathbb{R}$, cominciamo a controllare se il dominio è proprio $\mathbb{R}$. Intanto $1 + \sin x \geq 0$ per ogni $x \in \mathbb{R}$, dunque $\sqrt{1+\sin x}$ è sempre definita. Di conseguenza anche il primo radicale sarà sempre definito. Per il secondo dovrà essere

$$\sin x + 3 \geq 2\sqrt{2}\sqrt{1+\sin x}.$$

Trattandosi di disuguaglianza tra numeri positivi possiamo elevare al quadrato ambo i membri. Dopo semplificazione si ottiene

$$\sin^2 x - 2\sin x + 1 \geq 0 \quad \text{ovvero} \quad (\sin x - 1)^2 \geq 0,$$

che è sempre verificata.

Possiamo dunque concludere che $a \geq 0$, per ogni $x \in \mathbb{R}$. Osservato che a si presenta nella forma

$$\sqrt{\alpha+\beta} + \sqrt{\alpha-\beta},$$

e tenendo conto che, essendo $a \geq 0$, $a = \sqrt{a^2}$, proviamo ad elevare al quadrato la quantità data. Dobbiamo tenere conto, nel calcolo del doppio prodotto conseguente all'elevazione al quadrato, che in questo caso il prodotto dei due radicali è uguale alla radice del prodotto perché i radicandi sono positivi. Si ottiene, dopo semplificazione,

$$a^2 = 2\sin x + 6 + 2\sqrt{(\sin x - 1)^2} = 2\sin x + 6 + 2|\sin x - 1| = 2\sin x + 6 + 2(1 - \sin x) = 8,$$

dove abbiamo anche tenuto conto del fatto che $\sin x - 1 \leq 0$ Dunque $a = 2\sqrt{2}$. □

Esercizio 11.9. *Verificare che, per ogni intero positivo $n \geq 2$, esistono due angoli* acuti *α e β tali che*

$$\sin\alpha = \sqrt{\frac{n-\sqrt{n+1}}{2n}}, \quad \sin\beta = \sqrt{\frac{n+\sqrt{n+1}}{2n}},$$

e determinare che relazione esiste tra α e β.

Risoluzione. Bisognerà intanto verificare che

$$n \geq \sqrt{n+1}.$$

Trattandosi di disuguaglianza tra numeri positivi eleviamo al quadrato e semplifichiamo.

$$n^2 \geq n+1 \quad \Leftrightarrow n^2 - n \geq 1 \quad \Leftrightarrow n(n-1) \geq 1,$$

che è vera per l'ipotesi $n \geq 2$. Si deve poi verificare che le due quantità date verifichino la condizione

$$\sqrt{\frac{n-\sqrt{n+1}}{2n}} \leq 1, \quad \sqrt{\frac{n+\sqrt{n+1}}{2n}} \leq 1.$$

Trattandosi di disuguaglianze tra numeri positivi eleviamo al quadrato. Dopo semplificazione si ottiene, rispettivamente,

$$-n - \sqrt{n+1} \leq 0, \quad \text{e} \quad \sqrt{n+1} \leq n \ :$$

la prima è banalmente vera, la seconda l'abbiamo appena provata.

Per trovare la relazione tra α e β calcoliamo $\cos\alpha$ e $\cos\beta$. Si ottiene

$$\cos\alpha = \sqrt{\frac{n+\sqrt{n+1}}{2n}} = \sin\beta, \quad \cos\beta = \sqrt{\frac{n-\sqrt{n+1}}{2n}} = \sin\alpha.$$

Gli angoli α e β sono dunque complementari. □

Esercizio 11.10. *Risolvere il triangolo rettangolo* ABC *sapendo che l'ipotenusa* $\overline{\mathsf{BC}}$ *è* $2\sqrt{2}$ *e che la differenza delle tangenti dei due angoli acuti è* $2\sqrt{3}$.

Risoluzione. Essendo gli angoli γ e β complementari, il prodotto delle loro tangenti è 1. Dunque

$$\begin{cases} \tan\gamma - \tan\beta = 2\sqrt{3} \\ \tan\gamma\tan\beta = 1 \end{cases}.$$

Si tratta di un sistema di due equazioni in due incognite. Si trova, eliminando la soluzione negativa, in quanto si tratta di angoli acuti,

$$\tan\gamma = 2+\sqrt{3}, \quad \tan\beta = 2-\sqrt{3}.$$

A questo punto si può trovare, per esempio, $\sin\gamma$ e, successivamente $c = a\sin\gamma$. Il restante cateto si può trovare anche con il teorema di Pitagora. Si trova

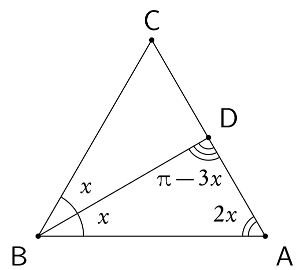
$$b = \sqrt{3}-1, \quad c = \sqrt{3}+1.$$ □

Esercizio 11.11. *Determinare gli angoli alla base di un triangolo isoscele di vertice* A, *sapendo che la base e la bisettrice di uno degli angoli della base misurano rispettivamente* 2 *e* $\sqrt{3}$.

Risoluzione. Con riferimento alla figura a lato, sia $\overline{\mathsf{BD}}$ la bisettrice dell'angolo in B e indichiamo con x l'angolo $\beta/2$. Si ha allora $\gamma = 2x$ e $\mathsf{B}\widehat{\mathsf{D}}\mathsf{C} = \pi - 3x$. Si può applicare il teorema dei seni al triangolo BDC. Tenendo conto che $\sin(\pi-3x) = \sin 3x$, si ottiene

$$\frac{2}{\sin 3x} = \frac{\sqrt{3}}{\sin 2x}$$

Dopo semplificazione si ottiene un'equazione in $\cos x$ che ha come unica soluzione accettabile $x = \pi/3$. □



Esercizio 11.12. *È dato un angolo retto* XOY *e sono dati due punti* A *e* B *sui lati* OX *e* OY *rispettivamente, in modo che* $|\overline{\mathsf{OA}}| = \sqrt{3}|\overline{\mathsf{OB}}|$. *Determinare un punto* P *interno all'angolo retto in modo che l'angolo* $\mathsf{O}\widehat{\mathsf{P}}\mathsf{A}$ *sia retto e che*

$$|\overline{\mathsf{OP}}|^2 + |\overline{\mathsf{PB}}|^2 = (4-2\sqrt{3})|\overline{\mathsf{OB}}|^2.$$

Risoluzione. Cominciamo con l'osservare che l'angolo $\mathsf{O}\widehat{\mathsf{P}}\mathsf{A}$ può essere retto se e solo se P sta sulla semicirconferenza di diametro $\overline{\mathsf{OA}}$ interna all'angolo retto dato. Dopodiché, posto $|\overline{\mathsf{OB}}| = a$, si ha $|\overline{\mathsf{OA}}| = a\sqrt{3}$. Scelto $x = \mathsf{A}\widehat{\mathsf{O}}\mathsf{P}$, si può trovare $|\overline{\mathsf{OP}}|$ come cateto del triangolo rettangolo AOP e, successivamente, $|\overline{\mathsf{PB}}|^2$ usando il teorema di Carnot nel triangolo POB, di cui l'angolo $\mathsf{P}\widehat{\mathsf{O}}\mathsf{B}$ è $\pi/2 - x$. Si ottiene, dopo semplificazione, l'equazione

$$6\cos^2 x - 2\sqrt{3}\cos x\sin x - 3 + 2\sqrt{3} = 0,$$

che può essere trasformata nella

$$\sqrt{3}\cos 2x - \sin 2x + 2 = 0.$$

Si trova l'unica soluzione $x = 5\pi/12$. Si veda la figura 11.57. □

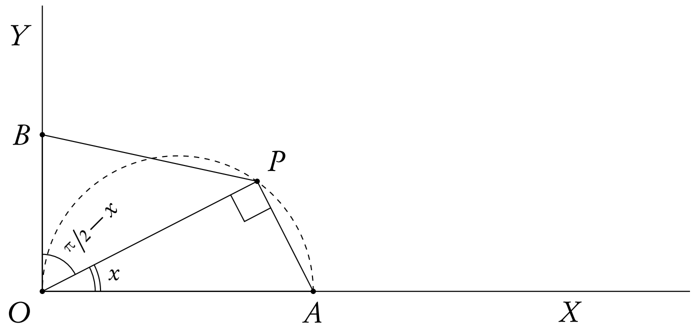


Figura 11.57.: *Determinazione di un punto* P *tale che* $|\overline{OP}|^2 + |\overline{PB}|^2 = (4-2\sqrt{3})|\overline{OB}|^2$

Esercizio 11.13. *Risolvere la disequazione*

$$\cos x + 1 < \cos\frac{x}{2}\,.$$

Risoluzione. Conviene preventivamente scrivere le funzioni trigonometriche in modo che abbiano come argomento lo stesso valore: in questo caso di deve scegliere tra x e $x/2$. È opportuno trasformare in modo da avere tutto in funzione di $x/2$, in quanto questo evita l'uso di radicali. Si può usare la formula

$$\cos\frac{x}{2} = 2\cos^2\frac{x}{2} - 1.$$

Dopo semplificazione si ottiene

$$2\cos^2\frac{x}{2} - \cos\frac{x}{2} < 0.$$

Posto $t = \cos x/2$, si ottiene una disequazione in t di immediata soluzione: $0 < t < 1/2$. Dunque

$$-\frac{\pi}{2} + 2k\pi < \frac{x}{2} < -\frac{\pi}{3} + 2k\pi \quad \vee \quad \frac{\pi}{3} + 2k\pi < \frac{x}{2} < \frac{\pi}{2} + 2k\pi.$$

Moltiplicando per 2 tutti i membri si ottiene infine

$$-\pi + 4k\pi < x < -\frac{2}{3}\pi + 4k\pi \quad \vee \quad \frac{2}{3}\pi + 4k\pi < x < \pi + 4k\pi.$$ □

Esercizio 11.14. *Risolvere la disequazione*

$$\cos(x + |x|) > 0.$$

Risoluzione. Se $x < 0$ si ha $|x| = -x$, e dunque la disequazione si riduce a $1 > 0$ che è sicuramente vera. Se $x \geq 0$ si ottiene invece $\cos 2x > 0$ che è verificata per

$$0 \leq 2x < \frac{\pi}{2} \quad \vee \quad -\frac{\pi}{2} + 2k\pi < 2x < \frac{\pi}{2} + 2k\pi, \quad k \in \mathbb{N} \setminus \{0\}.$$

Se ne deduca che le soluzioni sono

$$x < \frac{\pi}{4} \quad \vee \quad -\frac{\pi}{4} + k\pi < x < \frac{\pi}{4} + k\pi, \quad k \in \mathbb{N} \setminus \{0\}.$$ □

Esercizio 11.15. *Risolvere la disequazione*

$$\sqrt{5-2\sin x} \geq 6\sin x - 1.$$

Risoluzione. Posto $\sin x = t$ si ottiene la disequazione in t

$$\sqrt{5-2t} \geq 6t - 1,$$

che richiede $t \leq 5/2$ per il dominio e, successivamente,

$$t < \frac{1}{6} \quad \vee \quad \begin{cases} t \geq \dfrac{1}{6} \\ (\sqrt{5-2t})^2 \geq (6t-1)^2 \end{cases}.$$

Si trova

$$t \leq \frac{1}{2} \quad \Rightarrow \quad -\frac{7}{6}\pi + 2k\pi \leq x \leq \frac{\pi}{6} + 2k\pi.$$

□

Esercizio 11.16. *Risolvere la disequazione*

$$|\sin x - \cos x| < 1.$$

Risoluzione. La disequazione è equivalente a

$$-1 < \sin x - \cos x < 1 \quad \Rightarrow \quad \begin{cases} \sin x - \cos x < 1 \\ \sin x - \cos x > -1 \end{cases}.$$

Nella figura 11.58 sono visualizzate, sulla circonferenza goniometrica, le soluzioni delle due disequazioni e quelle del sistema. Questo tipo di visualizzazione sulla circonferenza è particolarmente significativa per trattare sistemi di disequazioni trigonometriche, quando il minimo periodo comune è 2π o suoi sottomultipli.

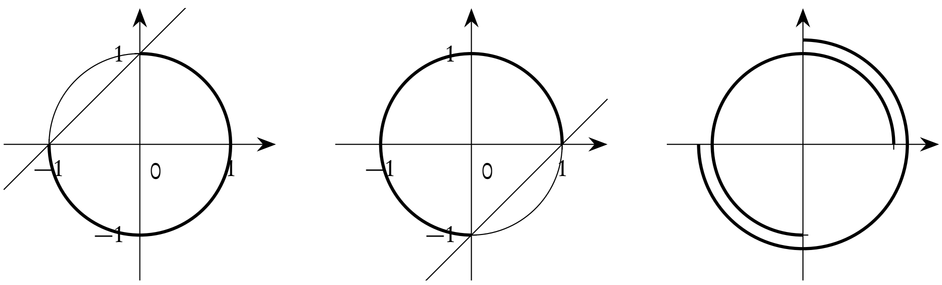


Figura 11.58.: *La disequazione* $|\sin x - \cos x| < 1$

Se ne deduce che le soluzioni sono

$$k\pi < x < \frac{\pi}{2} + k\pi.$$

Si noti che il minimo periodo è π.

□

Esercizio 11.17. *Risolvere la disequazione*

$$\frac{3\sin x - \sqrt{3}\cos x}{2\cos x - 1} \le 0.$$

Risoluzione. Si deve trovare il segno del numeratore e del denominatore, risolvendo le disequazioni

$$3\sin x - \sqrt{3}\cos x > 0 \quad \text{e} \quad 2\cos x - 1 > 0.$$

Si possono riportare i risultati su un grafico di segno, limitando l'indagine a un intervallo ampio 2π.

$+/-$	0		$\frac{\pi}{6}$		$\frac{\pi}{3}$		$\frac{7\pi}{6}$		$\frac{5\pi}{3}$		2π
$3\sin x - \sqrt{3}\cos x$	$-$	$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	$-$
$2\cos x - 1$	$+$	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$	$+$
Complessivo	$-$	$-$	0	$+$	$\times$	$-$	0	$+$	$\times$	$-$	$-$

Si ha dunque

$$-\frac{\pi}{3} + 2k\pi < x \le \frac{\pi}{6} + 2k\pi \quad \vee \quad \frac{\pi}{3} + 2k\pi < x \le \frac{7\pi}{6} + 2k\pi.$$

A volte, particolarmente quando la periodicità è 2π o suoi sottomultipli, può essere conveniente fare un grafico di segno direttamente usando la circonferenza goniometrica, come si può vedere dalla figura 11.59.

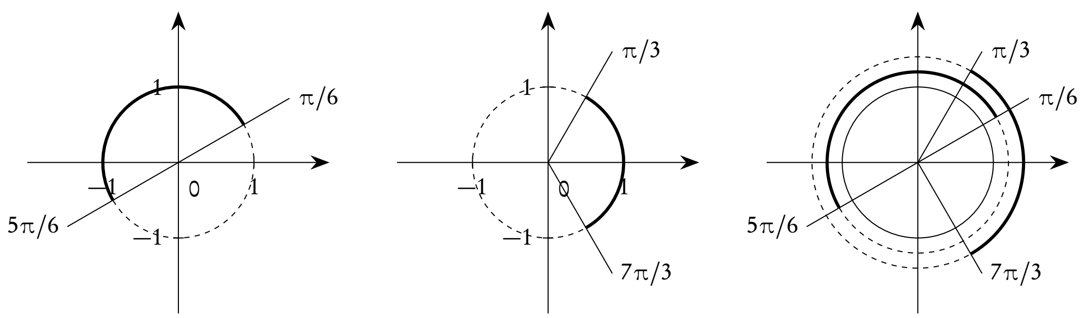


Figura 11.59.: *La disequazione* $(3\sin x - \sqrt{3}\cos x)/(2\cos x - 1) \le 0$

Sono stati evidenziati con linea continua o tratteggiata gli archi ove il numeratore e denominatore sono, rispettivamente, positivi o negativi: l'uso dei consueti segni "+" e "—" non è conveniente in una rappresentazione come questa. □

Esercizio 11.18. *Risolvere la disequazione*

$$\left(2\sin\frac{x}{2}-1\right)\left(2\cos x-\sqrt{3}\right)\leq 0.$$

Risoluzione. Si deve trovare il segno dei due fattori, risolvendo le disequazioni

$$2\sin\frac{x}{2}-1>0 \quad \text{e} \quad 2\cos x-\sqrt{3}>0.$$

Si tratta in entrambi i casi di disequazioni elementari che hanno per soluzioni, rispettivamente,

$$\frac{\pi}{6}+2k\pi<\frac{x}{2}<\frac{5\pi}{6}+2k\pi \quad \text{ovvero} \quad \frac{\pi}{3}+4k\pi<x<\frac{5\pi}{3}+4k\pi$$

e

$$-\frac{\pi}{6}+2k\pi<x<\frac{\pi}{6}+2k\pi.$$

Costruiamo un grafico di segno, limitando l'indagine all'intervallo $[0,4\pi]$ e osservando che, a differenza dell'esercizio 11.17, non conviene usare la circonferenza goniometrica, dove è difficile valutare periodi maggiori di 2π.

+/−	0		$\frac{\pi}{6}$		$\frac{\pi}{3}$		$\frac{5\pi}{3}$		$\frac{11\pi}{6}$		$\frac{13\pi}{6}$		$\frac{23\pi}{6}$		4π
$2\sin\frac{x}{2}-1$	−	−	−	−	0	+	0	−	−	−	−	−	−	−	−
$2\cos x-\sqrt{3}$	+	+	0	−	−	−	−	−	0	+	0	−	0	+	+
Complessivo	−	−	0	+	0	−	0	+	0	−	0	+	0	−	−

Le soluzioni sono dunque

$$-\frac{\pi}{6}+4k\pi\leq x\leq\frac{\pi}{6}+4k\pi \vee \frac{\pi}{3}+4k\pi\leq x\leq\frac{5\pi}{3}+4k\pi \vee \frac{11\pi}{6}+4k\pi\leq x\leq\frac{13\pi}{6}+4k\pi,$$

dove abbiamo raggruppato l'ultimo intervallo del grafico con il primo, per semplicità. □

12. Introduzione al calcolo combinatorio

In questo capitolo è proposta una introduzione ai fondamenti del calcolo combinatorio, limitatamente ai suoi aspetti più elementari.

In tutto questo capitolo, salvo quando diversamente specificato, i numeri m, n, k sono sempre naturali.

12.1. Introduzione

Il calcolo combinatorio, premessa indispensabile per lo studio del calcolo delle probabilità, si occupa della determinazione della cardinalità di certi insiemi finiti, assegnati mediante proprietà caratteristica. Il problema è usualmente molto complesso e, soprattutto, non esistono metodi standard per risolverlo: forse per questo è ritenuto argomento ostico e difficile.

In questo capitolo descriveremo le tecniche fondamentali e proporremo alcuni esempi risolti: l'esame di un buon numero di casi concreti è l'unico modo per acquisire dimestichezza nella risoluzione di questo tipo di problemi.

E' opportuno segnalare fin da subito che esistono situazioni in cui l'unica strategia che è possibile utilizzare è quella di scrivere l'insieme in questione e numerare, uno alla volta, i suoi elementi. L'esempio più classico è costituito dal seguente problema: *Dato un numero naturale n, determinare la cardinalità dell'insieme dei numeri primi minori di n.* Non resta altro da fare che scrivere pazientemente tutti i numeri richiesti e contarli.

Come già segnalato la risoluzione di problemi di Analisi Combinatoria richiede modi di ragionamento e tecniche usualmente poco familiari e che non si lasciano classificare in schemi standard. Ragionamenti intuitivi portano facilmente a risultati errati ed è opportuno ricondurre i problemi ad alcuni modelli astratti che di solito facilitano la ricerca della tecnica corretta. Utilizzeremo di norma i modelli, classici, dell'estrazione di oggetti da un'urna e della collocazione di oggetti in celle.

— *Modello dell'urna.* Consiste nell'estrarre da un'urna di composizione nota un certo numero di oggetti. In alcuni casi è prevista la reintroduzione nell'urna dell'oggetto estratto (modello con reimbussolamento), in altri casi no.
— *Modello a celle.* Si tratta di collocare un determinato numero di oggetti in un dato numero di celle, essendo a volte consentito che più oggetti stiano in una cella, altre volte no.

I due approcci sono assolutamente equivalenti, ma è utile familiarizzare con entrambi: a seconda del tipo di problema può essere più semplice schematizzarlo con uno invece che con l'altro.

Prima di entrare nei dettagli sono utili due definizioni.

Definizione 12.1. *Dato un naturale $n > 0$, si chiama* fattoriale *di n e si indica con $n!$ (n fattoriale) il prodotto di tutti i naturali da 1 a n:*

$$n! = 1 \cdot 2 \cdot 3 \dots n. \tag{12.1}$$

Si pone poi $0! = 1$.

È importante in molti esercizi, anche se praticamente ovvio, che $n! = n(n-1)! = n(n-1)(n-2)!$ e così via.

Definizione 12.2. *Dato un naturale $n > 0$, si chiama* semifattoriale *di n e si indica con $n!!$ (n semifattoriale) il prodotto di tutti i naturali che hanno la stessa parità di n, sempre da 1 a n:*

$$n \text{ pari } \Rightarrow n!! = 2 \cdot 2 \cdot 4 \dots n; \quad n \text{ dispari } \Rightarrow n!! = 1 \cdot 3 \cdot 5 \dots n. \tag{12.2}$$

Si ha $n! = n!! \cdot (n-1)!!$.

È utile anche richiamare il seguente risultato, di facile dimostrazione, sulla cardinalità del prodotto cartesiano di due o più insiemi finiti: la cardinalità del prodotto degli insiemi finiti $A_1, A_2, \dots, A_n$ è uguale al prodotto delle cardinalità dei singoli insiemi.

$$|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdots |A_n|. \tag{12.3}$$

La versione "combinatoria" di questo fatto può essere espressa con il seguente enunciato.

Teorema 12.3 (Principio fondamentale del calcolo combinatorio)**.** *Supponiamo che un problema P possa essere scomposto in k sottoproblemi indipendenti $P_1, P_2, \dots, P_k$, ciascuno dei quali può essere risolto in n_1, $n_2, \dots, n_k$ modi. Allora il problema P può essere risolto in*

$$n = n_1 \cdot n_2 \cdots n_k$$

modi diversi.

12.2. Disposizioni

12.2.1. Allineamenti ordinati

Il primo problema base del calcolo combinatorio può essere espresso come segue: dato un insieme A di n elementi in quanti modi è possibile costruire allineamenti ordinati di k di questi elementi?

Esempio 12.1. Dati i simboli 1, 2, X, quante colonne di 13 simboli siffatti si possono costruire? (Gioco del totocalcio).

Esempio 12.2. Dato un alfabeto di 26 simboli, quante sono le parole di tre simboli tutti distinti che si possono costruire (prescindendo dal significato!)?

Esempio 12.3. In una gara di Formula 1 con 20 piloti, quante sono le possibili terne candidate a salire sul podio?

Si noti che nel primo esempio gli allineamenti richiesti comportano la possibilità di ripetizioni, negli altri due no.

Utilizzando lo schema a celle il problema può essere visualizzato così: dati n simboli e k caselle numerate progressivamente, in quanti modi è possibile riempire le k caselle con gli n simboli? È chiaro che, se non sono consentite ripetizioni dei simboli, k non deve superare n, altrimenti non ci sono condizioni. Si veda la figura 12.1.

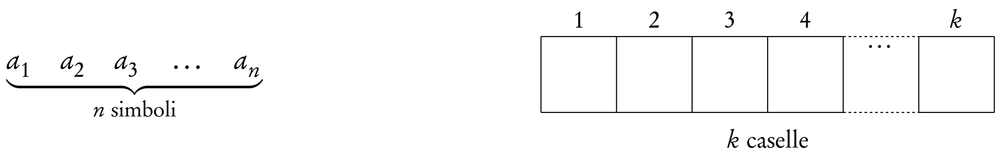


Figura 12.1.: *Disposizioni di "n" oggetti distinti in "k" caselle numerate*

Nel caso siano consentite ripetizioni è più utile pensare a un insieme di n "tipi di oggetti", anziché a un insieme di n oggetti: se ho tanti oggetti dello stesso tipo non ci sono difficoltà a prenderne più d'uno per costruire l'allineamento voluto.

Utilizzando lo schema dell'urna il problema può invece essere visualizzato così: data un'urna con n palline numerate da 1 ad n, in quanti modi è possibile estrarre k palline a caso, tenendo conto dell'ordine di apparizione? L'estrazione senza reimbussolamento equivale a considerare allineamenti senza possibilità di ripetizione, quella con reimbussolamento consente invece le ripetizioni.

Se non sono previste ripetizioni il problema può essere anche riformulato in uno dei modi seguenti.

— Dato un insieme di n elementi quanti sono i suoi sottoinsiemi ordinati di k ($\leq n$) elementi?
— Dati due insiemi, C di k elementi (nel caso in esame insieme dei contrassegni o delle caselle) e A di n elementi (nel caso in esame insieme degli oggetti da disporre), quante sono le funzioni *iniettive* di C in A?

Se sono previste ripetizioni il problema può essere anche riformulato nel seguente modo.

— Dati due insiemi, C di k elementi e A di n elementi, quante sono le funzioni di C in A?

L'interpretazione in termini di funzioni è particolarmente importante e si capisce subito se si pensa ad una funzione di questo tipo (cioè tra due insiemi finiti) come una tabella a doppia entrata: su una colonna (o riga) tutti gli elementi di C, sull'altra i corrispondenti elementi di A: se gli elementi di A non si possono ripetere si tratta di funzioni iniettive, altrimenti di funzioni generiche. Nella figura 12.2 sono proposti alcuni esempi, con $C = \{1,2,3\}$ e $A = \{a,b,c,d\}$; i primi due rappresentano funzioni generiche, gli ultimi due funzioni iniettive.

C	A
1	c
2	c
3	a

C	A
1	b
2	b
3	b

C	A
1	a
2	d
3	b

C	A
1	c
2	d
3	b

Figura 12.2.: *Disposizioni come funzioni tra insiemi finiti*

Diamo ora la definizione formale.

Definizione 12.4 (Disposizioni)**.** *Dati due insiemi, $C = \{1,2,\dots,k\}$ di k elementi*[(1)] *e A di n elementi, una funzione iniettiva di C in A si chiama una* disposizione semplice *di n oggetti a k a k, ovvero di classe*

[1]In realtà l'insieme C potrebbe essere un insieme generico, ma qui ci interessa proprio un insieme di "contrassegni" o caselline numerate.

k: in questo caso $k \leq n$; una funzione (anche non iniettiva) di C in A si chiama invece una disposizione con ripetizione *di n oggetti a k a k, ovvero di classe k: in questo caso non ci sono limiti per k. Una disposizione semplice si può anche identificare come un sottoinsieme ordinato di A costituito da k elementi.*

Il numero delle disposizioni semplici di n oggetti di classe k si indica[2] con V_n^k, quello delle disposizioni con ripetizione con ${}^{\mathsf{R}}\mathsf{V}_n^k$. Valgono le seguenti formule.

$$\mathsf{V}_n^k = \underbrace{n(n-1)(n-2)\dots(n-(k-1))}_{k \text{ fattori}} = \underbrace{n(n-1)(n-2)\dots(n-k+1)}_{k \text{ fattori}} = \frac{n!}{(n-k)!}\,; \tag{12.4}$$

$${}^{\mathsf{R}}\mathsf{V}_n^k = n^k. \tag{12.5}$$

Per la dimostrazione basta solo osservare che se si devono riempire k caselline con n simboli, nella prima casellina si può mettere uno qualunque dei simboli (e dunque ci sono n possibilità di scelta), nella seconda uno dei rimanenti, se non sono consentite ripetizioni (e dunque ci sono $n-1$ possibilità di scelta), ovvero un simbolo qualunque se sono consentite ripetizioni (e dunque con ancora n possibilità di scelta).

Osservazione 12.5. La formula 12.5 dà origine ad un simbolo comunemente usato nello studio delle funzioni tra insiemi, simbolo che abbiamo già incontrato nell'esercizio 1.19 nella pagina 35: l'insieme di tutte le funzioni da un insieme A in un insieme B si indica B^A. Anzi, se α e β sono due numeri cardinali qualunque, rispettivamente riferiti ad un insieme A e ad un insieme B, si definisce la potenza β^α come la cardinalità dell'insieme B^A. Questa definizione nel caso di cardinali finiti riproduce la classica definizione di potenza, definizione che però viene estesa anche ai cardinali transfiniti. Un caso interessante si ha quando B è costituito da due soli elementi: se li chiamiamo *true* e *false*, si vede subito che il numero di queste funzioni è uguale al numero dei sottoinsiemi dell'insieme A. Se infatti ad un elemento di A corrisponde il valore true, esso starà nel corrispondente sottoinsieme, altrimenti no. Per esempio il sottoinsieme vuoto è quello che corrisponde alla funzione che ad ogni elemento di A fa corrispondere false. Per questo motivo si usa indicare con il simbolo 2^A l'insieme delle parti di A, che avrà cardinalità 2^α. Ques'ultima simbologia trova una giustificazione anche nel calcolo delle combinazioni di n oggetti di classe k.

Esempio 12.4. Dato $A = \{a, b, c, d\}$, le disposizioni semplici di classe 2 sono in numero $\mathsf{V}_4^2 = 12$ e sono: ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc. Le disposizioni con ripetizione della stessa classe sono invece ${}^{\mathsf{R}}\mathsf{V}_4^2 = 16$ e sono le 12 precedenti con l'aggiunta di aa, bb, cc, dd.

12.2.2. Suddivisioni

Lo schema a celle può anche essere visualizzato in un modo complementare rispetto a quello visto prima. Per considerare una disposizione di n oggetti di classe k si può immaginare di avere n cassetti (tanti quanti sono gli oggetti) e k contrassegni che devono essere distribuiti nei cassetti: se ogni cassetto ne può contenere al massimo uno si tratta di disposizioni senza ripetizione, se invece ne può contenere anche di più si tratta di disposizioni con ripetizione. I contrassegni corrispondono, nella sostanza, al posto che ciascun oggetto occupa nell'allineamento considerato all'inizio.

[2] Qui, come del resto in tutto questo testo, abbiamo preferito usare, quando presente, la notazione UNI CEI ISO 80000-2 : 2010. In questo caso il simbolo V sta per "variazioni".

Consideriamo l'esempio, già trattato, delle disposizioni di quattro oggetti a due a due. Anziché disporre due dei quattro oggetti in due celle numerate, possiamo pensare di avere quattro cassetti, denominati a, b, c, d e due contrassegni, indicati con 1 e 2, da piazzare nei tre cassetti: la disposizione ab equivarrà a piazzare il contrassegno 1 nel cassetto a e il contrassegno 2 nel cassetto b, lasciando vuoti gli altri, e così via. Il totale delle disposizioni può ora essere schematizzato come mostra la figura 12.3, dove sono riportate sia le disposizioni come allineamenti ordinati di 4 oggetti a 2 a 2, sia come suddivisioni di 2 contrassegni in 4 cassetti.

	ab	ac	ad	ba	bc	bd	ca	cb	cd	da	db	dc	aa	bb	cc	dd
a	1	1	1	2			2			2			1, 2			
b	2			1	1	1		2			2			1, 2		
c		2			2		1	1	1			2			1, 2	
d			2			2			2	1	1	1				1, 2

Figura 12.3.: *Disposizioni di* 4 *oggetti a* 2 *a* 2 *e suddivisioni di* 2 *contrassegni distinti in* 4 *cassetti*

Questo schema è utile per trattare problemi di *suddivisione* del tipo: avendo k oggetti tra di loro distinguibili, determinare in quanti modi è possibile suddividerli in n cassetti, tenendo conto di quali oggetti finiscono nei cassetti, mettendone al più uno per cassetto (disposizioni senza ripetizione) o consentendo un numero arbitrario di oggetti nei vari cassetti (disposizioni con ripetizione).

Riassumendo quanto esposto relativamente alle disposizioni, possiamo concludere che una disposizione di n oggetti di classe k si può realizzare in uno dei modi seguenti.

— Un'estrazione, che tenga conto dell'ordine, di k palline da un'urna contenente n palline distinguibili (con reimbussolamento se si accettano le ripetizioni, senza nel caso contrario).
— Un riempimento di k celle numerate progressivamente con n simboli diversi, con la possibilità che ciascun simbolo sia ripetuto più di una volta (disposizioni con ripetizione) oppure al massimo una volta (disposizioni semplici).
— Una distribuzione di k oggetti distinguibili in n celle (cassetti) diverse, con possibilità che ciascuna cella contenga anche più oggetti (disposizioni con ripetizione) oppure al massimo una pallina (disposizioni semplici).

Esempio 12.5. Una colonna del totocalcio può essere vista

— come una successione di 13 estrazioni, con reimbussolamento, da un'urna contenente i simboli 1, X, 2.
— come il riempimento di 13 celle, numerate progressivamente, con i simboli 1, X, 2.
— come la sistemazione in tre cassetti, denominati 1, X, 2, dei numeri da 1 a 13, essendo consentito che qualcuno dei cassetti possa restare vuoto.

12.3. Permutazioni

Permutare significa "cambiare l'ordine" in cui certi oggetti sono scritti. Interessa sia il caso in cui gli oggetti sono tutti distinti, sia quello in cui ci sono ripetizioni.

12.3.1. Permutazioni fra elementi distinti

Il caso in cui le disposizioni semplici di n oggetti siano di classe n ha una particolare importanza, tanto da meritare un capitolo a sé nello studio del calcolo combinatorio. In questo caso si tratta di considerare le funzioni iniettive di un insieme in sé, oppure i possibili ordinamenti totali di un insieme di n elementi. In considerazione del fatto che gli insiemi che consideriamo sono finiti, ogni funzione iniettiva è anche biiettiva.

Si dà la seguente definizione.

Definizione 12.6. *Si chiama* permutazione *di un insieme A di n elementi una disposizione semplice degli elementi, di classe n.*

Il numero V_n^n si indica con P_n e si ha

$$\mathsf{P}_n = n! \tag{12.6}$$

Gli anagrammi di parole con lettere tutte distinte sono il più comune esempio di permutazioni. Naturalmente in questo problema si prescinde dal significato delle parole. Per esempio gli anagrammi di *cane* sono in numero di $4! = 24$ e sono: *cane*, caen, cean, *cena*, cnae, cnea, aecn, aenc, anec, *ance*, *acne*, acen, eacn, eanc, ecan, ecna, enac, enca, nace, naec, neac, neca, ncae, ncea. Di questi solo quelli in corsivo hanno un significato nel vocabolario italiano.

12.3.2. Permutazioni fra elementi non tutti distinti

Se riprendiamo in considerazione gli anagrammi, sappiamo che hanno interesse anche quelli di parole formate da lettere non tutte distinte (come "mamma"). E' chiaro che in questo caso il numero di anagrammi è notevolmente inferiore rispetto a quello di parole con lo stesso numero di lettere tutte distinte: per esempio nel citato caso di "mamma", uno scambio tra le tre m o tra le due a non provoca cambiamenti nella parola. In questa e altre situazioni ha dunque interesse anche il calcolo del numero delle permutazioni, o allineamenti, su un insieme di oggetti non tutti distinti.

Consideriamo allora un allineamento di n oggetti, di cui n_1 uguali ad un oggetto a_1, n_2 uguali ad un oggetto $a_2, \dots, n_k$ uguali ad un oggetto a_k, con $n = n_1 + n_2 + \cdots + n_k$, e non essendo escluso che qualcuno degli n_i (o magari anche tutti) sia 1. La determinazione del possibile numero di allineamenti è immediata se si suppone inizialmente che gli oggetti siano tutti distinti, si calcola il numero delle loro possibili permutazioni (che sono in numero di $n!$) e si tiene conto che le $n_i!$ permutazioni degli oggetti uguali ad a_i non danno luogo a situazioni distinte. Questo numero si indica con il simbolo

$$\mathsf{P}_n^{n_1,n_2,\dots,n_k}$$

e si ha

$$\mathsf{P}_n^{n_1,n_2,\dots,n_k} = \frac{n!}{n_1!n_2!\cdots n_k!}. \tag{12.7}$$

Questa formula comprende la (12.6) come caso particolare: la (12.6) si ottiene dalla (12.7) ponendo tutti gli n_i uguali a 1.

Poiché questi numeri sono legati allo sviluppo della potenza n-esima di un polinomio, si chiamano anche *coefficienti polinomiali* o *coefficienti multinomiali*.

Per esempio gli anagrammi della parola *mamma* sono *mamma*, mmmaa, mmaam, mmama, mamam, ammma, amamm, aammm, ammam, amamm. Di questi solo l'originale ha un significato nel vocabolario italiano.

Nello schema dell'urna il problema in considerazione equivale a considerare un'urna con n palline, di cui n_1 del colore 1, n_2 del colore 2, ..., n_k del colore k. Si suppone che i colori siano distinguibili, ma le palline dello stesso colore no. Si tratta allora di considerare il numero di possibili estrazioni di tutte queste palline, senza reimbussolamento.

12.3.3. Un problema di ripartizione

Se riprendiamo in esame gli anagrammi della parola "mamma" possiamo anche schematizzarli da un punto di vista complementare rispetto a quello già considerato. Possiamo infatti pensare di avere a disposizione due cassetti, denominati m ed a, e cinque contrassegni, indicati con 1, 2, 3, 4 e 5, da piazzare nei due cassetti: la permutazione *mamma* equivarrà a piazzare i contrassegni 1, 3 e 4 nel cassetto m, i contrassegni 2 e 5 nel cassetto a, e così via per le altre permnutazioni. In sostanza dovremo piazzare tre contrassegni nel cassetto m e due nel cassetto a, senza tenere conto dell'ordine in cui vengono piazzati.

In generale il problema potrà essere così formulato: dati n oggetti distinti e k cassetti numerati, il numero di modi in cui è possibile ripartire gli oggetti nei cassetti in modo che ce ne vadano n_1 nel primo, n_2 nel secondo, ..., n_k nell'ultimo, ovviamente in modo che $n = n_1 + n_2 + \cdots + n_k$, coincide con le permutazioni fra n elementi di cui n_1 uguali ad un oggetto a_1, n_2 uguali ad un oggetto a_2, ..., n_k uguali ad un oggetto a_k, con $n = n_1 + n_2 + \cdots + n_k$, e non essendo escluso che qualcuno degli n_i (o magari anche tutti) sia 1.

Esempio 12.6. Dovendo suddividere 15 studenti in tre classi in modo che ce ne vadano 5 nella prima, 4 nella seconda e 6 nella terza, avremo

$$\mathsf{P}_{15}^{5,4,6} = \frac{15!}{5! \times 4! \times 6!} = 630630$$

possibilità.

Riassumendo quanto esposto relativamente alle permutazioni, possiamo concludere che una permutazione di n oggetti, di cui n_1 del tipo 1, n_2 del tipo 2, ..., n_k del tipo k, con $n = n_1 + n_2 + \cdots + n_k$, e con la possibilità che uno o più degli n_i sia 1, si può vedere in uno dei modo seguenti.

— Un anagramma di una parola di n lettere, di cui n_1 uguali ad una stessa lettera, n_2 uguali ad una stessa lettera, e così via.
— Una ripartizione di n oggetti distinti in k cassetti, in modo che n_1 vadano nel primo cassetto, n_2 nel secondo cassetto, e così via.
— Un'estrazione di tutte le n palline contenute in un'urna, nell'ipotesi che ci siano n_1 palline del colore 1, n_2 palline del colore 2, ..., n_k palline del colore k, e che le palline dello stesso colore siano indistinguibili, mentre i colori sono distinguibili.

12.4. Combinazioni

"Combinare" significa prendere un certo numero di oggetti da un insieme, senza tenere conto dell'ordine con cui sono presi. Anche qui ha interesse sia il caso in cui sono consentite ripetizioni, sia quello in cui le ripetizioni non sono consentite.

12.4.1. Combinazioni semplici

Il terzo problema base del calcolo combinatorio è quello del computo del numero di sottoinsiemi di un dato insieme finito. È immediato che un sottoinsieme di k elementi, presi da un insieme di n elementi, può essere pensato come una collezione di k oggetti su un totale di n, in cui non conti l'ordine. Questo distingue il problema qui trattato da quello del calcolo delle disposizioni (semplici). Si dà in proposito la seguente definizione.

Definizione 12.7. *Dato un insieme A di n elementi, un suo sottoinsieme contenente k ($k \le n$) elementi si chiama una* combinazione semplice *degli n elementi di classe k oppure a k a k.*

Il numero delle combinazioni (semplici) di n elementi di classe k si indica con

$$\mathrm{C}_n^k \quad \text{oppure con} \quad \binom{n}{k}.$$

È evidente che, se si considera una qualunque combinazione di n elementi di classe k e si permutano in tutti i modi possibili, cioè in $k!$ modi, i suoi elementi, si ottengono le disposizioni di n elementi di classe k. Questo ci permette di concludere (tenendo presente, se $n = k$, che $0! = 1$) che, per le combinazioni in oggetto, vale la formula 12.8.

$$\mathrm{C}_n^k = \binom{n}{k} = \frac{\mathrm{V}_n^k}{k!} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}. \tag{12.8}$$

I numeri C_n^k si chiamano anche, perché legati allo sviluppo della potenza di un binomio, *coefficienti binomiali*. Se $k = 0$ il numero dei possibili sottoinsiemi è 1 (solo l'insieme vuoto) e perciò si pone

$$\binom{n}{0} = 1, \tag{12.9}$$

in accordo con la formula (12.8).

Nel modello dell'urna questo problema equivale a considerare un'estrazione di k palline da un'urna contenente n palline distinguibili, senza reimbussolamento e senza tener conto dell'ordine: si potrebbero anche estrarre tutte le k palline contemporaneamente.

Esempio 12.7. Dato $A = \{a, b, c, d\}$, le combinazioni semplici di classe 2 sono in numero di $\mathrm{C}_4^2 = 6$ e sono ab, ac, ad, bc, bd, cd.

12.4.2. Combinazioni con ripetizione

Riprendiamo ora in esame il modello dell'urna visto prima e supponiamo di reimbussolare ogni pallina dopo l'estrazione, avendo segnato il suo contrassegno, ma senza tenere conto dell'ordine. Se ripensiamo l'insieme $A = \{a, b, c, d\}$ dell'esempio precedente come insieme di palline in un'urna, un'estrazione di 2 palline con reimbussolamento dà luogo alle seguenti estrazioni: ab, ac, ad, bc, bd, cd, aa, bb, cc, dd.

L'esempio mostra che potremmo anche schematizzare il problema in questo modo, in analogia a quanto già detto per le disposizioni con ripetizione: supponiamo di avere un insieme A contenente n tipi diversi (anziché n oggetti) e di voler costruire, mediante essi, un insieme B di k elementi, prendendo

m_1 elementi del primo tipo, m_2 elementi del secondo,..., m_n elementi dell'ultimo (non escludendo che qualcuno degli m_i sia zero), ovviamente in modo che $m_1 + m_2 + \cdots + m_n = k$. La domanda che ci poniamo è: in quanti modi è possibile costruire B? Per esempio se abbiano oggetti di quattro tipi, che possiamo indicare (i tipi, non gli oggetti!) con a, b, c, d, e vogliamo costruire un insieme B di 2 elementi, le possibilità (indicando i tipi di oggetti e non gli oggetti) sono: ab, ac, ad, bc, bd, cd, aa, bb, cc, dd, esattamente come prima. Questa schematizzazione consente di mantenere l'idea che le combinazioni siano sottoinsiemi di un insieme. Se ciascuno degli m_i non supera 1, ricadiamo nel caso delle combinazioni semplici.

Possiamo dare la seguente definizione.

Definizione 12.8. *Dato un insieme di n (tipi di) oggetti distinti, si chiama* combinazione con ripetizioni *di quegli oggetti ogni sottoinsieme di k elementi, anche non tutti distinti, presi tra gli n (tipi di) oggetti dati. Trattandosi di sottoinsiemi l'ordine in cui gli oggetti sono presi non ha alcuna importanza.*

12.4.3. Un'equazione diofantea

Riprendendo in esame le combinazioni con ripetizione di 4 oggetti di classe 2, potremmo rappresentare ciascuna combinazione anche scrivendo quante volte compare ciascun oggetto. Le possibilità, in corrispondenza con le combinazioni, sono quelle indicate nella figura 12.4, rispettivamente per quelle semplici e quelle con ripetizione.

$\{a,b\}$	$\{a,c\}$	$\{a,d\}$	$\{b,c\}$	$\{b,d\}$	$\{c,d\}$
$\{1,1,0,0\}$	$\{1,0,1,0\}$	$\{1,0,0,1\}$	$\{0,1,1,0\}$	$\{0,1,0,1\}$	$\{0,0,1,1\}$

$\{a,a\}$	$\{b,b\}$	$\{c,c\}$	$\{d,d\}$
$\{2,0,0,0\}$	$\{0,2,0,0\}$	$\{0,0,2,0\}$	$\{0,0,0,2\}$

Figura 12.4.: *Combinazioni di* 4 *oggetti di classe* 2 *e numero di scelte degli oggetti*

Questo ci suggerisce un altro modo di pensare alle combinazioni con ripetizione di n oggetti di classe k: dato il numero naturale k si chiede in quanti modi esso possa essere scritto come somma di n numeri naturali. Se chiamiamo x_1, x_2, ..., x_n, i naturali incogniti, si tratta di risolvere l'equazione $x_1 + x_2 + \cdots + x_n = k$. Un'equazione di questo tipo, in cui i coefficienti delle incognite sono naturali e i valori ammessi sono naturali, si chiama *diofantea*, dal nome del matematico greco Diofanto. Se aggiungiamo la condizione che i numeri siano solo 0 o 1, allora si tratta di combinazioni semplici: naturalmente in questo caso dovremo avere $k \leq n$.

12.4.4. Suddivisioni

Esattamente come fatto per le disposizioni, possiamo schematizzare le combinazioni, in una maniera complementare a quanto già visto, come un problema di *suddivisione* di k oggetti in n cassetti, con la condizione che non conti quale oggetto capita in un cassetto, ma solo il numero di oggetti. Per considerare una combinazione di n oggetti di classe k possiamo immaginare di avere n oggetti e k contrassegni identici da distribuire nei cassetti: se ogni cassetto ne può contenere al massimo 1 si tratta di combinazioni semplici, se invece ne può contenere anche di più si tratta di combinazioni con ripetizione. I contrassegni indicano solo se un oggetto è presente o meno nel cassetto. Si veda la figura 12.5, relativa al caso delle combinazioni di 4 oggetti a 2 a 2, ovvero della suddivisione di 2 contrassegni in 4 cassetti, dove abbiamo segnato i contrassegni tutti con la $\times$, in quanto essi sono indistinguibili. Si confronti anche questa figura con la figura 12.3 nella pagina 421.

	ab	ac	ad	bc	bd	cd	aa	bb	cc	dd
a	$\times$	$\times$	$\times$				$\times\times$	$\times\times$	$\times\times$	$\times\times$
b	$\times$			$\times$	$\times$					
c		$\times$		$\times$		$\times$				
d			$\times$		$\times$	$\times$				

Figura 12.5.: *Combinazioni di* 4 *oggetti a* 2 *a* 2 *e suddivisioni di* 2 *contrassegni identici in* 4 *cassetti*

Si può rivisitare il problema della suddivisione in un altro modo ancora, a riprova del fatto che nel calcolo combinatorio non esistono metodi o procedure standard. Tra l'altro questo ci consentirà un facile calcolo del numero delle combinazioni con ripetizione.

Se indichiamo con il numero "1" gli n cassetti in cui piazzare gli oggetti e con il numero "0" i k oggetti da sistemare, possiamo immaginare di fare la suddivisione scrivendo tanti 1 quanti sono i contenitori, seguiti ciascuno da tanti zero quanti sono gli oggetti che vi piazziamo dentro. Per esempio nella disposizione di 2 oggetti in 4 cassetti scrivendo 101011 intendiamo la combinazione ab, ovvero un oggetto nel primo cassetto, seguito da un oggetto nel secondo cassetto e nessun oggetto negli altri due; scrivendo 111100 intendiamo la combinazione dd, ovvero nessun oggetto nei primi tre cassetti e due oggetti nell'ultimo. In sostanza le suddivisioni degli oggetti nei contenitori si possono vedere come numeri in binario, di lunghezza $n+k$. Poiché questi numeri iniziano sempre con 1, restano $n+k-1$ posti da riempire: dobbiamo scegliere k posti in cui piazzare il numero 0 e $n-1$ posti in cui piazzare il numero 1. Se si tratta di combinazioni semplici c'è la condizione che due zeri non siano mai appaiati, ma questo conto lo sappiamo già fare; se invece si tratta di combinazioni con ripetizione non c'è alcuna limitazione e dunque si tratterà solo di contare i sottoinsiemi di k o di $n-1$ elementi costruiti a partire da un insieme di $n+k-1$ elementi. Il problema si riduce a un problema di combinazioni semplici.

Se indichiamo il numero delle combinazioni con ripetizione di n (tipi di) oggetti, di classe k con

$$ {}^{\mathsf{R}}C_n^k, $$

si conclude che

$$ {}^{\mathsf{R}}C_n^k = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}. \tag{12.10} $$

È immediato, sulla base della formula (12.8), provare che i due coefficienti binomiali che compaiono nella (12.10) sono uguali.

12.5. La potenza di un binomio e di un polinomio

12.5.1. La potenza di un binomio

Abbiamo già discusso, vedi la tabella 3.1 nella pagina 84, del metodo basato sul triangolo di Tartaglia per calcolare i coefficienti dello sviluppo della potenza n-esima di un binomio. Tuttavia la formula che ora proporremo è di uso più semplice, in quanto non richiede di costruire tutte le righe precedenti l'ennesima prima di costruire l'ennesima stessa. Inoltre questa formula ha anche altri usi importanti.

Consideriamo dunque la potenza $(a+b)^n$, dove a e b sono reali e n è naturale. Per calcolare lo sviluppo si deve eseguire il prodotto: $(a+b)^n = (a+b)\cdot(a+b)\cdots(a+b)$ (n volte). È ovvio che il risultato sarà costituito dalla somma di tanti monomi del tipo $a^{n-k}b^k$, con k compreso tra 0 ed n, con opportuni coefficienti, per determinare i quali si tratta di contare quante volte compare ciascuno di questi monomi. Il calcolo è facile se si tiene conto che basterà scegliere $n-k$ volte il fattore a tra gli n fattori del tipo $(a+b)$, e di conseguenza k volte il fattore b nei rimanenti. Il numero cercato[3] è allora (visto che l'ordine non conta!) $C_n^k = C_n^{n-k}$. Si ottiene dunque

$$
\begin{aligned}
(12.11)\quad (a+b)^n &= \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k = \\
&= \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-2}a^2b^{n-2} + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n = \\
&= a^n + na^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-2}a^2b^{n-2} + nab^{n-1} + b^n.
\end{aligned}
$$

Questa formula è anche detta *formula di Newton* per la potenza di un binomio.

Esempio 12.8. Calcolare $(a-b)^5$. Si ha

$$
\begin{aligned}
(a-b)^5 &= \binom{5}{0}a^5 + \binom{5}{1}a^4(-b) + \binom{5}{2}a^3(-b)^2 + \binom{5}{3}a^2(-b)^3 + \binom{5}{4}a(-b)^4 + \binom{5}{5}(-b)^5 = \\
&= a^5 - 5a^4b + 10a^3b^2 - 10a^2b^3 + 5ab^4 - b^5.
\end{aligned}
$$

12.5.2. La potenza di un polinomio

La formula di Newton può essere generalizzata al calcolo della potenza n-esima di un polinomio:

$$
(a_1+a_2+\cdots+a_k)^n = \underbrace{(a_1+a_2+\cdots+a_k)\cdot(a_1+a_2+\cdots+a_k)\cdots(a_1+a_2+\cdots+a_k)}_{n \text{ fattori}}.
$$

Il risultato sarà costituito da una somma di monomi del tipo

$$
a_1^{n_1}a_2^{n_2}\cdots a_k^{n_k}
$$

con opportuni coefficienti, per determinare i quali basterà contare quante volte il monomio compare nello sviluppo, tenendo conto che si dovrà avere $n_1+n_2+\cdots+n_k = n$. Per trovare questo numero

[3] Sia in base alla definizione che per calcolo diretto è immediato che $C_n^k = C_n^{n-k}$.

occorre tenere conto che questo monomio si ottiene scegliendo a_1 in esattamente n_1 fattori, a_2 in esattamente n_2 fattori, e così via. Qualcuno degli n_i potrà essere anche zero; per gli a_i con esponente diverso da zero l'ordine in cui i fattori sono presi non conta e dunque si avrà una permutazione di elementi eventualmente non tutti distinti. Tenendo conto che $0! = 1$, i coefficienti si possono scrivere allora nella forma

$$\frac{n!}{n_1! \cdot n_2! \cdots n_k!}, \quad n_1 + n_2 + \cdots + n_k = n$$

e si avrà

$$(a_1 + a_2 + \cdots + a_k)^n = \sum \frac{n!}{n_1! \cdot n_2! \cdots n_k!} a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k}, \quad n_1 + n_2 + \cdots + n_k = n. \tag{12.12}$$

Questa formula è detta *formula di Leibniz* per la potenza di un polinomio. Nel caso $k = 2$ riproduce esattamente la formula di Newton.

Il fatto che $n_1 + n_2 + \cdots + n_k = n$ mostra che il numero di addendi dello sviluppo è dato da ${}^{\mathsf{R}}\mathsf{C}_k^n$: si presti attenzione al fatto che si tratta di ${}^{\mathsf{R}}\mathsf{C}_k^n$ e non ${}^{\mathsf{R}}\mathsf{C}_n^k$.

Esempio 12.9. Nello sviluppo di $(a + b + c + d)^7$, i coefficienti di a^2b^3cd e ab^5d sono, rispettivamente,

$$\frac{7!}{2! \times 3! \times 1! \times 1!} = 420 \quad \text{e} \quad \frac{7!}{1! \times 5! \times 0! \times 1!} = 42.$$

Il numero totale di addendi distinti dello sviluppo è

$${}^{\mathsf{R}}\mathsf{C}_4^7 = \binom{4+7-1}{7} = \binom{10}{7} = 120.$$

Esempio 12.10. Se vogliamo calcolare il coefficiente di a^2bc nello sviluppo di $(a + b - 2c)^4$ possiamo procedere come segue. Il coefficiente di $a^2b(-2c)$ è

$$\frac{4!}{2! \times 1! \times 1!} = 12;$$

si ha poi

$$12a^2b(-2c) = -24a^2bc.$$

Il coefficiente richiesto è dunque -24.

12.6. Proprietà dei coefficienti binomiali

I coefficienti binomiali godono di alcune importanti proprietà che vogliamo provare, anche come esempio di calcoli tipici con i coefficienti binomiali.

Proprietà di simmetria

$$\binom{n}{k}=\binom{n}{n-k}. \tag{12.13}$$

È una conseguenza della definizione. Scelti k elementi da un insieme di n, posso costruire con essi un insieme di k elementi e con i restanti un insieme di $n-k$ elementi: c'è corrispondenza biunivoca tra i sottoinsiemi di k e quelli di $n-k$ elementi. Si può anche ragionare sulla formula del binomio di Newton: il primo membro rappresenta il coefficiente si $a^{n-k}b^k$, il secondo quello di $a^k b^{n-k}$ ed è ovvio che la scelta di $n-k$ fattori uguali ad a implica la scelta di k fattori uguali a b, e viceversa. Si può anche fare un calcolo diretto, usando la formula (12.8).

$$\binom{n}{n-k}=\frac{n!}{(n-k)!\big(n-(n-k)\big)!}=\frac{n!}{k!(n-k)!}=\binom{n}{k}.$$

Numero dei sottoinsiemi di un insieme

$$2^n=\binom{n}{0}+\binom{n}{0}+\binom{n}{2}+\cdots+\binom{n}{n}. \tag{12.14}$$

È sufficiente applicare la formula del binomio con $a=b=1$. Questa formula ha un'importante applicazione. Dato infatti un insieme E di n elementi, il primo degli addendi del secondo membro rappresenta il numero degli insiemi con 0 elementi (1 insieme, l'insieme vuoto), il secondo degli addendi il numero degli insiemi con 1 elemento, e così via. La somma rappresenta dunque il totale del numero dei sottoinsiemi. Abbiamo già visto una prova di questo fatto nell'osservazione 12.5, nella pagina 420.

Somma a segno alterno di coefficienti binomiali

$$\binom{n}{0}-\binom{n}{0}+\binom{n}{2}-\cdots+(-1)^n\binom{n}{n}=0. \tag{12.15}$$

È sufficiente applicare la formula del binomio con $a=1$ e $b=-1$.

Formula di Stifel

Considerato un insieme E di n elementi, fissiamo un suo elemento, diciamolo a, e contiamo i sottoinsiemi di k elementi separando, nel conteggio, quelli che contengono a da quelli che non lo contengono. I primi sono C_{n-1}^{k-1}, in quanto basta scegliere $k-1$ elementi tra gli $n-1$ diversi da a e poi aggiungere a a ciascuno. I secondi sono, invece, C_{n-1}^{k}, in quanto basta scegliere k elementi tra gli $n-1$ diversi da a. Se ne deduce la seguente formula, nota come *formula di Stifel*:

$$\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k}. \tag{12.16}$$

Essa si può trarre anche per calcolo diretto.

$$\binom{n-1}{k-1}+\binom{n-1}{k}=\frac{(n-1)!}{(k-1)!(n-k)!}+\frac{(n-1)!}{k!(n-k-1)!}=\frac{(n-1)!}{(k-1)!(n-k-1)!}\left(\frac{1}{n-k}+\frac{1}{k}\right)=$$
$$=\frac{(n-1)!}{(k-1)!(n-k-1)!}\,\frac{n}{(n-k)k}=\frac{n!}{k!(n-k)!}=\binom{n}{k}.$$

È questa la formula che sta alla base della costruzione del triangolo di Tartaglia: basta scrivere una generica riga del triangolo e quella che la precede, si veda la figura 12.6.

$$\begin{array}{ccccccccccc}
 & \underbrace{\binom{n-1}{0} \quad \binom{n-1}{1}} & & \cdots & & \underbrace{\binom{n-1}{k-1} \quad \binom{n-1}{k}} & & \cdots & & \underbrace{\binom{n-1}{n-2} \quad \binom{n-1}{n-1}} & \\
\binom{n}{0} & \binom{n}{1} & & \cdots & & \binom{n}{k} & & \cdots & & \binom{n}{n-1} & \binom{n}{n}
\end{array}$$

Figura 12.6.: *Costruzione del triangolo di Tartaglia con la formula di Stifel*

Formula di convoluzione di Vandermonde

Consideriamo ora due insiemi A e B, disgiunti e di cardinalità, rispettivamente, n ed m. Se S è la loro unione, C_{n+m}^{k} rappresenta il numero dei suoi sottoinsiemi di k elementi. Questi sottoinsiemi si costruiscono prendendo r elementi da A e i restanti $k-r$ da B, con $r=1,2,3,\dots,k$. La scelta di r elementi da A si può fare in C_{n}^{r} modi, mentre la scelta di $k-r$ elementi da B si può fare in C_{m}^{k-r} modi; in totale la scelta anzidetta si potrà dunque fare in $\mathsf{C}_{n}^{r}\cdot\mathsf{C}_{m}^{k-r}$ modi. Per avere il numero totale di sottoinsiemi basterà sommare questi numeri per tutti i valori di r. Si ottiene la formula 12.17, detta *formula di convoluzione di Vandermonde*.

$$\binom{n+m}{k}=\binom{n}{0}\binom{m}{k}+\binom{n}{1}\binom{m}{k-1}+\cdots+\binom{n}{k}\binom{m}{0}= \tag{12.17}$$
$$=\sum_{r=0}^{k}\binom{n}{r}\binom{m}{k-r},\qquad k\le\min(n,m).$$

Somma dei quadrati di coefficienti binomiali

Ponendo $n=m=k$ nella formula (12.17), e tenendo presente la proprietà di simmetria (12.13), si ottiene la

$$\binom{2n}{n}=\binom{n}{0}^2+\binom{n}{1}^2+\binom{n}{2}^2+\cdots+\binom{n}{n}^2. \tag{12.18}$$

Estensione della definizione di coefficiente binomiale

In molte circostanze è utile estendere la definizione del coefficiente binomiale

$$\binom{n}{k}$$

al caso che n sia un reale qualunque. Precisamente si pone

$$\binom{r}{k}=\frac{r(r-1)(r-2)\cdots(r-k+1)}{k!},\quad r\in\mathbb{R},\ k\in\mathbb{N}. \tag{12.19}$$

Si noti che, se n è un naturale minore di k, si ha

$$\binom{n}{k}=0,\quad n\in\mathbb{N},\ k\in\mathbb{N},\ n<k.$$

Inoltre

$$\binom{-1}{k}=\frac{-1(-1-1)(-1-2)\cdots(-1-k+1)}{k!}=\frac{(-1)^k\,k!}{k!}=(-1)^k.$$

12.7. Esercizi

Esercizio 12.1. *In una gara con 25 concorrenti vengono premiati i primi cinque.*

1. *Quante sono le possibili assegnazioni dei premi?*
2. *Quante sono le possibili assegnazioni dei premi se si sa che il concorrente Rossi è sicuramente tra essi?*
3. *Quante sono le possibili assegnazioni dei premi se si sa che il concorrente Rossi arriverà secondo?*

Risoluzione. Nel primo caso basta fare $\mathrm{V}_{25}^5=6375600$. Nel secondo caso basta disporre i 24 concorrenti diversi da Rossi nelle quattro posizioni possibili ($\mathrm{V}_{24}^4=255024$) e poi piazzare Rossi, che potrà andare in uno dei cinque posti: $255024\cdot 5=1275120$. Nell'ultimo caso si dovranno solo piazzare i 2 concorrenti diversi da Rossi nelle 4 posizioni possibili: $\mathrm{V}_{24}^4=255024$. □

Esercizio 12.2. *Quante sono le possibili targhe di automobili costruite con due lettere, tre numeri, due lettere, se l'alfabeto è di* 26 *simboli e se la prima lettera del primo gruppo non deve essere la "z"?.*

Risoluzione. Per quanto riguarda le lettere si tratta di trovare le disposizioni con ripetizione di 26 oggetti, di classe 4, e poi togliere quelle che cominciano con zeta; le prime sono 26^4, le seconde sono le disposizioni di 26 oggetti di classe 3, ovvero 26^3. Si ottiene 439400. Naturalmente si potevano anche disporre tutte le 26 lettere nei posti due, tre, quattro e solo le 25 lettere diverse da zeta nel posto uno, ottenendo sempre lo stesso risultato. I numeri di tre cifre sono, naturalmente, 1000 (si tenga conto che, per esempio, 12 deve essere in realtà pensato come 012). In totale le targhe possibili sono dunque 439400000. □

Esercizio 12.3. *Risolvere l'equazione* $\mathrm{V}_x^5=3\mathrm{V}_x^4$.

Risoluzione. Si deve tenere conto che $x\in\mathbb{N}\wedge x\geq 5$. Dopodiché si può scrivere

$$x(x-1)(x-2)(x-3)(x-4)=3x(x-1)(x-2)(x-3).$$

Date le condizioni, l'unica soluzione accettabile è $x=7$. □

Esercizio 12.4. *In una gara devono venire premiati i primi tre classificati per ogni categoria di partecipanti. Se le categorie sono tre e i partecipanti sono* 10, 15, 18 *rispettivamente, quante sono le possibili premiazioni?*

Risoluzione. Si tratta di fare $\mathsf{V}_{10}^3 \cdot \mathsf{V}_{15}^3 \cdot \mathsf{V}_{18}^3$, ottenendo 9623577600. Come spesso succede nel calcolo combinatorio si tratta di un numero enorme! □

Esercizio 12.5. *Quanti sono i numeri minori o uguali a* 440000 *che hanno tre cifre* 4 *contigue e un solo 7?*

Risoluzione. La cosa più semplice da fare è di esaminare i casi possibili, contandoli in maniera opportuna. Intanto osserviamo che le tre cifre 4 contigue non possono stare ad inizio numero, e così pure il 7. Le possibilità sono quelle indicate di seguito.

— Ci sono 4×8 possibilità del tipo $*7444*$ (Al primo posto si può avere 0, 1, 2, 3; all'ultimo 0, 1, 2, 3, 5, 6, 8, 9).
— Ci sono $4 \cdot 8$ possibilità del tipo $*7*444$ (Al primo posto si può avere 0, 1, 2, 3; al terzo 0, 1, 2, 3, 5, 6, 8, 9).
— Ci sono $4 \cdot 9$ possibilità del tipo $*4447*$ (Al primo posto si può avere 0, 1, 2, 3; all'ultimo 0, 1, 2, 3, 4, 5, 6, 8, 9).
— Ci sono $4 \cdot 8$ possibilità del tipo $*444*7$ (Al primo posto si può avere 0, 1, 2, 3; al quinto 0, 1, 2, 3, 5, 6, 8, 9).
— Esaminiamo ora il caso $**4447$: al primo posto si può avere 0, 1, 2, 3 e al secondo posto 0, 1, 2, 3, 5, 6, 8, 9 (32 possibilità); se invece al primo posto si ha un 4, al secondo si deve avere 0, 1, 2, 3, cioè altre quattro possibilità, per un totale di 36.
— Esaminiamo infine il caso $**7444$: al primo posto si può avere 0, 1, 2, 3 e al secondo posto 0, 1, 2, 3, 4, 5, 6, 8, 9 (36 possibilità); se invece al primo posto si ha un 4, al secondo si deve avere 0, 1, 2, 3, cioè altre quattro possibilità, per un totale di 40.

In totale si hanno 208 possibilità. Questo esercizio è interessante perché mostra che, nella sostanza, conviene ridursi a fare un "conteggio bruto". □

Esercizio 12.6. *Quante sono le applicazioni iniettive di un insieme di* 6 *elementi in un insieme di* 8 *elementi?*

Risoluzione. Sono $\mathsf{V}_8^6 = 20160$. □

Esercizio 12.7. *Quanti sono i numeri di 5 cifre tutte distinte che iniziano con* 3*?*

Risoluzione. Si tratta di disporre le cifre diverse da 3 ai posti $2, 3, 4, 5$. Si ottiene $\mathsf{V}_9^4 = 3024$. □

Esercizio 12.8. *In quanti modi posso sistemare* 3 *camicie diverse (una Blu, una Rossa, una Gialla) in* 2 *cassetti, se è consentito che un cassetto possa anche restare vuoto?*

Risoluzione. Si tratta delle disposizioni, con ripetizione, di 2 oggetti a tre a tre: ${}^{\mathsf{R}}\mathsf{V}_2^3 = 2^3 = 8$. Le possiamo visualizzare nel seguente schema:

1	B-R-G	B	B-R	B-G		R-G	R	G
2		R-G	G	R	B-R-G	B	B-G	B-R

Le disposizioni richieste si possono anche visualizzare pensando a 3 caselle, denominate B,R,G, da riempire con i simboli $1, 2$ che indicano i cassetti dove le camicie vengono sistemate:

B	1	1	1	1	2	2	2	2
R	1	2	1	2	2	1	1	2
G	1	2	2	1	2	1	2	1

In questo schema il concetto di "ripetizione" è più evidente. □

Esercizio 12.9. *Uno studente deve sostenere 7 esami per ognuno dei suoi tre anni di corso, senza poter rimandare esami all'anno successivo, ma nell'ordine da lui scelto. In quanti modi lo può fare?*

Risoluzione.

$$7! \times 7! \times 7! = 128024064000.$$ □

Esercizio 12.10. *Quanti sono gli anagrammi della parola "luciano" che non cominciano per "l"?*

Risoluzione. Gli anagrammi della parola luciano sono 7!, perché le lettere sono tutte distinte. A questi bisogna togliere quelli che cominciano per l. Questi sono tutti quelli che sui sei posti dal 2 al 7 hanno una delle restanti sei lettere, e cioè 6!. Il numero richiesto è dunque $7! - 6! = 6 \times 6! = 4320$. □

Esercizio 12.11. *Un gruppo di* 4 *italiani,* 3 *francesi e* 5 *tedeschi devono sedersi ad uno stesso tavolo. Le persone di stessa nazionalità devono rimanere vicine. In quanti modi si può fare?*

Risoluzione. Ci sono 4! permutazioni degli italiani, 3! dei francesi, 5! dei tedeschi. Ci sono poi 3! permutazioni dei gruppi. In totale $4! \times 3! \times 5! \times 3! = 103680$. □

Esercizio 12.12. *I* 24 *studenti della classe quinta B devono essere divisi in tre gruppi di* 8 *studenti ciascuno, che lavoreranno indipendentemente per una ricerca. In quanti modi la cosa si può fare?*

Risoluzione. Si tratta in sostanza di calcolare in quanti modi si possono permutare i 24 studenti della classe, tenendo conto che una permutazione tra gli elementi del primo gruppo, o del secondo, o del terzo, non cambia il risultato. Il numero richiesto è allora $24!/(8! \times 8! \times 8!) = 9465511770$. □

Esercizio 12.13. *In quanti modi si possono far accomodare tre ospiti in un albergo che ha quattro stanze singole?*

Risoluzione. Basterà calcolare tutte le possibili permutazioni dei numeri di stanza ($4! = 24$) e poi sistemare i tre ospiti nelle prime tre (per esempio) stanze di ogni permutazione. □

Esercizio 12.14. *In quanti modi si possono sistemare* 20 *paia di calzini in* 3 *cassetti in modo che ce ne vadano* 10 *nel primo,* 7 *nel secondo e* 3 *nel terzo?*

Risoluzione. Basterà calcolare le permutazioni di 20 oggetti divisi in tre gruppi di 10, 7 e 3; si trovano 22170720 modi. □

Esercizio 12.15. *In quanti modi si possono distribuire le* 52 *carte del bridge a un giocatore? E in quanti modi lo stesso giocatore può ricevere carte tutte di valore diverso?*

Risoluzione. Per la prima parte si tratta di scegliere 13 carte da un totale di 52, quindi C_{52}^{13} (che equivale a 635013559600 modi). Per la seconda parte si osservi che si deve scegliere un asso su 4, un due su 4 e così via e che ciascuna scelta è indipendente dalle altre: si hanno $4^{13} = 67108864$ possibilità. □

Esercizio 12.16. *Quante sono le diagonali di un poligono di n lati?*

Risoluzione. Se il poligono ha n lati, ha anche n vertici. Si tratta di calcolare le combinazioni (non conta l'ordine perché, per esempio, il segmento $\overline{AB}$ e il segmento $\overline{BA}$ sono identici) dei vertici a due a due, e poi togliere i lati. Si ottiene

$$\mathrm{C}_n^2 - n = \frac{n(n-3)}{2}.$$

Si poteva anche fare un ragionamento elementare: da ogni punto si può condurre una diagonale ad ogni altro punto, tranne il punto stesso e i due "adiacenti", dunque $n(n-3)$; tenendo conto che in questo modo ogni diagonale viene contata due volte si ritrova il risultato precedente. □

Esercizio 12.17. *Dati n punti nel piano, quanti sono, al massimo i punti di intersezione delle rette che congiungono due di essi, se si escludono le intersezioni sui punti dati?*

Risoluzione. Ci sono C_n^2 rette che congiungono i punti. Detto r questo numero, in totale i punti di intersezione tra le rette sono al massimo C_r^2 (ogni coppia di rette può avere un punto di intersezione). Si deve poi tenere conto che, per ognuno degli n punti dati ci sono $n-1$ rette che si intersecano tra di loro in quel punto: bisogna togliere C_{n-1}^2 intersezioni dal numero precedente, per ogni punto. Nel caso del triangolo, per esempio, non ci sono intersezioni possibili del tipo detto. □

Esercizio 12.18. *Dieci copie di un libro devono essere ripartiti fra quattro scuole. In quanti modi si può fare? E se ogni scuola ne deve ricevere almeno uno?*

Risoluzione. Si tratta di un classico problema di suddivisione di k oggetti identici in n sottoinsiemi, problema equivalente alle combinazioni con ripetizione di n oggetti di classe k. Il numero richiesto è ${}^{\mathrm{R}}\mathrm{C}_4^{10} = \mathrm{C}_{4+10-1}^{10} = 286$. Se poi ogni scuola deve riceverne almeno uno allora si tratta di dividere, fra le quattro scuole, 6 libri invece di 10. Si trovano 84 possibilità. □

Esercizio 12.19. *Se si ha a disposizione una somma di* 10000€*, e si decide di investirla, in tranche da* 500€*, in azioni di quattro società, in quanti modi si può fare?*

Risoluzione. Si tratta in sostanza di distribuire 20 oggetti identici (le 20 tranche da 500€) in quattro contenitori diversi, ovvero di calcolare le combinazioni con ripetizione di quattro oggetti di classe 20: ${}^{\mathrm{R}}\mathrm{C}_4^20 = \mathrm{C}_{4+20-1}^{20} = 1771$. □

Esercizio 12.20. *Provare la formula*

$$\binom{n}{0} - \binom{n}{0} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n} = 0$$

facendo uso della formula di Stifel.

Risoluzione. Si possono scomporre tutti gli addendi del primo membro, tranne il primo e l'ultimo, usando la formula citata:

$$\binom{n}{0} - \binom{n}{0} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n} =$$

$$= \binom{n}{0} - \left[\binom{n-1}{0} + \binom{n-1}{1}\right] + \left[\binom{n-1}{1} + \binom{n-1}{2}\right] - \cdots$$
$$+ (-1)^{n-1}\left[\binom{n-1}{n-2} + \binom{n-1}{n-1}\right] + (-1)^n \binom{n}{n}.$$

Da qui si vede facilmente che gli addendi si elidono a coppie. □

Esercizio 12.21. *Dimostrare, per calcolo diretto, che*

$$\binom{n}{2} = \binom{k}{2} + \binom{n-k}{2} + k(n-k).$$

Ripetere la dimostrazione usando un ragionamento sui sottoinsiemi di cardinalità 2 *di un dato insieme.*

Risoluzione. Si ha

$$\binom{k}{2} + \binom{n-k}{2} + k(n-k) = \frac{k(k-1)}{2} + \frac{(n-k)(n-k-1)}{2} + k(n-k) = \frac{n(n-1)}{2} = \binom{n}{2}.$$

Ragionando invece sui sottoinsiemi, osserviamo che il primo membro della formula data fornisce il numero dei sottoinsiemi di cardinalità 2 di un dato insieme. Se immaginiamo di dividere l'insieme in due parti, uno di cardinalità k e l'altro di cardinalità $n-k$, per costruire i sottoinsiemi di cardinalità 2 si dovranno contare quelli della prima parte, quelli della seconda parte e, infine, quelli costruiti prendendo un elemento dalla prima parte e uno dalla seconda parte: sono i tre addendi del secondo membro. □

Esercizio 12.22. *Provare la formula*

$$\sum_{i=0}^{n} i\binom{n}{i} = 0\binom{n}{0} + 1\binom{n}{1} + \cdots + (n-1)\binom{n}{n-1} + n\binom{n}{n} = n \cdot 2^{n-1}.$$

Risoluzione. Sfruttiamo la proprietà che i coefficienti binomiali equidistanti dagli estremi sono uguali. Osserviamo poi che la somma dei coefficienti dei termini equidistanti dagli estremi è sempre n (se n è pari da questo discorso rimane escluso il termine centrale il cui valore è $n/2$). In ogni caso tutti i coefficienti possono essere sostituiti con $n/2$, senza alterare il valore della somma. Si ha allora

$$\sum_{i=0}^{n} i\binom{n}{i} = \frac{n}{2}\sum_{i=0}^{n}\binom{n}{i} = \frac{n}{2}2^n = n \cdot 2^{n-1}.$$ □

Esercizio 12.23. *Quanti sono i numeri di* 6 *cifre con almeno una cifra dispari? E quelli con almeno una cifra pari?*

Risoluzione. Intanto calcoliamo quanti sono i numeri di sei cifre; tenendo conto che non possono cominciare per 0, sono $9 \times 10 \times 10 \times 10 \times 10 \times 10 = 900000$. Posso poi calcolare i numeri che hanno solo cifre pari; in questo caso ho 4 scelte per il primo posto e 5 per ciascuno degli altri 5, quindi $4 \times 5^5 = 12500$. Invece i numeri che hanno solo cifre dispari sono 5^6 (lo zero non è dispari), cioè 15625. Si ottiene $900000 - 12500 = 887500$ per la prima richiesta, 884375 per la seconda. □

Esercizio 12.24. *Quante sono le soluzioni naturali dell'equazione* $x+y+z=33$*, con le condizioni* $x \geq 4$ *e* $y \geq 7$*?*

Risoluzione. Se non ci fossero condizioni si tratterebbe di trovare le combinazioni con ripetizione di 3 oggetti di classe 33, che sono 595. Per tenere conto delle condizioni basterà risolvere l'equazione $(x'+4)+(y'+7)+z=33$, ovvero $x'+y'+z=22$: ora le incognite non hanno limitazioni e quindi si devono cercare le combinazioni con ripetizione di 3 oggetti di classe 22, ovvero 276. □

Esercizio 12.25. *Si vogliono disporre su una scacchiera* 3 *torri, in modo che nessuna possa "mangiare" le altre. In quanti modi si può fare?*

Risoluzione. La prima torre può essere piazzata in ognuna delle 64 caselle (8^2 possibilità). Questa torre minaccia 7 caselle in orizzontale e 7 in verticale, per un totale di 14 caselle. La seconda torre allora non potrà essere piazzata né in queste 14 caselle, né naturalmente nella casella occupata dalla prima torre: le restano $64-14-1=49=7^2$ possibilità. Questa seconda torre minaccia ancora 14 caselle, ma due le ho già contate con la prima torre. Per la terza torre devo quindi escludere $14+12$ caselle minacciate, più 2 occupate. Ne restano $36=6^2$. In totale avrei $8^2 \times 7^2 \times 6^2$ possibilità, ma devo tenere conto che le tre torri sono indistinguibili, per cui una loro permutazione non modifica nulla. In totale si ha allora $8^2 \times 7^2 \times 6^2/3! = 18816$. La formula trovata si presta ad una facile generalizzazione "ad occhio". Con cinque torri avrei $8^2 \times 7^2 \times 6^2 \times 5^2 \times 4^2/5! = 376320$ possibilità. Con la stessa generalizzazione si può anche concludere che non potrei disporre più di 8 torri con questa condizione. □

Esercizio 12.26. *Si devono disporre su una fila di* 10 *sedie cinque coppie uomo-donna. In quanti modi la cosa si può fare se la disposizione può essere fatta alla rinfusa? E se le donne e gli uomini devono rimanere vicini tra di loro? E se le coppie devono rimanere unite?*

Risoluzione. Nel primo caso ci sono 10! (3628800) possibilità. Nel secondo si tratta di calcolare le possibili permutazioni dei maschi (5!), delle femmine (ancora 5!) e dei due gruppi (2!): 28800 possibilità. Nell'ultimo caso di tratta di permutare ciascuna coppia nei 2! modi possibili e poi le 5 coppie nei 5! modi possibili: 480 possibilità. □

Esercizio 12.27. *In una classe di* 22 *studenti, di cui* 12 *femmine e* 10 *maschi, si deve formare un gruppo per una ricerca costituito da* 3 *maschi e* 3 *femmine. In quanti modi la cosa si può fare se nei dieci maschi ci sono due gemelli e si decide che non possano stare insieme?*

Risoluzione. Se non ci fosse alcuna restrizione il numero delle scelte possibili sarebbe $C_{12}^3 \cdot C_{10}^3 = 26400$. Il numero delle possibilità tra le femmine non è alterato dalla restrizione. Tra i maschi basterà che contiamo in quanti modi si può operare una scelta che contenga i due gemelli: questo numero andrà sottratto dal totale delle possibili scelte tra i maschi. E' immediato che se un gruppo di 3 maschi deve averne due predefiniti (i gemelli), restano $10-2$ possibilità di scelta. Il numero cercato è allora $C_{12}^3\left(C_{10}^3-8\right)=24640$. □

Esercizio 12.28. *In quanti modi n palline indistinguibili possono venire collocate in n celle numerate, in modo che esattamente una cella rimanga vuota? Che cosa cambia se le palline sono distinguibili?*

Risoluzione. Bisogna innanzitutto scegliere la cella vuota (si può fare in n modi). Successivamente si deve collocare almeno una pallina nelle restanti $n-1$ celle e poi scegliere una cella dove collocare due palline (si può fare in $n-1$ modi). In totale la collocazione si può fare in $n(n-1)$ modi. Se invece le palline sono distinguibili, dopo aver scelto la cella vuota e quella che deve contenere due palline, cosa che si può fare in $n(n-1)$ modi come prima, occorre ancora scegliere quali palline mettere nelle varie celle. Ci sono C_n^2 modi di scegliere le due palline da mettere nella cella che le conterrà, e poi $(n-2)!$ modi di distribuire le restanti palline nelle restanti celle. In totale ci sono

$$n(n-1)\frac{n!}{(n-2)!\,2!}(n-2)! = n(n-1)\frac{n!}{2}$$

modi per collocare le palline come richiesto. □

Esercizio 12.29. *Se da un campione di n oggetti in cui ve ne sono k di difettosi se ne prendono a caso m, quante sono le possibilità che ci siano p pezzi difettosi ($p \le k$ e $p \le m$)?*

Risoluzione. È un problema classico di controllo di qualità a campione. Intanto si possono scegliere p oggetti difettosi tra i k in C_k^p modi. In un'estrazione di m pezzi, ciascuno di questi campioni difettosi si potrà combinare con uno costituito da oggetti tutti non difettosi: la scelta si può fare in C_{n-k}^{m-p} modi. Basterà fare il prodotto di questi due numeri per avere il numero richiesto:

$$\mathsf{C}_k^p \cdot \mathsf{C}_{n-k}^{m-p} = \binom{k}{p}\cdot\binom{n-k}{m-p}.$$ □

Esercizio 12.30. *Se si lancia* 8 *volte un dado, in quanti modi si possono ottenere* 4 *coppie diverse di numeri uguali?*

Risoluzione. Intanto bisognerà scegliere 4 tra i 6 numeri che compaiono nelle facce di un dado: ciò si può fare in C_6^4 modi diversi, ovvero in 15 modi. Fissata ora la scelta delle coppie, per esempio quelle formate dai numeri 1, 2, 3, 4, dobbiamo contare in quanti modi compaiono, in qualunque ordine, due 1, due 2, due 3, due 4 su otto caselle: si tratta delle permutazioni di otto oggetti divisi in quattro gruppi da 2, ovvero $\mathsf{P}_8^{2,2,2,2} = 2520$. In totale ci sono 37800 modi. □

Esercizio 12.31. *Tre ragazze e due ragazzi si siedono a tavola in cinque posti consecutivi. In quanti modi possono sedersi se ogni femmina vuole avere a fianco almeno un maschio e viceversa?.*

Risoluzione. La cosa più semplice da fare è tracciare uno schema ad albero ("grafo") esaminando le varie possibilità. Si può partire da un maschio o da una femmina e si ottiene uno schema come i seguenti, dove nell'ultima colonna abbiamo evidenziato le situazioni non accettabili e dove, nella costruzione dei rami, ci siamo fermati quando la situazione diventava inaccettabile.

M	M				no
	F	M	F	F	no
	F	F	M	F	
	F	F	F		no

F	M	M	F	F	no
	M	F	M	F	
	M	F	F	M	
	F				no

Dunque le uniche possibilità sono MFFMF (e qui ho 2 possibilità di scelta al primo posto, 3 al secondo, 2 al terzo e 1 per i due rimanenti, in totale 12), $FMFMF$ (e qui ho 3 possibilità di scelta al primo posto, 2 al secondo, 2 al terzo e 1 per i due rimanenti, in totale 12), FMFFM (e qui ho 3 possibilità di scelta al primo posto, 2 al secondo, 2 al terzo e 1 per i due rimanenti, in totale 12). In totale 36 possibili disposizioni. □

Esercizio 12.32. *Provare la formula*

$$2^n = \binom{n}{0} + \binom{n}{0} + \binom{n}{2} + \dots + \binom{n}{n}$$

contando in due modi diversi i numeri di n cifre in binario.

Risoluzione. I numeri di n cifre (contando anche quelli che cominciano per 0) sono le disposizioni con ripetizione di 2 simboli (0 ed 1) a n a n, e pertanto sono 2^n. Contiamo ora questi numeri in altro modo. Iniziamo con quelli che non contengono nessun 1: ce n'è uno solo, cioè C_n^0, quello cha tutti 0. I numeri che contengono un solo 1 sono n, ovvero C_n^1, corrispondenti alle possibili scelte di 1 posto su n dove mettere il numero 1. I numeri con due 1 sono C_n^2, corrispondenti alle possibili scelte delle due posizioni, su n in cui mettere gli 1. La conclusione è ora banale.

Il ragionamento può essere esteso alla verifica della formula

$$m^n = \sum_{k=0}^{n} \binom{n}{k} (m-1)^{n-k}.$$

Anch'essa può, come la precedente, discendere direttamente dalla formula del binomio applicato al caso $a = 1$, $b = m - 1$, ma può essere verificata contando i numeri di n cifre costruiti con un sistema di m simboli: $\{ s_1, s_2, \dots, s_m \}$. Il primo membro è il conteggio diretto mediante le disposizioni con ripetizione degli m simboli. Il secondo memebro si ottiene cominciando a contare i numeri che non contengono il simbolo s_1: sono tanti quanti i numeri che contengono solo i restanti simboli, ovvero

$$(m-1)^n = \binom{n}{0} (m-1)^n.$$

La quantità di numeri che contengono 1 volta il simbolo s_1 si ottiene contando i posti dove mettere il simbolo e moltiplicando per il numero di modi di disporre i restanti $m - 1$ simboli:

$$\binom{n}{1} (m-1)^{n-1}.$$

Anche ora la conclusione è immediata. □

Esercizio 12.33. *Provare la formula*

$$n! = \binom{n}{k} k!(n-k)!,$$

ragionando sulle permutazioni di un insieme di n oggetti.

Risoluzione. Il primo membro rappresenta le permutazioni indicate. Si tratta di provare a contarle in un altro modo. Si possono scegliere k degli n oggetti, in C_n^k modi diversi, e poi fare le $k!$ loro permutazioni e le $(n-k)!$ permutazioni dei rimanenti. In questo modo si è provata la formula richiesta. Si noti come questo ragionamento costituisca una dimostrazione alternativa della nota formula riguardante i coefficienti binomiali. □

Parte II.

Approfondimenti

La seconda parte del testo contiene una raccolta di argomenti sviluppati solo in alcuni corsi di scuola secondaria superiore. Il contenuto di questi capitoli può essere utilizzato anche come utile premessa ai corsi universitari dove questi argomenti saranno trattati con molto maggiore dettaglio.

13. Le funzioni iperboliche

13.1. Rivisitiamo le funzioni trigonometriche

Le funzioni trigonometriche, in particolare le funzioni seno e coseno, sono state definite a partire dalla circonferenza goniometrica: dato un numero reale x e il punto $\mathsf{A}(1,0)$ abbiamo individuato, sulla circonferenza goniometrica, il punto P tale che la misura dell'arco $\widehat{\mathsf{AP}}$ sia $|x|$, muovendoci in verso antiorario se $x>0$, in verso orario se $x<0$. Il valore della misura di questo arco coincide anche con la misura in radianti dell'angolo (generalizzato) al centro corrispondente. Il coseno e il seno di x sono rispettivamente l'ascissa e l'ordinata del punto P.

Considerato ora il settore circolare AOP rappresentato nella figura 13.1, notiamo che esso ha area $|x|/2$. Dunque possiamo anche definire il coseno e il seno di un numero reale x, compreso tra -2π e 2π come, rispettivamente, l'ascissa e l'ordinata del punto P sulla circonferenza goniometrica che individua il settore circolare AOP di area $|x|/2$ e che è ottenuto da A mediante una rotazione in senso orario se $x<0$, antiorario se $x>0$. Per gli altri valori basterà procedere per periodicità.

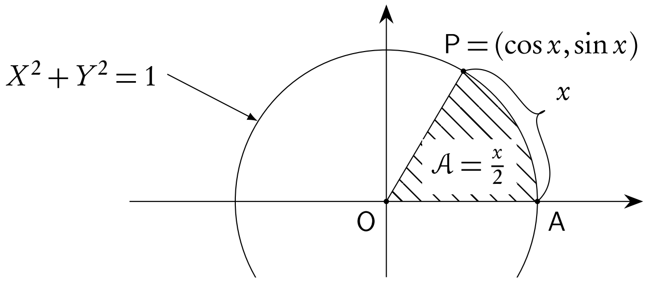


Figura 13.1.: *Settore circolare e funzioni trigonometriche*

Sostituiamo ora la circonferenza goniometrica con il ramo "destro" dell'iperbole equilatera di equazione $X^2-Y^2=1$ e, dato il numero reale x, individuiamo sull'iperbole il punto P tale che l'area del settore iperbolico AOP evidenziato nella figura 13.2 sia $|x|/2$ e che sta nel primo o terzo quadrante a seconda che x sia positivo o negativo. Si dà a questo punto la seguente definizione.

Definizione 13.1. *L'ascissa e l'ordinata del punto P individuato come sopra descritto sul ramo destro dell'iperbole equilatera $X^2-Y^2=1$ si chiamano, rispettivamente,* coseno iperbolico *e* seno iperbolico *del numero reale x e si scrive*

$$X_{\mathsf{P}}=\cosh x,\quad Y_{\mathsf{P}}=\sinh x. \tag{13.1}$$

I nomi sono legati al fatto che queste funzioni godono di proprietà straordinariamente simili a quelle delle funzioni trigonometriche, in particolare per quanto riguarda le formule di addizione e sottrazione e quelle che ne conseguono, come vedremo.

Si può verificare, ma la cosa esula dal nostro contesto, che l'area del settore iperbolico in questione diventa infinitamente grande se P si allontana sul ramo di iperbole, per cui queste funzioni hanno dominio tutto l'asse reale e, a differenza di quelle trigonometriche, non sono periodiche.

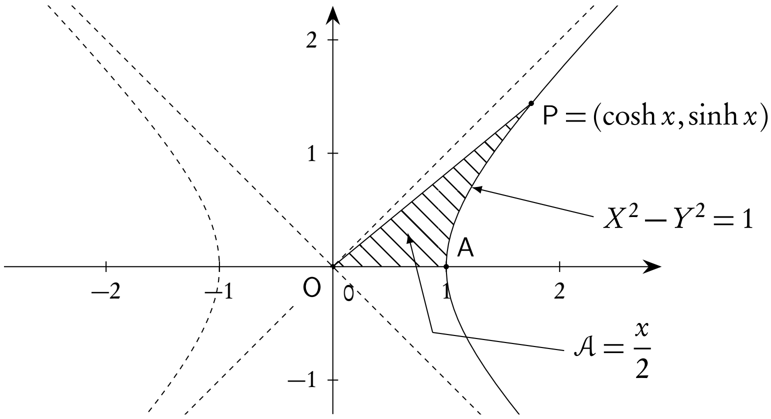


Figura 13.2.: *Settore iperbolico e funzioni iperboliche*

Queste funzioni non aumentano il numero delle funzioni elementari, in quanto si può provare (ma anche questo esula dal nostro contesto) che esse sono legate alla funzione esponenziale secondo le formule (13.2).

$$\cosh x = \frac{e^x + e^{-x}}{2}, \quad \sinh x = \frac{e^x - e^{-x}}{2}. \tag{13.2}$$

Si noti che dalle (13.2) si ottiene

$$e^x = \cosh x + \sinh x. \tag{13.3}$$

Dunque si sarebbe anche potuto definire la funzione esponenziale a partire dalle funzioni iperboliche, introducendo in sostanza il concetto di potenza di base ed esponente reale in maniera alternativa a quanto accennato nel paragrafo 7.3 del capitolo 7.

I grafici di queste funzioni sono proposti nella figura 13.3.

La funzione coseno iperbolico è una funzione pari (come il coseno) e sempre positiva, con minimo di valore 1 nell'origine; la funzione seno iperbolico è una funzione dispari (come il seno) e strettamente crescente: tutto questo risulta anche dalla definizione geometrica che ne abbiamo dato usando i settori iperbolici.

In analogia con quanto fatto per le funzioni trigonometriche si definiscono anche altre funzioni iperboliche. Ci occuperemo solo della *tangente iperbolica* definita come rapporto tra seno e coseno iperbolici:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{13.4}$$

Il suo grafico è proposto nella figura 13.4. La funzione è dispari (come la tangente) e strettamente crescente, inoltra ha due asintoti orizzontali, $y = \pm 1$.

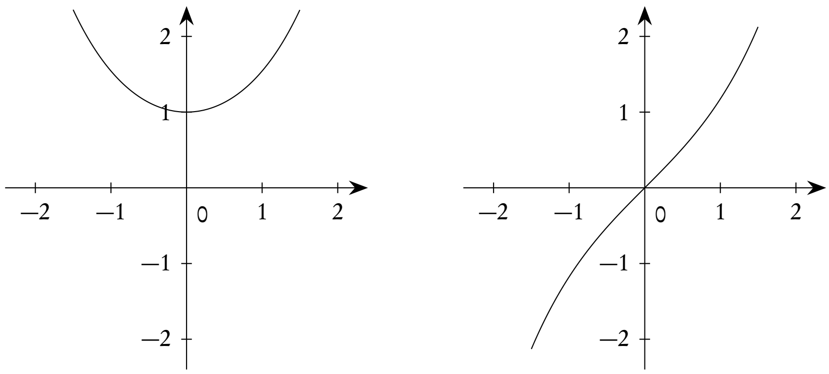


Figura 13.3.: *Grafici delle funzioni coseno e seno iperbolico*

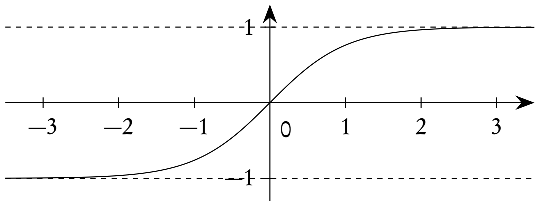


Figura 13.4.: *Grafico della funzione tangente iperbolica*

13.2. Le funzioni iperboliche inverse

Il problema di costruire le inverse delle funzioni iperboliche è decisamente più semplice dell'analogo problema con le funzioni trigonometriche. Per le funzioni seno iperbolico e tangente iperbolica, che sono strettamente crescenti, non ci sono difficoltà, per la funzione coseno iperbolico si opera una restrizione sul dominio ai reali positivi. Per il seno non si opera nessuna restrizione nemmeno sul codominio, per il coseno si restringe il codominio all'intervallo $[1,+\infty[$, per la tangente all'intervallo $]-1,1[$.

Le funzioni inverse, nella lingua italiana, prendono il nome di *settore coseno iperbolico*, *settore seno iperbolico* e *settore tangente iperbolica*. Questi nomi sono significativi, in ragione della definizione geometrica che abbiamo dato: a partire dal coseno, dal seno o dalla tangente iperbolica esse forniscono il numero x il cui modulo fornisce la doppia area del settore iperbolico che è servito per costruirle, con la restrizione che, nel caso del coseno iperbolico, si ottiene solo il valore positivo. Per le funzioni trigonometriche, in perfetta analogia, a partire dalle funzioni coseno, seno e tangente, le inverse forniscono quell'arco che serve a definire il coseno, il seno e tangente, con opportune restrizioni per evitare di avere funzioni plurivoche: per questo motivo le funzioni si chiamano arccoseno, arcseno, arctangente. La simbologia ufficiale, a cui abbiamo deciso di adeguarci, si basa però sul generico nome *area hyperbolic function*, anch'esso decisamente significativo. Si adottano le seguenti scritture:

— per l'inversa del coseno iperbolico, $\operatorname{arcosh}(x)$;
— per l'inversa del seno iperbolico, $\operatorname{arsinh}(x)$;
— per l'inversa della tangente iperbolica, $\operatorname{artanh}(x)$.

Segnaliamo anche altri nomi di uso comune: settsinh(x), argsinh(x), arcsinh(x), asinh(x), e analoghi per le altre funzioni. Si presti attenzione al fatto che nella convenzione UNI CEI ISO si usa "arsinh" (area seno iperbolico) e *non* arcsinh, proprio per distinguere questo caso da quello delle funzioni trigonometriche, dove si usa proprio (e coerentemente) "arcsin" (arco seno).

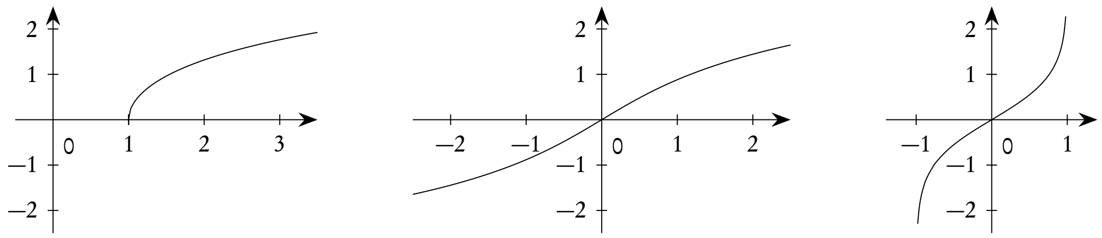


Figura 13.5.: *Grafici delle funzioni inverse di coseno, seno e tangente iperboliche*

Le funzioni iperboliche si possono esprimere utilizzando la funzione esponenziale; le loro inverse si possono esprimere utilizzando la funzione logaritmo naturale. È un utile esercizio eseguire i calcoli necessari.

Per il settore seno iperbolico si ha:

$$y=\frac{e^x-e^{-x}}{2}\ \Rightarrow\ e^{2x}-2ye^x-1=0\ \Rightarrow\ e^x=y+\sqrt{y^2+1}\ \Rightarrow\ x=\ln(y+\sqrt{y^2+1}),$$

dove abbiamo tenuto del fatto che $e^x>0$. Rinominando[1] la variabile indipendente come x si ottiene

$$\text{(13.5)}\qquad \operatorname{arsinh}(x)=\ln(x+\sqrt{x^2+1}),\quad x\in\mathbb{R}.$$

Per il settore coseno iperbolico si ha:

$$y=\frac{e^x+e^{-x}}{2}\ \Rightarrow\ e^{2x}-2ye^x+1=0\ \Rightarrow\ e^x=y+\sqrt{y^2-1}\ \Rightarrow\ x=\ln(y+\sqrt{y^2-1}),$$

dove abbiamo tenuto conto del fatto che vogliamo avere $x>0$. Rinominando, come prima, la variabile indipendente si ottiene

$$\text{(13.6)}\qquad \operatorname{arcosh}(x)=\ln(x+\sqrt{x^2-1}),\quad x>1.$$

Infine per la tangente iperbolica si ha:

$$y=\frac{e^x-e^{-x}}{e^x+e^{-x}}\ \Rightarrow\ e^{2x}(y-1)+y+1=0\ \Rightarrow\ e^{2x}=\frac{1+y}{1-y}\ \Rightarrow\ x=\frac{1}{2}\ln\left(\frac{1+y}{1-y}\right),$$

da cui

$$\text{(13.7)}\qquad \operatorname{artanh}(x)=\frac{1}{2}\ln\left(\frac{1+x}{1-x}\right),\quad -1<x<1.$$

[1] Si tratta di una situazione standard: quando si vuole ricavare l'espressione analitica dell'inversa di una funzione si può partire dall'equazione $y=f(x)$ e ricavare la x in funzione di y. Terminati i calcoli si usa scambiare i nomi delle variabili in modo da rispettare la tradizione che vuole la x come variabile indipendente e la y come variabile dipendente.

13.3. Formule coinvolgenti le funzioni iperboliche

Le formule che riguardano le funzioni iperboliche sono molto simili a quelle relative alle funzioni trigonometriche e questo giustifica la nomenclatura utilizzata. Riportiamo qui solo quelle relative alle funzioni seno e coseno.

Formula fondamentale

La formula fondamentale sulle funzioni iperboliche discende direttamente dal fatto che il punto di coordinate $(\cosh x, \sinh x)$ appartiene all'iperbole equilatera di equazione $X^2 - Y^2 = 1$:

$$\cosh^2 x - \sinh^2 x = 1. \tag{13.8}$$

Formule di addizione e sottrazione

Si possono ricavare applicando le (13.2) e ricordando le proprietà delle potenze.

$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y; \tag{13.9}$$
$$\cosh(x-y) = \cosh x \cosh y - \sinh x \sinh y; \tag{13.10}$$
$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y; \tag{13.11}$$
$$\sinh(x-y) = \sinh x \cosh y - \cosh x \sinh y. \tag{13.12}$$

Formule di duplicazione

Dalle formule di addizione, ponendo $x = y$ si ottiene:

$$\cosh 2x = \cosh^2 x + \sinh^2 x = 2\cosh^2 - 1 = 2\sinh^2 + 1; \tag{13.13}$$
$$\sinh 2x = 2\sinh x \cosh x. \tag{13.14}$$

Formule di prostaferesi

Per completare il quadro delle similitudini tra le funzioni trigonometriche e quelle iperboliche, riportiamo anche le formule di prostaferesi.

$$\cosh x + \cosh y = 2\cosh\frac{x+y}{2}\cosh\frac{x-y}{2}; \tag{13.15}$$
$$\cosh x - \cosh y = -2\sinh\frac{x+y}{2}\sinh\frac{x-y}{2}; \tag{13.16}$$
$$\cosh x + \cosh y = 2\sinh\frac{x+y}{2}\cosh\frac{x-y}{2}; \tag{13.17}$$
$$\cosh x - \cosh y = 2\cosh\frac{x+y}{2}\sinh\frac{x-y}{2}. \tag{13.18}$$

14. I numeri complessi

In questo capitolo proponiamo una introduzione, a livello elementare, del corpo dei complessi, con lo scopo principale di fissare i concetti fondamentali. Abbiamo intenzionalmente utilizzato uno schema semplificato, partendo da una definizione “operativa”, per poi giungere alla sistemazione formale dei concetti.

14.1. Perché i numeri complessi?

I numeri reali vengono introdotti per risolvere alcuni problemi legati alle misure di grandezze altrimenti intrattabili: il più famoso (e storicamente importante) è quello della misura della diagonale di un quadrato il cui lato sia assunto come unità. La possibilità di risolvere i problemi di misura è legata ad una proprietà cruciale dell’insieme dei reali, proprietà che non si ritrova nell’insieme dei numeri razionali, la cosiddetta proprietà dell’estremo superiore: ogni insieme non vuoto e superiormente limitato di reali o ha massimo o è tale che l’insieme dei suoi maggioranti ha minimo. L’introduzione di questo insieme di numeri può essere considerata conclusiva anche dal punto di vista della struttura algebrica: si tratta di un insieme in cui le operazioni di somma e prodotto godono di tutte le proprietà desiderabili (associativa, esistenza dell’elemento neutro, esistenza del simmetrico, distributiva del prodotto rispetto alla somma, commutativa), proprietà del resto già godute dall’insieme dei razionali.

Purtroppo rimane un grosso problema, insolubile anche in un insieme così raffinato come è quello dei reali, e precisamente quello della risoluzione di certe equazioni di grado superiore al primo. L’esempio più semplice è quello dell’equazione $x^2+1=0$ che non ha soluzioni, pur essendo molto simile all’equazione $x^2-1=0$, che invece ha due soluzioni. In sostanza è abbastanza “fastidioso” che certe equazioni abbiano un numero di soluzioni (tenendo conto della molteplicità) pari al loro grado, mentre altre, anche non molto diverse, non hanno alcuna soluzione, o ne hanno meno di quanto sia il loro grado. La cosa non è limitata alle equazioni di secondo grado. Un esempio può essere l’equazione $x^3+2x^2+3x+2=0$ che ha solo la soluzione $x=-1$, mentre l’equazione $x^3+2x^2-x-2=0$ ha tre soluzioni distinte e precisamente $x=-1$, $x=1$, $x=-2$.

Questo ci spinge ad ampliare l’insieme dei reali introducendo un nuovo insieme di numeri che, ed è un vero colpo di fortuna, non risolverà solo questo problema, ma presenterà una serie di proprietà tali da renderlo adatto a trattare un grande numero di situazioni nei più disparati campi della matematica e delle sue applicazioni.

14.2. Un’introduzione informale

Il problema che ci spinge a introdurre un nuovo insieme di numeri è, dunque, quello della risoluzione di certe equazioni. Cominciamo dalla più semplice: $x^2+1=0$. Ci si accorge immediatamente che l’impossibilità di risolvere questa equazione è legata al fatto che non si può estrarre la radice quadrata di -1: non esiste alcun numero reale il cui quadrato sia -1. Se il problema è questo, proviamo ad affrontarlo "inventando", anzi "immaginando", un nuovo oggetto che indichiamo con "i", e chiamiamo *unità immaginaria*, con la proprietà che il suo quadrato sia il numero reale -1. Poniamo cioè, per definizione,

$$\mathrm{i}^2 = -1.$$

È chiaro che la cosa, detta in questi termini, ha poco (o nessun) senso. Per intanto però andiamo avanti ugualmente cercando di fare tutte le operazioni che siamo abituati a fare con i reali, trattando questo nuovo oggetto come se fosse una lettera in un'espressione algebrica, con in più la caratteristica speciale che il suo quadrato vale -1. Naturalmente vogliamo che anche tutte le solite proprietà delle operazioni siano verificate.

Poiché ora vogliamo procedere in modo assolutamente informale, proviamo a fare subito alcuni calcoli, via via più complessi, e vediamo che cosa succede. Dobbiamo procedere come nel calcolo letterale, solo aggiungendo la nuova "regola" che abbiamo introdotto ad hoc.

- $\mathrm{i}^3 = \mathrm{i}^2 \cdot \mathrm{i} = (-1)\mathrm{i} = -\mathrm{i}$.
- $\mathrm{i}^4 = \mathrm{i}^2 \cdot \mathrm{i}^2 = (-1)(-1) = 1$.
- $(1+\mathrm{i})(1-\mathrm{i}) = 1 - \mathrm{i}^2 = 1 - (-1) = 2$.
- $(1-2\mathrm{i})(3+\mathrm{i}) = 3 + \mathrm{i} - 6\mathrm{i} - 2\mathrm{i}^2 = 3 + \mathrm{i} - 6\mathrm{i} - 2(.1) = 5 - 5\mathrm{i}$.
- $(2+\mathrm{i})^3 = 8 + 12\mathrm{i} + 6\mathrm{i}^2 + \mathrm{i}^3 = 8 + 12\mathrm{i} + 6(-1) + (-\mathrm{i}) = 2 + 11\mathrm{i}$.
- $\dfrac{2+\mathrm{i}}{3+2\mathrm{i}} = \dfrac{(2+\mathrm{i})(3-2\mathrm{i})}{(3+2\mathrm{i})(3-2\mathrm{i})} = \dfrac{6-4\mathrm{i}+3\mathrm{i}-2\mathrm{i}^2}{9-4\mathrm{i}^2} = \dfrac{8-\mathrm{i}}{13} = \dfrac{8}{13} - \dfrac{1}{13}\mathrm{i}$.
- $1 + 2\mathrm{i} + \dfrac{\mathrm{i}}{2-\mathrm{i}} = 1 + 2\mathrm{i} + \dfrac{\mathrm{i}(2+\mathrm{i})}{(2-\mathrm{i})(2+\mathrm{i})} = 1 + 2\mathrm{i} + \dfrac{-1+2\mathrm{i}}{5} = \dfrac{4}{5} + \dfrac{12}{5}\mathrm{i}$.

Come si vede dagli esempi, semplificando opportunamente le espressioni date che coinvolgono addizioni, moltiplicazioni, divisioni, si riesce sempre a scrivere il risultato nella forma $a + b\mathrm{i}$ dove a e b sono numeri reali. In sostanza operando su oggetti del tipo $a + b\mathrm{i}$ si ottengono ancora oggetti dello stesso tipo. Ma torniamo un momentino indietro: eravamo partiti con l'idea di risolvere equazioni e ci eravamo accorti che, almeno nel caso semplice considerato ($x^2 + 1 = 0$), le difficoltà incontrate erano legate all'impossibilità di calcolare le radici dei numeri negativi. Nel caso di -1 il problema l'abbiamo risolto con uno stratagemma; nel caso degli altri negativi, che cosa succede? Lasciamoci guidare dalla stessa idea di prima: quello che chiediamo è che, a parte il fatto nuovo che i^2 deve fare -1, tutto il resto deve rimanere invariato. Allora per calcolare la radice, per esempio, di -5 potremo procedere così:

- $\sqrt{-5} = \sqrt{(-1) \times 5} = \sqrt{-1}\sqrt{5} = \sqrt{5}\,\mathrm{i}$.

Dunque tutto funziona anche per gli altri numeri negativi. Attenzione, però: qui, come già nei reali, c'è anche un altro numero(1) che elevato al quadrato dà -5 e precisamente $-\sqrt{5}\,\mathrm{i}$, ma questo è un altro problema, su cui torneremo più avanti in maniera dettagliata. È proprio tutto a posto, almeno per quanto riguarda il calcolo delle radici quadrate? In realtà un piccolo sospetto ci dovrebbe venire: abbiamo introdotto una entità misteriosa per poter calcolare le radici dei numeri negativi, questo ci ha condotto

[1] Del resto anche di numeri che elevati al quadrato danno -1, se devono valere le usuali regole, ce ne sono due $\pm\mathrm{i}$.

a lavorare con oggetti di un nuovo tipo (che si scrivono nella forma $a+b\mathrm{i}$), ma ora possiamo calcolare le radici di questi nuovi oggetti? Se la risposta fosse negativa dovremmo chiaramente abbandonare l'idea che ci è venuta. Non ci resta quindi che provare, procedendo secondo lo schema già utilizzato sopra.

Consideriamo il numero $1+i$ e poniamoci la seguente domanda: esistono cose del tipo $x+\mathrm{i}y$ con valori opportuni di x e y tali che $(x+\mathrm{i}y)^2 = 1+\mathrm{i}$? Se oggetti di questo tipo ci sono, bisogna che $x^2+2\mathrm{i}xy-y^2 = 1+\mathrm{i}$. Invocando il "principio di identità dei polinomi" concludiamo, uguagliando i "termini noti" e i "coefficienti di i", che deve essere

$$\begin{cases} x^2-y^2=1 \\ 2xy=1 \end{cases} .$$

Risolvendo il sistema si trovano le due soluzioni

$$\sqrt{\frac{\sqrt{2}+1}{2}}+\mathrm{i}\sqrt{\frac{\sqrt{2}-1}{2}} \quad \text{e} \quad -\sqrt{\frac{\sqrt{2}+1}{2}}-\mathrm{i}\sqrt{\frac{\sqrt{2}-1}{2}}.$$

Come ulteriore esempio della potenza di questa idea, risolviamo l'equazione $x^3+1=0$, che equivale a trovare i numeri che elevati al cubo danno -1. Sappiamo già che il numero -1 stesso, ovvero $-1+0\mathrm{i}$, soddisfa questo requisito: vediamo se ce ne sono altri. Procedendo come prima si trovano le seguenti soluzioni

$$\frac{1}{2}\pm\frac{\sqrt{3}}{2}\mathrm{i}.$$

Il tutto si fa veramente interessante: il risultato dell'ultimo esempio equivale a dire che l'equazione $x^3+1=0$ ha tre soluzioni distinte (cioè tante quant'è il suo grado) e questo era proprio il nostro problema di partenza! Naturalmente è presto per concludere che tutto funziona, ma perlomeno siamo sulla strada giusta: l'idea di introdurre un nuovo oggetto con la strana e inusuale proprietà che il suo quadrato sia -1, comincia a piacerci proprio. Non possiamo però procedere oltre senza dare una giustificazione formalmente corretta a quanto abbiamo finora detto, ma, come vedremo, dopo questa "chiacchierata" non ci saranno grossi problemi tecnici (almeno all'inizio!).

Per intanto riassumiamo il contenuto di questo paragrafo: per costruire un nuovo insieme di numeri in cui sia possibile fare sempre le radici quadrate, abbiamo introdotto un nuovo simbolo, i, con la proprietà che il suo quadrato valga -1, e, di conseguenza, abbiamo imparato ad operare con oggetti del tipo $a+\mathrm{i}b$, dove a e b sono numeri reali. In sostanza l'idea che abbiamo avuto ci ha portato ad operare con un oggetto misterioso, i, e con coppie di numeri reali. È proprio partendo dalla considerazione di queste coppie che introdurremo in modo formalmente corretto i nuovi numeri. L'idea di utilizzare coppie di oggetti già noti l'avevamo già sfruttata per costruire i numeri razionali, partendo da "frazioni" con un numeratore e un denominatore.

14.3. Definizioni e proprietà

Nell'insieme $\mathbb{R}^2 = \{\,(a,b)\mid a\in\mathbb{R}\wedge b\in\mathbb{R}\,\}$ introduciamo due operazioni, che chiameremo addizione e moltiplicazione, e per le quali useremo gli abituali simboli, nel seguente modo

— $(a,b)+(c,d)=(a+c,b+d)$

— $(a,b)\cdot(c,d)=(ac-bd,ad+bc)$

La prima operazione è una “somma per componenti”, nella seconda non è difficile riconoscere che si tratta della formalizzazione dei calcoli informali che abbiamo fatto nel precedente paragrafo.

È un utile esercizio provare per calcolo diretto che valgono le usuali proprietà delle operazioni già viste in $\mathbb{R}$.

— Associativa.
  - $[(a,b)+(c,d)]+(e,f)=(a+c,b+d)+(e,f)=(a+c+e,b+d+f)$
    $(a,b)+[(c,d)+(e,f)]=(a,b)+(c+e,d+f)=(a+c+e,b+d+f)$.
  - $[(a,b)(c,d)](e,f)=(ac-bd,ad+bc)(e,f)=(ace-bde-adf-bcf,acf-bdf+ade+bce)$
    $(a,b)[(c,d)(e,f)]=(a,b)(ce-df,cf+de)=(ace-bde-adf-bcf,acf-bdf+ade+bce)$.

— Esistenza dell'elemento neutro.
  - $(a,b)+(0,0)=(0,0)+(a,b)=(a,b)$, ovvero $(0,0)$ è l'elemento neutro dell'addizione.
  - $(a,b)(1,0)=(1,0)(a,b)=(a,b)$, ovvero $(1,0)$ è l'elemento neutro della moltiplicazione.

— Esistenza del simmetrico.
  - $(a,b)+(-a,-b)=(0,0)$, ovvero $(-a,-b)$ è l'elemento simmetrico (opposto) di (a,b) nell'addizione.
  - Se $(a,b)\neq(0,0)$, $(a,b)\left(\dfrac{a}{a^2+b^2},\dfrac{b}{a^2+b^2}\right)=(1,0)$, ovvero $\left(\dfrac{a}{a^2+b^2},\dfrac{b}{a^2+b^2}\right)$ è l'elemento simmetrico (reciproco) di (a,b) nella moltiplicazione.

— Commutativa.
  - $(a+b)+(c+d)=(a+c,b+d)=(c,d)+(a,b)$.
  - $(a,b)(c,d)=(ac-bd,ad+bc)=(c,d)(a,b)$.

— Distributiva della moltiplicazione rispetto all'addizione (questa proprietà è una specie di “regola di convivenza” tra le due operazioni).
  - $(a,b)[(c,d)+(e,f)]=(a,b)(c+e,d+f)=(ac+ae-bd-bf,ad+af+bd+bf)$.
    $(a,b)[(c,d)+(e,f)]=(a,b)(c,d)+(a,b)(e,f)=(ac-bd,ad+bc)+(ae-bf,af+be)=(ac+ae-bd-bf,ad+af+bd+bf)$.

Se valgono tutte queste proprietà siamo autorizzati a chiamare *numeri* questi nuovi oggetti, cioè questa nuova struttura algebrica costituita da $\mathbb{R}^2$ come supporto e con le due operazioni indicate: per distinguerli dagli altri numeri che già conosciamo li chiameremo *numeri complessi*. L'insieme $\mathbb{R}^2$, munito di questa struttura algebrica, sarà indicato con $\mathbb{C}$.

Si vede facilmente che i numeri del tipo $(a,0)$ se sommati o moltiplicati tra di loro rimangono sempre dello stesso tipo (cioè con seconda coordinata nulla); anche il reciproco e l'opposto di un numero di questo tipo è ancora un numero di questo tipo. In sostanza il sottoinsieme di $\mathbb{C}$ costituito dai numeri del tipo (a,o) è chiuso rispetto alle operazioni di addizione e moltiplicazione e, come è facile vedere, gode di tutte le proprietà dell'insieme dei reali. Questo ci autorizza ad identificare i nuovi numeri del tipo

$(a,0)$ con i vecchi numeri reali, tanto che scriveremo semplicemente a al posto di $(a,0)$. È la stessa cosa che si fa, per esempio, con i razionali: le frazioni $^2/_1$, oppure $^4/_2$, $^6/_3$, ecc. sono formalmente diverse dal vecchio numero razionale intero 2, ma si comportano esattamente allo stesso modo, tanto che scriviamo sempre 2 e a nessuno viene in mente di scrivere $^2/_1$: sarebbe solo una perdita di tempo. Dunque possiamo ritenere che l'insieme $\mathbb{C}$ sia un ampliamento di $\mathbb{R}$.

Osserviamo ora che $(a,b)=(a,0)+(0,1)(b,0)$, cosa che possiamo anche scrivere, tenuto conto di quanto appena osservato, come $(a,b)=a+(0,1)b$. Con questo trucco possiamo rappresentare i nuovi numeri come somma di uno dei vecchi numeri più uno dei vecchi numeri moltiplicato sempre per la coppia $(0,1)$: siccome questa coppia compare sempre, tanto vale usare un simbolo speciale per rappresentarla, per non appesantire la scrittura. Il simbolo giusto ci viene suggerito da questo calcolo: $(0,1)^2=(0,1)(0,1)=(-1,0)=-1$. Il numero $(0,1)$ è proprio l'oggetto misterioso che, nella introduzione informale, avevamo introdotto come un artificio per risolvere l'equazione $x^2+1=0$. Quindi sarà logico porre, per definizione,

$$(0,1)=\mathrm{i}. \tag{14.1}$$

Il numero $(0,1)=\mathrm{i}$ si chiama *unità immaginaria*. Questa posizione ci consente di scrivere il numero complesso (a,b) nella forma

$$(a,b)=a+\mathrm{i}b\,, \tag{14.2}$$

detta *forma algebrica* (in inglese *rectangular form*) del numero complesso. È tradizione indicare i numeri complessi con z. Nella scrittura $z=a+\mathrm{i}b$, a si chiama *parte reale* del numero complesso z, e viene indicata con $\operatorname{Re} z$, mentre b si chiama *parte immaginaria* (a volte *coefficiente dell'immaginario*) del numero complesso z, e viene indicata con $\operatorname{Im} z$.

La scrittura (14.2) e il fatto che $(0,1)^2=\mathrm{i}^2=-1$ ci consente di eseguire i calcoli algebrici con i complessi come indicato nell'introduzione informale 14.2, dimenticando le regole di calcolo sopra indicate: almeno per le operazioni di addizione e moltiplicazione si può operare esattamente come con i reali, con la condizione aggiuntiva che $\mathrm{i}^2=-1$.

14.4. Proprietà dei complessi

14.4.1. Complessi e ordine

L'insieme dei numeri complessi è dunque un ampliamento dei reali, di cui mantiene alcune caratteristiche molto importanti, quali la struttura algebrica. C'è però una sostanziale differenza, che apparirà naturale quando parleremo del piano di Gauss: in questo insieme non è possibile introdurre un ordine che abbia un qualche interesse.

Ricordiamo che su ogni insieme è possibile introdurre addirittura un buon ordinamento, ovvero un ordinamento simile a quello che c'è sui naturali (ogni insieme non vuoto e inferiormente limitato ha un minimo), ma in generale perché un ordinamento su un insieme con struttura abbia interesse occorre che esso sia compatibile con le operazioni della struttura stessa. Quello che ci interessa è che l'insieme, diciamolo K, sia un corpo (commutativo) ordinato, ovvero che succeda quanto segue.

1. Siano definite due operazioni (addizione e moltiplicazione) dotate di

a) proprietà associativa;
b) esistenza dell'elemento neutro dell'addizione, 0, e della moltiplicazione , 1,
c) esistenza dell'opposto e del reciproco (tranne il reciproco dell'elemento neutro dell'addizione);
d) proprietà commutativa;
e) proprietà distributiva della moltiplicazione rispetto all'addizione.

2. Sull'insieme sia definito un ordine totale (due elementi sono sempre confrontabili).
3. Se definiamo $K^+ = \{ x \in K \mid x > 0 \}$, valgano le seguenti proprietà:

a) $x, y \in K^+ \Rightarrow x + y \in K^+ \wedge xy \in K^+$;
b) $x \neq 0 \Rightarrow x \in K^+ \vee -x \in K^+$;
c) $0 \notin K^+$;
d) $x \leq y \Leftrightarrow (y - x = 0 \vee y - x \in K^+)$.

Gli elementi di K^+ si dicono positivi, gli altri, escluso lo 0, negativi.

Da qui si possono dedurre alcune ulteriori proprietà valide in ogni corpo (commutativo) ordinato.

— Se $k > 0$, allora $-k < 0$, altrimenti $k + (-k) > 0$, in contrasto con la 3c.
— Per ogni $k \in K, \neq 0$ si ha $k^2 > 0$. Infatti se $k > 0$, a maggior ragione $k^2 > 0$ (proprietà 3a), se invece $k < 0$, allora $-k > 0$ e quindi $(-k)^2 = k^2 > 0$. Ne segue che $1 = (1)^2$ è positivo e quindi che $-1 < 0$.

Tenendo conto di queste ultime due proprietà, possiamo affermare che $\mathbb{C}$ non può essere un corpo ordinato. Se infatti consideriamo il numero i, si deve avere o $\mathrm{i} > 0$ o $-\mathrm{i} > 0$, ma allora $\mathrm{i}^2 > 0$, mentre sappiamo che $\mathrm{i}^2 = -1$ e dunque, come abbiamo appena provato, $\mathrm{i}^2 < 0$.

Dunque non esiste possibilità di introdurre nell'insieme dei complessi una qualche forma di ordine: dati due complessi qualunque potremo solo dire se essi sono uguali o diversi, mentre non avremo alcuna possibilità di decidere quale dei due è il più grande né, tantomeno, di decidere se un complesso è positivo o negativo.

14.4.2. Il modulo di un numero complesso

Dato un numero complesso $z = a + \mathrm{i}b$, chiameremo suo *modulo* il numero reale positivo

$$\varrho = |z| = \sqrt{a^2 + b^2}. \tag{14.3}$$

Per il modulo valgono proprietà analoghe a quelle già viste per i numeri reali. In particolare segnaliamo che:

— $|z + w| \leq |z| + |w|$;
— $|zw| = |z||w|$.

La dimostrazione costituisce un utile esercizio. Poniamo $z = a + \mathrm{i}b$ e $w = c + \mathrm{i}d$.

$$|z + w| \leq |z| + |w| \Leftrightarrow \sqrt{(a + c)^2 + (b + d)^2} \leq \sqrt{a^2 + b^2} + \sqrt{c^2 + d^2} \Leftrightarrow$$
$$a^2 + c^2 + 2ac + b^2 + d^2 + 2bd \leq a^2 + b^2 + c^2 + d^2 + 2\sqrt{a^2 + b^2}\sqrt{c^2 + d^2} \Leftrightarrow$$

$$ac+bd \le \sqrt{a^2+b^2}\sqrt{c^2+d^2} \iff a^2c^2+b^2d^2+2abcd \le a^2c^2+a^2d^2+b^2c^2+b^2d^2 \iff 0 \le (ad-bc)^2.$$

Questo basta per concludere, visto che l'ultima disuguaglianza è vera.

$$|zw|=|z||w| \iff \sqrt{(ac-bd)^2+(ad+bc)^2}=\sqrt{a^2+b^2}\sqrt{c^2+d^2} \iff a^2c^2+b^2d^2-2acbd+a^2d^2+b^2c^2+2adbc=a^2c^2+a^2d^2+b^2c^2+b^2d^2.$$

Anche questo basta per concludere, visto che l'ultima uguaglianza è ovvia.

È importante notare alcune differenze operative tra il modulo nei reali e quello nei complessi. Esaminiamo alcuni esempi.

— $|z|^2 \neq z^2$. Il primo infatti è un numero reale positivo, il secondo no: $|z|^2 = a^2+b^2$, $z^2 = a^2-b^2+2\mathrm{i}ab$.
— $|z|^2 = |z^2|$. Il primo membro vale a^2+b^2; per il secondo si ha

$$\left|z^2\right| = \sqrt{(a^2-b^2)^2+4a^2b^2} = \sqrt{(a^2+b^2)^2} = a^2+b^2,$$

da cui l'uguaglianza.
— $|z| \neq \sqrt{z^2}$. Per convincersi si può provare prendendo $z=\mathrm{i}$: al primo membro si ottiene 1, al secondo membro $\pm\mathrm{i}$. Questa osservazione è particolarmente importante, in quanto nei reali il modulo di un numero x può anche essere definito mediante

$$|x| = \sqrt{x^2} \quad (x \in \mathbb{R}).$$

Si noti anche che abbiamo scritto che il secondo membro vale $\pm\mathrm{i}$. La cosa può apparire strana, in quanto nei reali con il simbolo $\sqrt{a}$ si indica un solo numero, se $a>0$, non due. La convenzione è diversa sui complessi, come vedremo.

14.4.3. Il coniugato di un numero complesso

Dato un numero complesso $z=a+\mathrm{i}b$, si chiama suo *coniugato* il numero

$$\overline{z} = a - \mathrm{i}b. \tag{14.4}$$

L'operazione che fa passare da un complesso al suo coniugato gode di alcune interessanti proprietà che si possono verificare per esercizio.

— $\overline{z+w} = \overline{z}+\overline{w}$.
— $\overline{z\cdot w} = \overline{z}\cdot\overline{w}$.
— $\overline{\overline{z}} = z$.
— $z\cdot\overline{z} = |z|^2$.
— $z=\overline{z} \iff z\in\mathbb{R}$.
— $z+\overline{z}\in\mathbb{R}$, $\forall z\in\mathbb{C}$.

14.5. Forma trigonometrica

14.5.1. Il piano di Gauss

Ancor più che nel caso dei numeri reali, una opportuna rappresentazione grafica dei complessi è cruciale per trattare i problemi che li coinvolgono. Il modo con cui questi numeri sono stati introdotti porta immediatamente a concludere che la rappresentazione grafica "naturale" è quella in un piano cartesiano, che chiameremo *piano di Gauss* o anche di *Argand-Gauss*. L'asse delle ascisse si chiama anche *asse reale* mentre quello delle ordinate *asse immaginario*.

Tenendo conto di come è stata definita la somma, i numeri complessi potranno essere anche pensati come vettori di questo piano (con l'origine in O): la somma si fa esattamente come la somma di vettori, con la regola del parallelogramma. Si veda la figura 14.1.

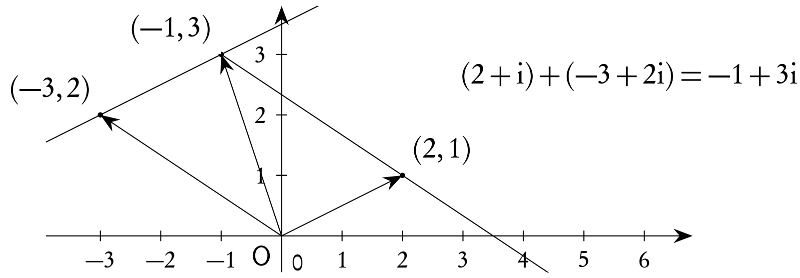


Figura 14.1.: *Somma di complessi nel piano di Gauss*

14.5.2. Forma trigonometrica o polare

Più difficile l'interpretazione grafica del prodotto, per la quale conviene introdurre una diversa scrittura dei numeri complessi. Questa scrittura sarà utile anche per altri scopi.

Per questo rappresentiamo il numero complesso $z = x + \mathrm{i}y$ nel piano di Gauss e consideriamone il modulo (che non è altro che la distanza del punto (x, y) dall'origine). Consideriamo poi la semiretta orientata uscente dall'origine e passante per z. Si chiama *argomento principale* o *anomalia* di z, e si indica con $\arg z$, l'angolo ϑ che questa semiretta forma con il semiasse positivo delle ascisse, orientato in senso antiorario e appartenente all'intervallo $[0°, 360°[$, o $[0, 2\pi[$ se si usano, come è preferibile, i radianti. La convenzione può essere anche diversa: $[0°, 360°[$ e $[0, 2\pi[$ sono sostituiti con $]-180°, 180°]$ e $]-\pi, \pi]$. È chiaro che la rappresentazione di un numero complesso $x + \mathrm{i}y$ mediante una coppia (modulo,argomento principale) è univoca, con la sola eccezione del punto 0, per il quale non risulta definito un argomento principale. In molte questioni è utile rimuovere le limitazioni poste per l'argomento principale, ammettendo anche valori dell'angolo superiori o inferiori a quelli dell'argomento principale: in questo caso si parla di solito di argomento senza l'aggettivo principale ed è ovvio che la rappresentazione non è più univoca. Questo però non porta complicazioni particolari nel trattare i complessi e anzi facilita molti calcoli.

Dunque un numero complesso può essere rappresentato con la coppia (modulo,argomento), anziché con la coppia (parte reale,parte immaginaria): nel primo caso si avrà la *forma trigonometrica*, detta a volte anche *polare*, nel secondo la ormai tradizionale forma algebrica. Alcuni software, per distinguere le due situazioni, usano la notazione tradizionale di coppia con le parentesi tonde e la virgola, (x, y), per la forma algebrica, notazioni con le parentesi tonde ma il punto e virgola o i due punti come separatore: $(\varrho;\vartheta)$ oppure $(\varrho:\vartheta)$. Di solito non ci sono problemi di confusione e quindi useremo la notazione tradizionale anche per la coppia (modulo,argomento): la differenza risulterà chiara dal contesto.

Si noti che due numeri complessi $z_1 = x_1 + \mathrm{i}y_1 = (\varrho_1, \vartheta_1)$ e $z_2 = x_2 + \mathrm{i}y_2 = (\varrho_2, \vartheta_2)$ sono uguali se e solo se

$$\begin{cases} x_1 = x_2 \\ y_1 = y_2 \end{cases}, \quad \text{oppure} \quad \begin{cases} \varrho_1 = \varrho_2 \\ \vartheta_1 = \vartheta_2 + 2k\pi \end{cases}, \tag{14.5}$$

a seconda che i numeri siano in forma algebrica o polare.

14.5.3. Passaggio da una forma all'altra

Il passaggio dalla forma polare a quella algebrica si fa in modo banale, tenendo conto della definizione delle funzioni seno e coseno:

$$\begin{cases} x = \varrho\cos\vartheta \\ y = \varrho\sin\vartheta \end{cases}. \tag{14.6}$$

Da qui si deduce la seguente scrittura

$$z = x + \mathrm{i}y = \varrho(\cos\vartheta + \mathrm{i}\sin\vartheta). \tag{14.7}$$

Quasi sempre è proprio l'ultima delle (14.7) che prende il nome di *forma trigonometrica*.

Il passaggio inverso è immediato per quanto riguarda il modulo:

$$\varrho = |z| = \sqrt{a^2 + b^2}.$$

Per la ricerca dell'argomento basta utilizzare le formule inverse delle 14.6:

$$\begin{cases} \cos\vartheta = \dfrac{x}{\sqrt{x^2+y^2}} \\ \sin\vartheta = \dfrac{y}{\sqrt{x^2+y^2}} \end{cases}. \tag{14.8}$$

Si presti attenzione a determinare preventivamente il quadrante in cui si trova ϑ, esaminando il segno del suo coseno e del suo seno, dopodiché mediante una delle due formule si potrà trovare l'argomento.

14.5.4. Prodotto e quoziente in forma trigonometrica

Dati due complessi $z = x + \mathrm{i}y = \varrho(\cos\vartheta + \mathrm{i}\sin\vartheta)$ e $w = s + \mathrm{i}t = r(\cos\varphi + \mathrm{i}\sin\varphi)$, calcoliamone il prodotto e il quoziente, considerando la forma trigonometrica.

$$z \cdot w = \varrho r\left[(\cos\vartheta\cos\varphi - \sin\vartheta\sin\varphi) + \mathrm{i}(\cos\vartheta\sin\varphi + \sin\vartheta\cos\varphi)\right] = \varrho r\left[\cos(\vartheta+\varphi) + \mathrm{i}\sin(\vartheta+\varphi)\right].$$

$$\frac{z}{w} = \frac{\varrho}{r}\,\frac{\cos\vartheta + \mathrm{i}\sin\vartheta}{\cos\varphi + \mathrm{i}\sin\varphi} = \frac{\varrho}{r}\,\frac{(\cos\vartheta + \mathrm{i}\sin\vartheta)(\cos\varphi - \mathrm{i}\sin\varphi)}{(\cos\varphi + \mathrm{i}\sin\varphi)(\cos\varphi - \mathrm{i}\sin\varphi)} = \cdots = \frac{\varrho}{r}\left[\cos(\vartheta-\varphi) + \mathrm{i}\sin(\vartheta-\varphi)\right].$$

Questo ci autorizza a scrivere le seguenti importanti formule

$$(\varrho_1,\vartheta_1)\cdot(\varrho_2,\vartheta_2)=(\varrho_1\varrho_2,\vartheta_1+\vartheta_2) \quad , \quad \frac{(\varrho_1,\vartheta_1)}{(\varrho_2,\vartheta_2)}=\left(\frac{\varrho_1}{\varrho_2},\vartheta_1-\vartheta_2\right), \tag{14.9}$$

che si possono enunciare a parole come segue: il modulo del prodotto e del quoziente di due numeri complessi è uguale rispettivamente al prodotto e quoziente dei moduli, mentre l'argomento del prodotto è la somma degli argomenti, l'argomento del quoziente è la differenza degli argomenti.

Tenendo conto della formula del binomio di Newton e della prima delle (14.9), possiamo ora scrivere l'espressione per la potenza n-esima di un complesso, sia scritto in forma algebrica che in forma trigonometrica.

$$z^n:\begin{cases} (x+\mathrm{i}y)^n=\binom{n}{0}x^n+\binom{n}{1}x^{n-1}(\mathrm{i}y)+\cdots+\binom{n}{n}(\mathrm{i}y)^n \\ (\varrho,\vartheta)^n=(\varrho^n,n\vartheta), \quad \text{oppure} \quad \big(\varrho(\cos\vartheta+\mathrm{i}\sin\vartheta)\big)^n=\varrho^n(\cos n\vartheta+\mathrm{i}\sin n\vartheta)\end{cases}. \tag{14.10}$$

La seconda formula, detta *formula di de Moivre*, è particolarmente semplice. La sua dimostrazione richiede anche l'uso del principio di induzione, che però esula dagli scopi di questo testo. Essa è inoltre valida per ogni valore intero di n, anche non positivo. Infatti se $n=0$ diventa banale, se $n<0$ si ha

$$(\cos\vartheta+\mathrm{i}\sin\vartheta)^n=\frac{1}{(\cos\vartheta+\mathrm{i}\sin\vartheta)^{-n}}=\frac{1}{\cos(-n\vartheta)+\mathrm{i}\sin(-n\vartheta)}==\frac{1}{\cos(n\vartheta)-\mathrm{i}\sin(n\vartheta)}=$$

$$=\frac{1}{\cos(n\vartheta)-\mathrm{i}\sin(n\vartheta)}\,\frac{\cos(n\vartheta)+\mathrm{i}\sin(n\vartheta)}{\cos(n\vartheta)+\mathrm{i}\sin(n\vartheta)}=\frac{\cos(n\vartheta)+\mathrm{i}\sin(n\vartheta)}{\cos^2(n\vartheta)+\sin^2(n\vartheta)}=\cos(n\vartheta)+\mathrm{i}\sin(n\vartheta).$$

Applicando la formula di de Moivre e la formula del binomio di Newton al calcolo della potenza n-esima di un complesso scritto in forma trigonometrica si possono calcolare facilmente $\cos(n\vartheta)$ e $\sin(n\vartheta)$ in funzione di $\cos\vartheta$ e $\sin\vartheta$. Per esempio per $n=3$ si ha

$$\begin{aligned}(\cos\vartheta+\mathrm{i}\sin\vartheta)^3&=\cos(3\vartheta)+\mathrm{i}\sin(3\vartheta)=\\ &=\binom{3}{0}\cos^3\vartheta+\binom{3}{1}\cos^2\vartheta(\mathrm{i}\sin\vartheta)+\binom{3}{2}\cos\vartheta\,\mathrm{i}^2\sin^2\vartheta+\binom{3}{3}\mathrm{i}^3\sin^3\vartheta=\\ &=\cos^3\vartheta-3\cos\vartheta\sin^2\vartheta+\mathrm{i}(3\cos^2\vartheta\sin\vartheta-\sin^3\vartheta).\end{aligned}$$

Da qui, uguagliando le parti reale e immaginaria, si ottiene

$$\begin{aligned}\cos(3\vartheta)&=\cos^3\vartheta-3\cos\vartheta\sin^2\vartheta=4\cos^3\vartheta-3\cos\vartheta,\\ \sin(3\vartheta)&=3\cos^2\vartheta\sin\vartheta-\sin^3\vartheta=3\sin\vartheta-4\sin^3\vartheta,\end{aligned}$$

in accordo con le (11.32) e (11.33).

14.5.5. Prodotto e rotazioni nel piano di Gauss

La formula del prodotto di due complessi in forma trigonometrica consente di dare una semplice interpretazione geometrica al prodotto stesso: il prodotto tra il numero complesso $z = x + iy = \varrho(\cos\vartheta + i\sin\vartheta)$ e il numero complesso $w = s + it = r(\cos\varphi + i\sin\varphi)$ può essere interpretato come una dilatazione del vettore che rappresenta z (dilatazione del fattore $r = |w|$) seguita da una sua rotazione di un angolo φ. Se, in particolare, w ha modulo 1 si tratta di una semplice rotazione. Molto importante il caso in cui $w = i$, nel qual caso si tratta di una rotazione di 90°. Si veda la figura 14.2.

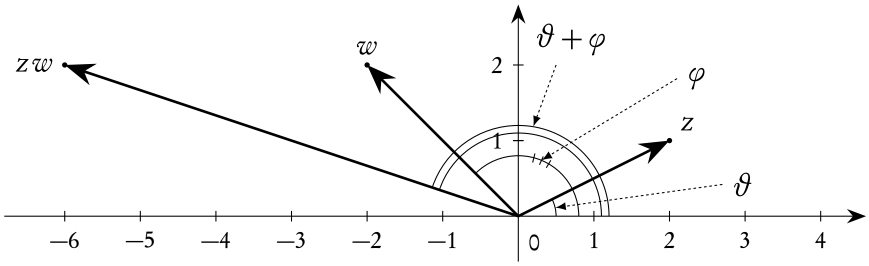


Figura 14.2.: *Prodotto di complessi e rotazioni del piano*

Questa proprietà dei complessi introduce un importante legame tra questi numeri e le trasformazioni geometriche del piano ed ha numerose applicazioni nella fisica e in particolare nell'elettrotecnica (studio di circuiti in corrente alternata).

14.6. Radici nei complessi

La formula di de Moivre permette di dimostrare facilmente il seguente teorema.

Teorema 14.1. *Per ogni numero complesso $\alpha = a + ib = (\varrho, \vartheta)$ esistono esattamente n radici n-esime distinte, ovvero n numeri complessi diversi, che indichiamo con $z_0, z_1, \dots, z_{n-1}$ tali che $(z_k)^n = \alpha$. Vale la seguente formula.*

$$z_k = \left(\sqrt[n]{\varrho}, \frac{\vartheta + 2k\pi}{n}\right) = \sqrt[n]{\varrho}\left(\cos\frac{\vartheta + 2k\pi}{n} + i\sin\frac{\vartheta + 2k\pi}{n}\right), \quad k \in \{0, 1, 2, \dots, n-1\}. \tag{14.11}$$

Dimostrazione. È sufficiente osservare che, posto $z = (r, \varphi)$, si deve avere $(r^n, n\varphi) = (\varrho, \vartheta)$, cioè

$$\begin{cases} r^n = \varrho \\ n\varphi = \vartheta + 2k\pi \end{cases}, \quad \text{ovvero} \quad \begin{cases} r = \sqrt[n]{\varrho} \\ \varphi = \dfrac{\vartheta + 2k\pi}{n} \end{cases}.$$

Esistono solo n valori distinti di k che producono numeri complessi distinti (per la consueta periodicità delle funzioni trigonometriche): si possono scegliere i valori da 0 ad $n - 1$, come abbiamo fatto sopra e come si fa d'abitudine. □

Come è noto, nell'insieme dei numeri reali si verifica che:

— per ogni numero a esiste una ed una sola radice di indice dispari, ossia uno ed un solo numero reale x tale che, se n è dispari, $x^n = a$; questa radice si indica con $\sqrt[n]{a}$;
— per ogni numero reale $a > 0$ esistono due radici distinte di indice pari, ossia due numeri x tali che, se n è pari, $x^n = a$; le due radici sono tra di loro opposte (quindi una positiva e una negativa) e quella positiva si indica con $\sqrt[n]{a}$, mentre quella negativa con $-\sqrt[n]{a}$;
— se $a < 0$, non esiste alcun numero reale x tale che, se n è pari, $x^n = a$;
— se $a = 0$, l'unico numero reale x tale che $x^n = 0$ è 0 stesso (zero ha una sola radice, sia pari che dispari).

In sostanza il simbolo $\sqrt[n]{a}$ indica un unico numero reale e precisamente l'unica radice di a se n è dispari (radice che può essere sia positiva che negativa), mentre indica solo la radice positiva di a se n è pari. La scrittura nei reali è facilitata dal fatto che, al massimo, per un numero reale esistono due radici.

Tutto si complica nei complessi, dove le radici n-esime sono n. Inoltre non essendo possibile definire il concetto di positivo e negativo non è facile stabilire un criterio per privilegiare (almeno nella scrittura) una delle radici sulle altre (nei reali abbiamo privilegiato la positiva sulla negativa).

Per questo, dato il numero complesso α, si è scelto di indicare con il simbolo $\sqrt[n]{\alpha}$ l'insieme di tutte le radici n-esime di α. Dunque:

$$\alpha = a + \mathrm{i}b = (\varrho, \vartheta) \quad \Rightarrow \quad \sqrt[n]{\alpha} = \left\{ \left(\sqrt[n]{\varrho}, \frac{\vartheta + 2k\pi}{n} \right), \, k \in \{0, 1, 2, \dots, n-1\} \right\}, \tag{14.12}$$

oppure anche

$$\sqrt[n]{\alpha} = \left\{ \sqrt[n]{\varrho} \left(\cos \frac{\vartheta + 2k\pi}{n} + \mathrm{i} \sin \frac{\vartheta + 2k\pi}{n} \right), \, k \in \{0, 1, 2, \dots, n-1\} \right\}. \tag{14.13}$$

Questa notazione può ingenerare confusione perché usa un simbolo, comunemente adottato per rappresentare un unico numero, per rappresentare invece un insieme di n numeri. Alcuni autori hanno introdotto qualche modifica per evitare questi problemi, per esempio scrivendo $\sqrt[n]{\alpha}\rceil$, ma questo tipo di simboli non ha avuto fortuna. Non resta che prestare attenzione: di solito tutto è chiarito dal contesto.

In questo spirito la formula risolutiva per le equazioni di secondo grado si può scrivere in maniera più compatta, evitando il doppio segno davanti alla radice, in quanto il simbolo di radice quadrata indica già (solo nei complessi!) due numeri diversi.

$$ax^2 + bx + c = 0 \Rightarrow x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad b^2 - 4ac \geq 0;$$
$$az^2 + bz + c = 0 \Rightarrow z_{1,2} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad \forall\, b^2 - 4ac.$$

Non è comunque il caso di sottilizzare troppo: in ogni caso anche le radici complesse sono una l'opposta dell'altra (anche se non sono una positiva e una negativa!).

Si noti anche che le usuali proprietà dei radicali in $\mathbb{R}$ vanno opportunamente modificate, di solito interpretandole nel senso della teoria degli insiemi. Vediamo alcuni esempi.

— $(\sqrt[n]{\alpha})^n$ è sempre uguale ad α, qualunque sia il valore di $\sqrt[n]{\alpha}$.

— Ha senso scrivere $\alpha \in \sqrt[n]{\alpha^n}$ e non $\alpha = \sqrt[n]{\alpha^n}$, in quanto α è un numero complesso, $\sqrt[n]{\alpha^n}$ è un insieme di numeri complessi. Per esempio

$$\sqrt{2^2} = \{-2, +2\}, \quad \sqrt[3]{2^3} = \left\{2, -\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}, -\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}\right\}.$$

Si noti ancora una volta che questa convenzione è diversa da quanto succede in $\mathbb{R}$.

— Ha senso scrivere $\sqrt[n]{\sqrt[m]{\alpha}} = \sqrt[nm]{\alpha}$ se si tiene conto che si tratta di un'uguaglianza tra due insiemi costituiti entrambi da $n \cdot m$ numeri.

— Ho senso scrivere $\sqrt[n]{\alpha} \cdot \sqrt[n]{\beta} = \sqrt[n]{\alpha\beta}$ se si interpreta correttamente l'uguaglianza. Nessun problema per il secondo membro che rappresenta un insieme di n numeri. Per quanto riguarda invece il primo membro osserviamo che $\sqrt[n]{\alpha}$ e $\sqrt[n]{\beta}$ sono due insiemi di n numeri ciascuno e si deve pensare il prodotto come l'insieme costituito dal prodotto di ciascun elemento del primo insieme per ciascun elemento del secondo insieme: anche se in teoria si ottengono n^2 numeri, in realtà solo n sono tra di loro distinti. Per esempio

$$\sqrt{4} = \{-2, 2\}, \quad \sqrt{9} = \{-3, 3\}, \quad \sqrt{4} \cdot \sqrt{9} = \{6, -6, -6, 6\} = \{-6, 6\} = \sqrt{36}.$$

A costo di sembrare prolissi, ripetiamo un'ultima volta che tutte queste scritture in $\mathbb{R}$ sono errate.

14.6.1. Radici dell'unità

Il caso $\alpha = 1$ è particolarmente interessante e merita un'analisi particolare.

Ora tutte le radici n-esime (qualunque sia n) hanno modulo 1 e quindi nel piano di Gauss si situano sulla circonferenza di centro l'origine e raggio 1. Inoltre, poiché 1^n vale ovviamente 1, il numero 1 stesso appartiene all'insieme delle radici n-esime, per ogni n. Se si osserva che gli argomenti delle radici sono, nell'ordine, 0, $2\pi/n$, $3\pi/n$, $\ldots$,$(n-1)\pi/n$, si conclude subito che esse si situano sui vertici di un poligono regolare di n lati, inscritto nella circonferenza e con un vertice nel punto $(1, 0)$. Questa osservazione riconduce il problema del calcolo esplicito delle radici in forma algebrica a quello generale della ciclotomia, ovvero della divisione della circonferenza in n parti uguali. Si veda la figura 14.3.

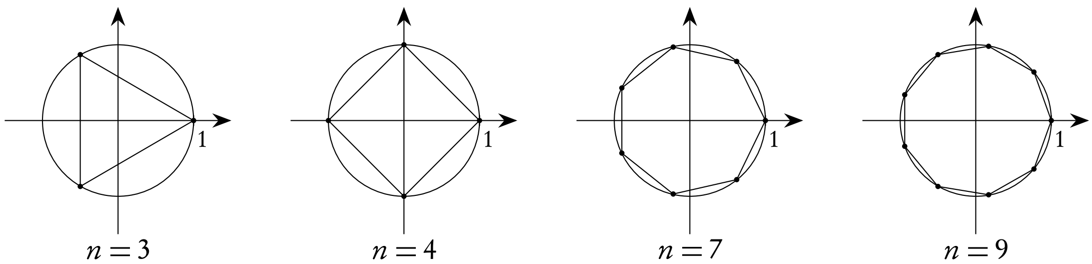


Figura 14.3.: *Radici n-esime dell'unità in $\mathbb{C}$ e ciclotomia*

Se consideriamo poi un complesso α qualunque, potremo scrivere $\sqrt[n]{\alpha} = \sqrt[n]{1 \cdot \alpha} = \sqrt[n]{1} \cdot \sqrt[n]{\alpha}$. Ci sono n radici n-esime sia di 1 che di α. Si deduce allora che per trovare le radici n-esime di α basterà trovarne una e poi moltiplicarla per le n radici n-esime di 1: se costruiamo una tabella contenente le radici n-esime di 1, ci sarà molto più facile trovare le radici degli altri numeri complessi.

14.6.2. Radici quadrate

Si tratta di un caso particolarmente importante nelle applicazioni, in quanto legato al problema della risoluzione di equazioni di secondo grado.

Nelle applicazioni, infatti, ha spesso interesse, oltre alla forma trigonometrica di un complesso, anche la forma algebrica. Osserviamo esplicitamente che la formula precedentemente trovata per il calcolo delle radici n-esime, permette un agevole calcolo dell'argomento di qualunque radice n-esima solo se si conosce l'argomento del numero di cui si vuole estrarre la radice. In generale il numero di cui trovare le radici viene dato nella sua forma algebrica, e quindi è possibile ricavare esplicitamente il seno ed il coseno dell'argomento, mentre l'argomento stesso deve venire espresso utilizzando le funzioni inverse arcsin e arccos, tranne alcuni casi favorevoli.

Inoltre, una volta trovati gli argomenti delle radici n-esime, il passaggio alla forma algebrica richiede il calcolo del seno e coseno di questi argomenti, cosa esplicitamente possibile solo in casi favorevoli.

Consideriamo qualche esempio per chiarire il concetto.

$$\begin{aligned}\sqrt[3]{-1}=\sqrt[3]{(1,\pi)}=\left\{\left(1,\,\frac{\pi+2k\pi}{3}\right),\,k\in\{\,0,1,2\,\}\right\}=\\ =\left\{\cos\frac{\pi+2k\pi}{3}+\mathrm{i}\sin\frac{\pi+2k\pi}{3},\,k\in\{\,0,1,2\,\}\right\},\end{aligned}$$

e i calcoli del coseno e seno necessari possono essere agevolmente eseguiti.

$$\begin{aligned}\sqrt[7]{-1}=\sqrt[7]{(1,\pi)}=\left\{\left(1,\,\frac{\pi+2k\pi}{7}\right),\,k\in\{\,0,1,2,3,4,5,6\,\}\right\}=\\ =\left\{\cos\frac{\pi+2k\pi}{7}+\mathrm{i}\sin\frac{\pi+2k\pi}{7},\,k\in\{\,0,1,2,3,4,5,6\,\}\right\},\end{aligned}$$

e questa volta i calcoli del coseno e seno necessari non possono essere eseguiti, in quanto non è possibile dividere con riga e compasso in sette parti la circonferenza.

$$\begin{aligned}\sqrt[3]{\frac{1}{2}+\mathrm{i}\frac{\sqrt{3}}{2}}=\sqrt[3]{\left(1,\,\frac{\pi}{3}\right)}=\left\{\left(1,\,\frac{\pi/3+2k\pi}{3}\right),\,k\in\{\,0,1,1\,\}\right\}=\\ =\left\{\cos\frac{\pi/3+2k\pi}{3}+\mathrm{i}\sin\frac{\pi/3+2k\pi}{3},\,k\in\{\,0,1,2\,\}\right\},\end{aligned}$$

e i calcoli del coseno e seno necessari non possono essere eseguiti mediante radicali quadratici, in quanto l'angolo di $\pi/3$ non può essere trisecato con riga e compasso.

Nel caso delle radici quadrate (e in realtà anche delle radici n-esime dove n è una potenza di 2), i calcoli si possono invece sempre eseguire, utilizzando eventualmente le formule di bisezione. Anche in questo caso proponiamo un esempio per chiarire il concetto.

$$\sqrt{3-4\mathrm{i}}=\sqrt{(5,\vartheta)}=\left\{\left(\sqrt{5},\,\frac{\vartheta}{2}\right),\left(\sqrt{5},\,\frac{\vartheta}{2}+\pi\right)\right\}=\pm\sqrt{5}\left(\cos\frac{\vartheta}{2}+\mathrm{i}\sin\frac{\vartheta}{2}\right),\quad \begin{cases}\cos\vartheta=\dfrac{3}{5}\\ \sin\vartheta=\dfrac{4}{5}\end{cases}.$$

Anche se non si riesce ad esprimere l'argomento ϑ in termini di angoli notevoli, si possono comunque eseguire i calcoli indicati con le citate formule di bisezione, osservando che $3\pi/2 < \vartheta < 2\pi$ e quindi $3\pi/4 < \vartheta/2 < \pi$. Si trova facilmente:

$$\sqrt{3-4\mathrm{i}} = \pm\sqrt{5}\left(-\frac{2}{\sqrt{5}} + \mathrm{i}\frac{1}{\sqrt{5}}\right) = \pm(-2+\mathrm{i}).$$

Naturalmente il calcolo delle radici quadrate si poteva anche fare senza l'uso della forma trigonometrica, semplicemente osservando che si trattava di calcolare x ed y tali che $(x+\mathrm{i}y)^2 = 3-4\mathrm{i}$ e giungendo ad un sistema di due equazioni in due incognite.

14.7. Il teorema fondamentale dell'algebra

Consideriamo un polinomio $P(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, $n \geq 1$, $a_n \neq 0$. Come è noto si chiama radice del polinomio un numero complesso w che annulla il polinomio.

Il teorema fondamentale dell'algebra afferma che ogni polinomio del tipo citato ha almeno una radice (reale o complessa).

La dimostrazione di questo teorema esula dagli scopi di queste pagine, ma segnaliamo che si tratta di un teorema di fondamentale, tanto che la sua validità nell'insieme dei complessi è uno dei motivi che rendono questo insieme numerico così importante.

Interessiamoci ora di qualche conseguenza di questo teorema.

Se $P(z)$ è un polinomio di grado n, e z_1 è una sua radice, per il teorema di Ruffini avremo $P(z) = (z-z_1)Q_1(z)$ ove $Q_1(z)$ è il polinomio di grado $n-1$ ottenuto dividendo P per $z-z_1$. Anche $Q_1(z)$ è un polinomio e dunque avrà almeno una radice, per cui potremo scrivere $P(z) = (z-z_1)(z-z_2)Q_2(z)$. Iterando il procedimento potremo concludere che il polinomio P si potrà sempre scrivere come il prodotto di una costante (che deve necessariamente essere uguale ad a_n) per n fattori di primo grado:

$$P(z) = a_n(z-z_1)(z-z_2)(z-z_3)\cdots(z-z_n) \tag{14.14}$$

I numeri $z_1, z_2, \ldots, z_n$ sono le radici di $P(z)$. Tali radici possono essere tutte diverse o alcune possono essere ripetute. Un esempio banale in cui sono tutte diverse è $P(z) = z^n - 1$, in cui le radici sono le n radici n-esime dell'unità; un esempio in cui sono tutte uguali è $P(z) = (z-1)^n$, in cui tutte le radici sono uguali a 1. Il numero di volte che una radice si ripete si chiama molteplicità della radice stessa. E' ovvio che la somma delle molteplicità di tutte le radici è esattamente n.

È di particolare interesse il caso in cui i coefficienti del polinomio siano numeri reali. In questo caso consideriamo una radice w e la sua complessa coniugata $\overline{w}$. Utilizzando le proprietà del coniugato, e ricordando che il coniugato di un reale è il numero stesso, si trova:

$$\begin{aligned}
a_n w^n + a_{n-1}w^{n-1} + \cdots + a_1 w + a_0 = 0 \quad &\Rightarrow \quad \overline{a_n w^n + a_{n-1}w^{n-1} + \cdots + a_1 w + a_0} = 0 \\
&\Rightarrow \quad \overline{a_n w^n} + \overline{a_{n-1}w^{n-1}} + \cdots + \overline{a_1 w} + \overline{a_0} = 0 \\
&\Rightarrow \quad a_n \overline{w^n} + a_{n-1}\overline{w^{n-1}} + \cdots + a_1 \overline{w} + a_0 = 0 \\
&\Rightarrow \quad a_n \overline{w}^n + a_{n-1}\overline{w}^{n-1} + \cdots + a_1 \overline{w} + a_0 = 0
\end{aligned}$$

Se ne deduce che anche $\overline{w}$ è una radice: in un polinomio a coefficienti reali le radici si presentano sempre a coppie di radici tra di loro coniugate. Nella decomposizione del polinomio, se si moltiplicano tra di loro i due fattori che si riferiscono a radici coniugate si trova

$$(z-w)(z-\overline{w}) = z^2 - zw - z\overline{w} - z\overline{w} + w\overline{w} = z^2 - z(w+\overline{w}) + w\overline{w},$$

espressione che non contiene numeri complessi, ma è di secondo grado. Dunque un polinomio a coefficienti reali si può sempre decomporre nel prodotto di fattori non contenenti numeri complessi, al massimo di grado due. Questo fatto ha come conseguenza che un polinomio di grado dispari, a coefficienti reali, ha almeno una radice reale.

15. Algebra lineare e geometria analitica

In questo capitolo proponiamo una introduzione, a livello elementare, dei concetti fondamentali dell'algebra lineare, con applicazioni allo studio della geometria analitica del piano e, molto sommariamente, dello spazio. La trattazione proposta è molto semplificata e, come al solito, non sono previste le dimostrazioni dei risultati via via presentati, tuttavia può essere utile anche a chi si appresta ad affrontare il corso di geometria del primo anno universitario.

15.1. Matrici e operazioni tra matrici

Abbiamo già fatto un cenno (vedi la definizione 5.6 nel capitolo 5) al concetto di matrice quadrata di ordine 2. Estendiamo ora questo concetto al caso di matrici, anche non quadrate, con un qualunque numero di righe e di ccolonne.

Definizione 15.1 (Matrice $m \times n$)**.** *Una tabella di numeri reali*[1]*, con m righe e n colonne, indicata con la scrittura seguente*[2]

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \tag{15.1}$$

si chiama matrice $m \times n$. *Se serve precisare il numero di righe e di colonne si può scrivere*

$$A_{[m\times n]} \quad \textit{oppure} \quad A_{m\times n} \quad \textit{o semplicemente} \quad A_{m,n}\,.$$

I numeri a_{ij}, $i = 1,2,\dots,m$, $j = 1,2,\dots,n$ si chiamano *elementi* della matrice e si usa scrivere anche

$$A = \left(a_{ij}\right), \quad i = 1,2,\dots,m, \quad j = 1,2,\dots,n\,. \tag{15.2}$$

In alcuni casi, per evitare confusioni, si scrive $a_{i,j}$ invece di a_{ij}.

Se $m \neq n$ la matrice si dice *rettangolare*. Se $m = n$ la matrice si dice *quadrata* e il comune numero di righe e di colonne si chiama *ordine* della matrice. Nelle matrici quadrate gli elementi a_{ii} costituiscono la *diagonale principale*. A volte ha interesse considerare, sempre per le matrici quadrate, anche l'altra diagonale, detta *diagonale secondaria*. Se $m = 1$ (una sola riga) la matrice si dice anche un *vettore riga*; se $n = 1$ (una sola colonna) la matrice si dice anche un *vettore colonna*.

[1] Si possono naturalmente anche considerare matrici con elementi complessi, ma non ce ne occupiamo in questo capitolo. Segnaliamo tuttavia che esse sono molto importanti in numerose applicazioni, per esempio alla fisica.

[2] Al posto delle parentesi tonde si possono usare come delimitatori anche parentesi quadre, come indicato nell'elenco delle notazioni utilizzate.

Due matrici $\boldsymbol{A}$ e $\boldsymbol{B}$, con lo stesso numero m di righe e n di colonne, sono *uguali* se e soltanto se

$$a_{ij} = b_{ij}, \quad i = 1,2,\dots,m, \quad j = 1,2,\dots,n,$$

cioè se e solo se hanno ordinatamente uguali tutti gli elementi

Definizione 15.2 (Matrice trasposta). *Data una matrice $\boldsymbol{A}_{m,n}$, la matrice $n \times m$ ottenuta da $\boldsymbol{A}$ scambiando le righe con le colonne si chiama la* trasposta, $\boldsymbol{A}^\mathsf{T}$, *di $\boldsymbol{A}$:*

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad , \quad \boldsymbol{A}^\mathsf{T} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}. \tag{15.3}$$

Per esempio

$$\text{se } \boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}, \quad \text{si ha } \boldsymbol{A}^\mathsf{T} = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}.$$

In particolare la matrice trasposta di un vettore colonna è un vettore riga e viceversa. Poiché si ha spesso bisogno di utilizzare vettori colonna, questa osservazione può facilitare le scritture:

$$\text{invece di scrivere } \boldsymbol{A} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix}, \text{ si può scrivere } \boldsymbol{A} = \begin{pmatrix} 1 & 2 & \dots & n \end{pmatrix}^\mathsf{T}.$$

Per evitare confusioni, se i vettori colonna sono rappresentati con la trasposta di un vettore riga si usano spesso delle virgole per separare i singoli elementi, come nelle n-uple. I vettori colonna si indicano usualmente con lettere minuscole in grassetto o sormontate da una freccia:

$$\boldsymbol{v} = \vec{v} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \dots & n \end{pmatrix}^T = (1,2,\dots,n)^T. \tag{15.4}$$

Alcuni tipi di matrici quadrate ricorrono spesso nelle applicazioni e per esse si usa la seguente terminologia e i seguenti simboli:

- *matrice simmetrica*: $a_{ij} = a_{ji}$, ovvero $\boldsymbol{A} = \boldsymbol{A}^\mathsf{T}$;
- *matrice diagonale*: $a_{ij} = 0$ se $i \neq j$ (gli elementi fuori dalla diagonale principale sono nulli);
- *matrice triangolare superiore*: gli elementi sotto la diagonale principale sono nulli;
- *matrice triangolare inferiore*: gli elementi sopra la diagonale principale sono nulli;
- *matrice unità* o *matrice identica*: gli elementi della diagonale principale sono tutti uguali a 1, mentre gli altri sono nulli; essa sarà indicata con $\boldsymbol{I}$ o con $\boldsymbol{E}$.

La matrice, anche non quadrata, con tutti gli elementi nulli si chiama anche *matrice nulla* e si può indicare con $\mathtt{0}$. La matrice i cui elementi sono gli opposti degli elementi di A si chiama *opposta* di A e si indica con $-A$. Questa denominazione è giustificata dal fatto che questa matrice è proprio l'opposta della matrice A nell' operazione di somma che sarà tra poco definita.

Definizione 15.3 (Somma di matrici)**.** *Date due matrici A e B, entrambe $m \times n$, si chiama* somma *di A e B, $A+B$, la matrice C ottenuta sommando gli elementi corrispondenti:*

$$A+B=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}. \tag{15.5}$$

Definizione 15.4 (Prodotto di una matrice per uno scalare)**.** *Data una matrice A e un reale c, si chiama* prodotto *di A per c la matrice, cA, ottenuta moltiplicando per c tutti gli elementi di A:*

$$cA=\begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}. \tag{15.6}$$

Le operazioni ora introdotte godono delle seguenti proprietà, di immediata verifica:

- $(A+B)+C=A+(B+C)$: proprietà associativa della somma;
- $A+B=B+A$: proprietà commutativa della somma;
- $A+\mathtt{0}=A$: la matrice nulla è elemento neutro della somma;
- $A+(-A)=\mathtt{0}$: esistenza dell'opposto per la somma;
- $c(A+B)=cA+cB$;
- $(c+d)A=cA+dA$;
- $0A=\mathtt{0}$;
- $-1A=-A$.

Definizione 15.5 (Prodotto righe per colonne di due matrici)**.** *Date due matrici $A_{m,p}$ e $B_{p,n}$ si chiama loro* prodotto righe per colonne *la matrice $C_{m,n}$ i cui elementi c_{ij} sono ottenuti moltiplicando ordinatamente gli elementi della i-esima riga di A per la j-esima colonna di B e sommando i prodotti così ottenuti:*

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}=\sum_{k=1}^{n}a_{ik}b_{kj}\,. \tag{15.7}$$

L'algoritmo indicato è un po' noioso da esprimere a parole: cercheremo di sintetizzarlo in modo da ottenere una formulazione facilmente utilizzabile, ragionando su un esempio. Supponiamo di dover moltiplicare una matrice $A_{2\times 3}$ per una matrice $B_{3\times 4}$: il prodotto sarà una matrice $C_{2\times 4}$. Potremo simbolicamente rappresentare la situazione con la (15.8), dove gli elementi delle matrici sono rappresentati con box da riempire.

$$\begin{pmatrix} \square & \square & \square \\ \square & \square & \square \end{pmatrix}\cdot\begin{pmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{pmatrix}=\begin{pmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \end{pmatrix}. \tag{15.8}$$

Ciascuno dei box da riempire della matrice $\boldsymbol{C}$ ha un ben preciso numero di riga e di colonna, come evidenziato nella (15.9).

$$\begin{pmatrix} \square & \square & \square \\ \square & \square & \square \end{pmatrix} \cdot \begin{pmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{pmatrix} = \begin{pmatrix} \boxed{1,1} & \boxed{1,2} & \boxed{1,3} & \boxed{1,4} \\ \boxed{2,1} & \boxed{2,2} & \boxed{2,3} & \boxed{2,4} \end{pmatrix}. \tag{15.9}$$

Consideriamo allora uno dei box da riempire, per esempio il box $\boxed{2,3}$. Basterà moltiplicare ciascun elemento della 2ª riga della prima matrice (che ha 3 elementi), per *il corrispondente* elemento della 3ª colonna della seconda matrice (che ha sempre 3 elementi) e poi sommare i prodotti così ottenuti: il risultato andrà ad occupare il box $\boxed{2,3}$ della matrice prodotto. La (15.10) illustra il procedimento.

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \boxed{2,1} & \boxed{2,2} & \boxed{2,3} \end{pmatrix} \cdot \begin{pmatrix} \cdot & \cdot & \boxed{1,3} & \cdot \\ \cdot & \cdot & \boxed{2,3} & \cdot \\ \cdot & \cdot & \boxed{3,3} & \cdot \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \boxed{2,3} & \cdot \end{pmatrix}. \tag{15.10}$$

Un esempio numerico chiarirà ancora meglio come si opera.

$$\begin{pmatrix} 1 & -2 & 4 \\ \boxed{-3} & \boxed{-1} & \boxed{5} \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & \boxed{-2} & 3 \\ 1 & -1 & \boxed{3} & 4 \\ 5 & 2 & \boxed{1} & 3 \end{pmatrix} = \begin{pmatrix} 20 & 11 & -4 & 7 \\ 18 & 8 & \boxed{8} & 2 \end{pmatrix}$$

Il numero 8, al posto $(2,3)$ della matrice prodotto, è stato ottenuto con il seguente calcolo:

$$(-3)(-2)+(-1)(3)+(5)(1)=8\,.$$

Si noti che la definizione ha senso solo se il numero di colonne della prima matrice è uguale al numero di righe della seconda. Questo significa che, in generale, dato il prodotto $\boldsymbol{AB}$, non ha nemmeno senso considerare il prodotto $\boldsymbol{BA}$.

Nel caso particolarmente importante di prodotto tra matrici quadrate, esse devono avere lo stesso ordine; in questo caso hanno sempre senso sia $\boldsymbol{AB}$ che $\boldsymbol{BA}$, ma può benissimo succedere che $\boldsymbol{AB} \neq \boldsymbol{BA}$. Se $\boldsymbol{AB} = \boldsymbol{BA}$ si dice che le matrici $\boldsymbol{A}$ e $\boldsymbol{B}$ *commutano*, altrimenti che *non commutano*.

Esempio 15.1. Per calcolare il prodotto $\boldsymbol{C} = \boldsymbol{AB}$ tra

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & -1 & 4 \\ 1 & 0 & -1 & 2 \end{pmatrix} \quad \text{e} \quad \boldsymbol{B} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & -3 \\ 0 & 0 & 1 \\ 1 & 0 & 2 \end{pmatrix},$$

si osserva intanto che $\boldsymbol{A}_{2,4}\boldsymbol{B}_{4,3} = \boldsymbol{C}_{2,3}$; poi si ha:

— $c_{11} = 2 \times 1 + 3 \times 2 + (-1) \times 0 + 4 \times 1 = 10$ (prima riga × prima colonna);

- $c_{12} = 2 \times 0 + 3 \times 1 + (-1) \times 0 + 4 \times 0 = 3$ (prima riga $\times$ seconda colonna);
- $c_{13} = 2 \times 2 + 3 \times (-3) + (-1) \times 1 + 4 \times 2 = 2$ (prima riga $\times$ terza colonna);
- $c_{21} = 1 \times 1 + 0 \times 2 + (-1) \times 0 + 2 \times 1 = 3$ (seconda riga $\times$ prima colonna);
- $c_{22} = 1 \times 0 + 0 \times 1 + (-1) \times 0 + 2 \times 0 = 0$ (seconda riga $\times$ seconda colonna);
- $c_{23} = 1 \times 2 + 0 \times (-3) + (-1) \times 1 + 2 \times 2 = 7$ (seconda riga $\times$ terza colonna).

La matrice prodotto è dunque:

$$C = \begin{pmatrix} 10 & 3 & 2 \\ 3 & 0 & 7 \end{pmatrix}.$$

Esempio 15.2. Le due matrici A e B seguenti non commutano:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Infatti

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \neq \quad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Il prodotto tra matrici gode delle seguenti proprietà (si suppone che le operazioni indicate siano definite):

- $A(BC) = (AB)C$: proprietà associativa del prodotto tra matrici;
- $A(B+C) = AB + AC$, $(A+B)C = AC + BC$: proprietà distributiva del prodotto rispetto alla somma;
- $c(AB) = (cA)B$.

Nell'insieme delle matrici quadrate di ordine n vale anche la

- $AI = IA = A$, proprietà che giustifica il nome di matrice unità dato alla matrice I.

È molto importante il fatto che, nell'insieme delle matrici quadrate di ordine n, data una matrice A, in generale *non esiste* una matrice *inversa*, ovvero una matrice B tale che $AB = BA = I$. Le matrici che godono di questa proprietà si dicono *invertibili*. Si può provare che le matrici invertibili hanno una sola inversa.

Esempio 15.3. Sia data la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

Per verificare se è invertibile e, in caso affermativo, calcolarne l'inversa si deve verificare se esiste una matrice

$$B = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

tale che $AB = BA = I$. Si ottiene:

$$AB = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} x + 2z & y + 2t \\ 2x - z & 2y - t\,. \end{pmatrix}$$

Si deve dunque avere:

$$\begin{cases} x+2z=1 \\ y+2t=0 \\ 2x-z=0 \\ 2y-t=1 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x+2z=1 \\ 2x-z=0 \end{cases} \wedge \begin{cases} y+2t=0 \\ 2y-t=1 \end{cases} .$$

La risoluzione dei due sistemi può essere fatta con la regola di Cramer, già studiata nel capitolo 5, e si ottiene:

$$x = \frac{\begin{vmatrix} 1 & 2 \\ 0 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}} = \frac{1}{5},\ z = \frac{\begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}} = \frac{2}{5},\ y = \frac{\begin{vmatrix} 0 & 2 \\ 1 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}} = \frac{2}{5},\ t = \frac{\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}} = -\frac{1}{5}.$$

La matrice cercata esiste ed è:

$$\begin{pmatrix} 1/5 & 2/5 \\ 2/5 & -1/5 \end{pmatrix}.$$

È facile ora verificare che si ha non solo $\boldsymbol{AB} = \boldsymbol{I}$, ma anche $\boldsymbol{BA} = \boldsymbol{I}$.

La matrice inversa di una data matrice $\boldsymbol{A}$, se esiste, si indica con

$$A^{-1}\colon AA^{-1} = A^{-1}A = I\,. \tag{15.11}$$

È molto importante la proprietà seguente: se $\boldsymbol{A}$ e $\boldsymbol{B}$ sono due matrici invertibili allora

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\,. \tag{15.12}$$

Come conseguenza della definizione di prodotto tra matrici, si possono definire anche le *potenze* di una matrice quadrata:

$$A^2 = AA, \quad A^3 = AAA, \quad \text{ecc.}, \tag{15.13}$$

e si può convenire che

$$A^1 = A, \quad A^0 = \boldsymbol{I}\,. \tag{15.14}$$

15.1.1. Determinante di una matrice quadrata

Daremo solo una definizione ricorsiva di determinante, basata su un teorema dovuto a Laplace: data la definizione per la matrici di ordine 1, la definizione per matrici di ordine superiore si riconduce, mediante passaggi successivi, al caso $n = 1$. Ricordo che abbiamo già definito il determinante per una matrice di ordine 2: naturalmente la definizione che daremo comprenderà quella già nota come caso particolare.

Per il determinante di una matrice A useremo le notazioni già utilizzate per quelle di ordine 2:

$$|A| = \det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\,. \end{vmatrix}. \tag{15.15}$$

Definizione 15.6 (Minori di una matrice). *Data una matrice A si chiama sottomatrice di A ogni matrice ottenuta da A sopprimendo un certo numero di righe e un certo numero di colonne (anche non consecutive). Si chiama* minore *di A una sottomatrice quadrata di A. Si può usare il termine minore anche per indicare il determinante del minore stesso.*

Definizione 15.7 (Complemento algebrico o cofattore). *Dato un elemento a_{ij} di una matrice quadrata A si chiama suo* complemento algebrico *o* cofattore*, e si indica con A_{ij}, il determinante, moltiplicato per $(-1)^{i+j}$, del minore che si ottiene sopprimendo la riga e la colonna di A che si intersecano in a_{ij}; questo minore è anche detto minore complementare di a_{ij}.*

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & m_{m,j} & a_{m,j+1} & \cdots & a_{m,n} \end{pmatrix}, \tag{15.16}$$

$$A_{ij} = (-1)^{i+j} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,n} \end{vmatrix}$$

Definizione 15.8 (Determinante). *Premesso che per una matrice di ordine* 1 *il determinante coincide con l'unico elemento, data una matrice quadrata A di ordine n e considerata una sua riga o colonna qualunque, il* determinante *di A è il numero ottenuto moltiplicando gli elementi della riga o colonna scelta per i rispettivi cofattori e sommando i risultati ottenuti. A partire dalla riga r-esima di A si ha:*

$$|A| = a_{r1}A_{r1} + a_{r2}A_{r2} + \cdots + a_{rn}A_{rn}\,; \tag{15.17}$$

a partire dalla colonna p-esima di A si ha invece:

$$|A| = a_{1p}A_{1p} + a_{2p}A_{2p} + \cdots + a_{np}A_{np}\,. \tag{15.18}$$

In sostanza, per calcolare il determinante una matrice di ordine n si devono calcolare n determinanti di matrici di ordine $n-1$, per calcolare i quali si devono calcolare $n-1$ determinanti di matrici di ordine $n-2$, e così via. Naturalmente si prova che nella definizione data la scelta della riga o colonna è ininfluente ai fini del risultato. È immediato che per matrici di ordine 2 la definizione porge il già noto risultato:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

Esempio 15.4. Per calcolare il determinante di

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ -2 & 1 & 1 \end{pmatrix},$$

scegliendo la prima colonna, si ha

$$\begin{aligned} |A| &= 1 \cdot (-1)^{1+1} \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} + 3 \cdot (-1)^{2+1} \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} + (-2) \cdot (-1)^{3+1} \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = \\ &= 1 \cdot (1-2) - 3 \cdot (2+1) - 2 \cdot (4+1) = -17. \end{aligned}$$

Valgono le seguenti proprietà del determinante.

— $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$ (regola di Binet);
— se le matrici $\boldsymbol{A}$ e $\boldsymbol{B}$ differiscono solo per lo scambio di due linee parallele, allora $\det(\boldsymbol{A}) = -\det(\boldsymbol{B})$;
— se una matrice $\boldsymbol{A}$ ha due righe uguali o proporzionali, allora $\det(\boldsymbol{A}) = 0$;
— aggiungendo a una linea di una matrice $\boldsymbol{A}$ un'*altra* linea della matrice stessa, eventualmente moltiplicata per un numero, il determinante di $\boldsymbol{A}$ non cambia;
— moltiplicando una linea di $\boldsymbol{A}$ per un numero c, il determinante di $\boldsymbol{A}$ risulta moltiplicato per c.

15.1.2. Calcolo dell'inversa di una matrice

Teorema 15.9 (Condizione di invertibilità di una matrice)**.** *Una matrice quadrata* $\boldsymbol{A}$ *è invertibile se e solo se il suo determinante è diverso da* 0.

Teorema 15.10 (Calcolo dell'inversa)**.** *Data una matrice* $\boldsymbol{A}$ *con determinante diverso da* 0*, per ottenere la matrice inversa si procede nel seguente modo:*

1. *si trova la trasposta* $\boldsymbol{A}^{\mathsf{T}}$ *di* $\boldsymbol{A}$*;*
2. *si costruisce la matrice i cui elementi sono i complementi algebrici della trasposta* $\boldsymbol{A}^{\mathsf{T}}$*;*
3. *si divide questa matrice per il determinante di* $\boldsymbol{A}$*.*

È immediato che sarebbe del tutto equivalente costruire prima la matrice dei complementi algebrici e poi farne la trasposta.

Osservazione 15.11. La matrice costruita con i complementi algebrici della trasposta di una matrice $\boldsymbol{A}$ è chiamata, da taluni autori, *aggiunta* della matrice $\boldsymbol{A}$ e indicata con $\operatorname{agg}\boldsymbol{A}$. Questa nomenclatura non è però universale e, anzi, il termine *matrice aggiunta* è usato nelle norme ISO con il significato di *matrice trasposta coniugata* (trasposta della matrice ottenuta prendendo i complessi coniugati degli elementi di una matrice: nel caso dei reali coincide con la trasposta). Per questo motivo evitiamo di usare il termine matrice aggiunta in questo testo.

Esempio 15.5. Data

$$A = \begin{pmatrix} 1/2 & 0 & 1 \\ 0 & 4 & 1 \\ 3 & 0 & 2 \end{pmatrix},$$

si trova $|A| = -8 \neq 0$, e la matrice dei complementi algebrici della trasposta è

$$\begin{pmatrix} 8 & 0 & -4 \\ 3 & -2 & -1/2 \\ -12 & 0 & 2 \end{pmatrix}.$$

Dunque

$$A^{-1} = \begin{pmatrix} -1 & 0 & 1/2 \\ -3/8 & 1/4 & 1/16 \\ 3/2 & 0 & -1/4. \end{pmatrix}$$

15.1.3. Rango di una matrice

Data una matrice (anche rettangolare) $A_{m,n}$, da essa possiamo estrarre minori di ordine $1, 2, \dots, r$, con $1 \leq r \leq \min(m, n)$.

Definizione 15.12 (Rango). *Il* rango *o* caratteristica *di una matrice $A_{m,n}$,* rg(A)*, è il massimo ordine dei suoi minori con determinante non nullo (brevemente il massimo ordine dei suoi minori non nulli).*

In sostanza la definizione data implica che se $\mathrm{rg}(A) = p$

- esiste almeno un minore di ordine p non nullo (cioè con determinante diverso da zero);
- tutti gli eventuali minori di ordine $p + 1$ sono nulli.

Per la matrice nulla si pone $\mathrm{rg}(A) = 0$. Se A è una matrice quadrata di ordine n allora

$$\mathrm{rg}(A) = n \quad \Leftrightarrow \quad \det(A) \neq 0 \quad \Leftrightarrow \quad A \text{ è invertibile}.$$

Esempio 15.6. La matrice

$$A = \begin{pmatrix} -1 & 3 & 2 & 5 \\ 6 & -2 & 4 & 3 \\ -2 & 6 & 4 & 10 \end{pmatrix}$$

ha 4 minori di ordine 3, e precisamente:

$$\begin{pmatrix} -1 & 3 & 2 \\ 6 & -2 & 4 \\ -2 & 6 & 4 \end{pmatrix}, \quad \begin{pmatrix} -1 & 3 & 5 \\ 6 & -2 & 3 \\ -2 & 6 & 10 \end{pmatrix}, \quad \begin{pmatrix} -1 & 2 & 5 \\ 6 & 4 & 3 \\ -2 & 4 & 10 \end{pmatrix}, \quad \begin{pmatrix} 3 & 2 & 5 \\ -2 & 4 & 3 \\ 6 & 4 & 10 \end{pmatrix},$$

e tutti i loro determinanti sono nulli. Essa ha inoltre 18 minori di ordine 2 e si verifica subito che il minore

$$\begin{pmatrix} 3 & 2 \\ -2 & 4 \end{pmatrix}$$

ha determinante diverso da 0. Se ne conclude che la matrice data ha rango 2.

In pratica per calcolare il rango di una matrice si procede secondo il seguente schema:

- si comincia col calcolare i minori di ordine massimo possibile, $p = \min(m, n)$: se si trova un minore non nullo, allora il rango è p;

— se tutti i minori di ordine p sono nulli, si procede a calcolare i minori di ordine $p-1$: se si trova un minore non nullo, allora il rango è $p-1$;
— se tutti i minori di ordine $p-1$ sono nulli, si procede successivamente con $p-2$, $p-3$, ecc.

Non si confonda il concetto di rango con quello di determinante: il rango è un numero naturale (che può essere zero solo se per la matrice nulla), il determinante è invece un numero reale (positivo, negativo o nullo).

15.2. Sistemi lineari

La risoluzione di sistemi di equazioni lineari in n incognite si può fare generalizzando il metodo di Cramer già visto per i sistemi di due equazioni in due incognite. Accenneremo anche al metodo di riduzione o eliminazione di Gauss, per la sua straordinaria efficienza.

15.2.1. Definizioni

Un sistema di m equazioni lineari in n incognite è un sistema del tipo:

$$\left\{\begin{array}{lllllllll} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \cdots & & & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}\right. . \tag{15.19}$$

Esattamente come per i sistemi di due equazioni in due incognite, un sistema di m equazioni in n incognite si dice

— *compatibile*, se ha soluzioni, e in questo caso si dice
  - *determinato*, se ha una sola soluzione,
  - *indeterminato*, se ha infinite soluzioni;
— *incompatibile*, se non ha soluzioni.

Si tenga ben presente che soluzione di un sistema in n incognite è un'n-upla di numeri reali che sostituiti ordinatamente al posto delle incognite rendono vera ciascuna equazione del sistema. Quando si parla di *una sola* soluzione si intende dunque *una sola* n-upla.

La matrice

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \tag{15.20}$$

è detta *matrice dei coefficienti*, o *matrice incompleta* del sistema. Se si considerano i vettori colonna

$$\vec{x}=(x_1, x_2, \dots, x_n)^{\mathsf{T}}, \quad \vec{b}=(b_1, b_2, \dots, b_m)^{\mathsf{T}}, \tag{15.21}$$

il sistema può essere scritto nella forma compatta

$$A\vec{x} = \vec{b}\,, \tag{15.22}$$

che è formalmente simile nella scrittura a un'equazione di primo grado in un'incognita, $ax = b$.

La matrice

$$A|b = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right), \tag{15.23}$$

ottenuta da A aggiungendo la colonna dei termini noti, è detta *matrice completa* del sistema.

Osservazione 15.13 (Incognite e parametri). Nella risoluzione di un sistema di equazioni le x_i sono le *incognite*, cioè variabili appartenenti a un dato insieme (nel nostro caso l'insieme dei reali) e soggette alle condizioni espresse dalle equazioni stesse: la risoluzione permette di trovare gli eventuali valori delle incognite per cui le equazioni sono verificate.

Capita frequentemente nelle applicazioni che i coefficienti a_{ij} non siano costanti (numeri reali "fissati"), ma che dipendano da una o più variabili, che però giocano il ruolo di *parametri*, cioè sono variabili *non* soggette ad alcuna condizione (tranne eventualmente quelle per l'esistenza).

È indispensabile prestare molta attenzione alla differenza tra le incognite (che si potrebbero chiamare anche *variabili condizionate*) e i parametri (ovvero *variabili non condizionate*). Chiariamo il fatto con un esempio. Nell'equazione "in x"

$$(t+1)x = 3t\,,$$

x è l'incognita, t è un parametro: t può assumere qualunque valore, mentre, una volta scelto t, x va determinato in modo da soddisfare l'equazione data. Se per esempio scelgo $t = 1$, x *deve necessariamente valere* $3/2$; se scelgo $t = 2$, x *deve necessariamente valere* 2; se scelgo $t = -1$ non esiste alcun valore di x che verifichi l'equazione, e così via. Risolvere un'equazione (o un sistema) con parametri significa risolverlo per ogni valore consentito del parametro. Naturalmente non si dovrà in generale risolvere esplicitamente l'equazione o il sistema dando ai parametri tutti i valori consentiti (cosa del resto impossibile visto che di solito le scelte possono essere infinite): si dovranno piuttosto esaminare solo alcuni casi tipo. Nell'esempio sopra riportato basterà considerare solo due casi:

$$\begin{aligned} &t = -1\colon \text{l'equazione non ha alcuna soluzione;} \\ &t \neq -1\colon x = \frac{3t}{t+1}\,. \end{aligned}$$

15.2.2. Il metodo di Cramer

Il teorema fondamentale sui sistemi lineari è quello di Rouché-Capelli, che fornisce una condizione sulla risolubilità basata unicamente sulle caratteristiche delle matrici A e $A|b$, estendendo quanto già detto a proposito dell'equazione $ax = b$, la cui risolubilità dipende dai valori di a e b.

Teorema 15.14 (Rouchè-Capelli). *Un sistema lineare di m equazioni in n incognite è compatibile se e soltanto se le matrici incompleta e completa hanno lo stesso rango. Il valore comune di questo rango è detto* rango del sistema.

Una volta controllata la compatibilità del sistema si può procedere alla sua risoluzione seguendo il percorso che segue.

Consideriamo un sistema lineare di m equazioni in n incognite e supponiamo che il rango comune delle due matrici completa e incompleta sia r ($\leq \min(m,n)$). Consideriamo inoltre il minore (di ordine r) della matrice incompleta che abbiamo usato per determinarne il rango e che quindi ha determinante non nullo. Allora:

1. si sopprimono le eventuali equazioni le cui righe dei coefficienti non compaiono nel minore detto;
2. si portano a secondo membro i termini contenenti le eventuali incognite le cui colonne dei coefficienti non compaiono nel minore detto (queste incognite assumeranno il ruolo di parametri, cioè non saranno sottoposte ad alcuna condizione: potrebbe essere utile indicarle anche formalmente con un opportuno nome, di solito t, u, ...);
3. si ottiene così un sistema quadrato di r equazioni in r incognite, che si risolve con la regola di Cramer che è proposta nel successivo teorema 15.15 e che estende quanto già detto per i sistemi di due equazioni in due incognite.

Teorema 15.15 (Regola di Cramer). *Sia dato un sistema di r equazioni in r incognite, con matrice dei coefficienti, $\mathbf{A}$, a determinante non nullo*[3]*. Si considerino inoltre le matrici $\mathbf{A}_i$ ottenute sostituendo alla prima, seconda, ecc., colonna, la colonna dei termini noti. Si ha:*

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad i = 1, 2, \ldots, r\,. \tag{15.24}$$

Esempio 15.7. Risolvere il sistema

$$\begin{cases} x_1 + x_2 - x_3 + x_4 = 1 \\ x_1 + 2x_3 - 2x_4 = 1 \\ 4x_1 + 3x_2 - x_3 + x_4 = 4 \\ 2x_1 + x_2 + x_3 - x_4 = 2 \end{cases}.$$

La matrice $A|b$ è:

$$A|b = \left(\begin{array}{cccc|c} 1 & 1 & -1 & 1 & 1 \\ 1 & 0 & 2 & -1 & 1 \\ 4 & 3 & -1 & 1 & 4 \\ 2 & 1 & 1 & -1 & 2 \end{array}\right).$$

Con un po' di pazienza (e magari con un po' d'occhio nel semplificare la matrice!), si trova che il determinante di A è nullo, e così pure tutti quelli dei minori di ordine 4 della matrice completa. Stesso discorso per i minori di ordine 3, sia della matrice incompleta che di quella completa. Ci sono invece molti minori di ordine 2 della matrice incompleta che sono non nulli. Dunque il rango del sistema è 2. Scegliamo, per esempio, il minore

$$\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix},$$

[3] Un sistema siffatto è sicuramente compatibile in quanto la matrice completa, avendo le stesse righe di quella incompleta, non può avere rango maggiore di r, né, d'altro canto può avere rango minore di r, in quanto è una "sopramatrice" della matrice incompleta.

che ha determinante -6. La prima e la quarta equazione possono essere eliminate; i termini contenenti la prima e la quarta incognita si portano a secondo membro. Ponendo $x_1 = t$ e $x_4 = u$ si ottiene:

$$\begin{cases} & & 2x_3 & = & 1-t+2u \\ 3x_2 & - & x_3 & = & 4-4t-u \end{cases}.$$

Applicando la regola di Cramer si ha quindi:

$$x_1 = t\,, \quad x_2 = \frac{\begin{vmatrix} 1-t+2u & 2 \\ 4-4t-u & -1 \end{vmatrix}}{\begin{vmatrix} 0 & 2 \\ 3 & -1 \end{vmatrix}}, \quad x_3 = \frac{\begin{vmatrix} 0 & 1-t+2u \\ 3 & 4-4t-u \end{vmatrix}}{\begin{vmatrix} 0 & 2 \\ 3 & -1 \end{vmatrix}}, \quad x_4 = u\,,$$

ovvero

$$x_1 = t\,, \quad x_2 = \frac{3-3t}{2}, \quad x_3 = \frac{1-t+2u}{2}, \quad x_4 = u\,.$$

Le soluzioni dipendono da due parametri, t e u. Si usa esprimere questo fatto dicendo che il sistema ha

$$\infty^2 \text{ soluzioni.}$$

In generale se un sistema in n incognite ha rango r, si dice che ha

$$\infty^{n-r} \text{ soluzioni},$$

con la convenzione che, se $n-r=0$, si intende che il sistema ha una sola soluzione. Il numero $n-r$ rappresenta, in un senso che qui non interessa precisare, la *dimensione* dello spazio delle soluzioni.

15.2.3. Il metodo di eliminazione di Gauss

La risoluzione di un sistema lineare con il *metodo di riduzione* o *eliminazione* di Gauss prevede una serie di operazioni, da eseguirsi sia sulla matrice completa che su quella incompleta, atte a trasformare la matrice incompleta del sistema nella forma cosiddetta *a scala*, che nella sua versione più generale ha l'aspetto indicato nella (15.25).

$$\begin{pmatrix} 0 & \cdots & 0 & p_1 & * & \cdots & * & * & * & \cdots & * & * & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & p_2 & * & \cdots & * & * & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & p_3 & * & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & \cdots \end{pmatrix} \tag{15.25}$$

Nella (15.25)

— eventuali colonne formate da tutti zeri[4] si trovano sull'estrema sinistra;
— eventuali righe formate da tutti zeri si trovano in fondo;
— le altre righe, a partire dall'alto, hanno un primo valore p_i, detto *pivot*, diverso da zero, situato su colonne via via crescenti muovendosi dall'alto verso il basso.

Una volta ridotta la matrice incompleta nella forma a scala, osserviamo che il numero r dei pivot è, come è facile provare, il rango della matrice incompleta. Se indichiamo con n il numero di incognite effettive (cioè trascurando quelle corrispondenti alle colonne formate da tutti zeri) e, al solito, con m il numero di equazioni, si può concludere che il sistema è compatibile se e solo se in corrispondenza alle $m-r$ righe finali della matrice incompleta composte da tutti zeri, anche i corrispondenti termini noti sono nulli. Le incognite corrispondenti alle colonne in cui non compaiono i pivot sono completamente arbitrarie e i corrispondenti termini possono essere "portati a secondo membro": il sistema avrà ∞^{n-r} soluzioni. La risoluzione è a questo punto immediata: dall'ultima delle r equazioni non identicamente nulle si ricava l'incognita corrispondente al pivot p_r; si effettua una sostituzione nell'equazione precedente e si ricava l'incognita corrispondente al pivot p_{r-1}; si procede così "all'indietro" fino a trovare tutte le incognite corrispondenti ai vari pivot.

Per ottenere la riduzione a scala della matrice incompleta si procede secondo il seguente schema, operando anche sui termini noti, cioè direttamente sulla matrice completa.

— Si eliminano eventuali colonne tutte nulle (che comunque nelle situazioni comuni non sono presenti).
— Si opera, eventualmente con uno scambio di righe, in modo che sulla prima riga, in corrispondenza alla prima colonna non nulla, ci sia un numero diverso da zero: questo numero prende il nome di pivot numero 1 o p_1.
— Con una opportuna combinazione lineare tra righe si fa in modo che dalla seconda riga in poi, sotto al pivot, compaiano solo zeri. Per fare questo si usa fare la seguente combinazione lineare: se la riga i ha, sotto al pivot p_1, l'elemento a non nullo, si sostituisce la riga i con la combinazione lineare ottenuta sommando la riga i stessa con la riga 1 moltiplicata per $-a/p_1$. È chiaro che non è l'unica scelta possibile, ma questa ha il vantaggio, se la matrice è quadrata, di non modificare il determinante della matrice. Se si tiene conto che l'operazione precedente può al più, sempre per le matrici quadrate, cambiare di segno il determinante, si capisce come questa tecnica possa essere vantaggiosamente applicata anche al calcolo dei determinanti.
— Si ripetono le operazioni indicate, considerando la matrice che si ottiene da quella così trasformata sopprimendo la prima riga. Successivamente si opera con le righe successive.

Esempio 15.8. Risolvere il sistema

$$\begin{cases} x_2+x_3+2x_4=-1 \\ -x_1-2x_2+2x_3=0 \\ x_1+3x_2+x_3-x_4=-2 \end{cases}.$$

[4]In pratica una colonna tutta nulla significa che la corrispondente incognita non compare nel sistema: nelle situazioni comuni questo fatto non succede mai, ma abbiamo voluto comunque metterci nella situazione più generale possibile.

Scriviamo la matrice completa del sistema.

$$A|b = \left(\begin{array}{cccc|c} 0 & 1 & 1 & 2 & -1 \\ -1 & -2 & 2 & 0 & 0 \\ 1 & 3 & 1 & -1 & -2 \end{array}\right)..$$

Operiamo ora secondo le indicazioni fornite, evidenziando i pivot man man che si ottengono.

$$\left(\begin{array}{cccc|c} 0 & 1 & 1 & 2 & -1 \\ -1 & -2 & 2 & 0 & 0 \\ 1 & 3 & 1 & -1 & -2 \end{array}\right); \quad r_1 \leftrightarrow r_3: \quad \left(\begin{array}{cccc|c} \boxed{1} & 3 & 1 & -1 & -2 \\ -1 & -2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 \end{array}\right); \quad r_2 \to r_2 + r_1:$$

$$\left(\begin{array}{cccc|c} \boxed{1} & 3 & 1 & -1 & -2 \\ 0 & \boxed{1} & 3 & -1 & -2 \\ 0 & 1 & 1 & 2 & -1 \end{array}\right); \quad r_3 \to r_3 + (-1)r_2: \left(\begin{array}{cccc|c} \boxed{1} & 3 & 1 & -1 & -2 \\ 0 & \boxed{1} & 3 & -1 & -2 \\ 0 & 0 & \boxed{-2} & 3 & 1 \end{array}\right).$$

Posto $x_4 = t$, il sistema si può riscrivere nella forma

$$\begin{cases} x_1 + 3x_2 + x_3 = t - 2 \\ x_2 + 3x_3 = t - 2 \\ -2x_3 = -3t + 1 \end{cases}.$$

Le $\infty^{4-3} = \infty^1$ soluzioni sono allora

$$x_1 = 10t, \quad x_2 = -\frac{1}{2} - \frac{7}{2}t, \quad x_3 = -\frac{1}{2} + \frac{3}{2}t, \quad x_4 = t.$$

Esempio 15.9. Risolvere il sistema di matrice completa

$$A|b = \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & -5 & 2 & 1 \\ 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & 2 \\ 3 & 0 & -2 & -1 \end{array}\right).$$

Operiamo ora secondo le indicazioni fornite, evidenziando i pivot man man che si ottengono.

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & 0 \\ 0 & -5 & 2 & 1 \\ 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & 2 \\ 3 & 0 & -2 & -1 \end{array}\right); \begin{array}{l} r_3 \to r_3 + (-2)r_1 \\ r_4 \to r_4 + r_1 \\ r_5 \to r_5 + (-3)r_1 \end{array} : \left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & 0 \\ 0 & -5 & 2 & 1 \\ 0 & -5 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & -6 & 1 & -1 \end{array}\right); r_4 \leftrightarrow r_2:$$

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & 0 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & -5 & 2 & 1 \\ 0 & -5 & 2 & 1 \\ 0 & -6 & 1 & -1 \end{array}\right); \begin{array}{l} r_3 \to r_3 + 5r_2 \\ r_4 \to r_4 + 5r_2 \\ r_5 \to r_5 + 6r_2 \end{array} :$$

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & 0 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & \boxed{7} & 11 \\ 0 & 0 & 7 & 11 \\ 0 & 0 & 7 & 11 \end{array}\right); \quad \begin{array}{l} r_4 \to r_4 + (-1)r_3 \\ r_5 \to r_5 + (-1)r_3 \end{array} : \quad \left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & 0 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & \boxed{7} & 11 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

La risoluzione è a questo punto immediata. Si ottiene

$$x_1 = \frac{5}{7}, \quad x_2 = \frac{3}{7}, \quad x_3 = \frac{11}{7}.$$

15.3. Vettori nello spazio ordinario

15.3.1. Definizioni

Consideriamo nello spazio ordinario le coppie ordinate di punti (A,B), ove A è detto primo estremo e B secondo estremo. Le coppie (A,B) sono anche dette *segmenti orientati* o *vettori applicati* e indicate con $\overrightarrow{\mathsf{AB}}$, o con $\mathsf{B}-\mathsf{A}$[5]. La lunghezza del segmento orientato $\overrightarrow{\mathsf{AB}}$ (rispetto ad una prefissata unità di misura) indica la distanza tra i punti A e B e si rappresenta, oltreché con $|\overline{\mathsf{AB}}|$, anche con $|\overrightarrow{\mathsf{AB}}|$. La direzione della retta individuata da A e B si chiama direzione del segmento orientato $\overrightarrow{\mathsf{AB}}$. Il segmento orientato individua anche, sulla retta AB, un verso: quello in cui A precede B.

Nell'insieme dei segmenti orientati si introduce una relazione di equivalenza, detta di *equipollenza*:

$$\overrightarrow{\mathsf{AB}} \text{ è equipollente a } \overrightarrow{\mathsf{CD}}$$

se

- $|\overline{\mathsf{AB}}| = |\overline{\mathsf{CD}}|$;
- le rette AB e CD sono parallele;
- i versi di $\overrightarrow{\mathsf{AB}}$ e $\overrightarrow{\mathsf{CD}}$ sono uguali.

Un modo equivalente, e più compatto, per dare la definizione di equipollenza è il seguente:

$$\overrightarrow{\mathsf{AB}} \text{ è equipollente a } \overrightarrow{\mathsf{CD}} \text{ se i punti medi di } \overline{\mathsf{AD}} \text{ e } \overline{\mathsf{BC}} \text{ coincidono.}$$

Trattandosi di una relazione di equivalenza, essa ripartisce l'insieme di tutti i segmenti orientati dello spazio in classi di equivalenza: come è d'uso, la classe di equivalenza individuata dal segmento orientato $\overrightarrow{\mathsf{AB}}$ è indicata con $[\overrightarrow{\mathsf{AB}}]$.

Si dà ora la seguente definizione:

[5] La notazione $\mathsf{B}-\mathsf{A}$ per indicare un segmento orientato è stata introdotta da William Rowan Hamilton (1805-1865), matematico irlandese. Si tratta di una notazione particolarmente felice e utile, come vedremo in seguito. Qui segnaliamo solo che la scrittura di un segmento orientato come *differenza di due punti* rende palese il diverso ruolo dei due estremi del segmento, esattamente come succede nella sottrazione ordinaria di numeri. Occorre tenere ben presente che da questa notazione non si può dedurre alcun concetto di somma di due punti: $\mathsf{B}-\mathsf{A}$ ha un ben preciso significato, nessun significato si attribuisce alla scrittura $\mathsf{B}+\mathsf{A}$.

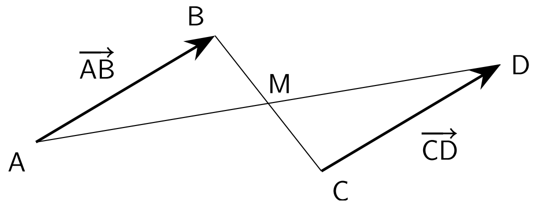


Figura 15.1.: *Segmenti orientati equipollenti*

Definizione 15.16 (Vettore). *Si chiama* vettore libero *o, semplicemente,* vettore *una classe di equivalenza di segmenti orientati equipollenti.*

I vettori saranno indicati con una lettera minuscola in grassetto, o sormontata da una freccia (come già fatto per le matrici ad una sola colonna, i vettori colonna, e vedremo in seguito il perché di questa coincidenza di notazioni), cioè si pone:

$$\boldsymbol{u} = \vec{u} = [\overrightarrow{AB}]. \tag{15.26}$$

È evidente che se $\overrightarrow{AB}$ e $\overrightarrow{CD}$ sono segmenti orientati equipollenti, si avrà

$$\vec{u} = [\overrightarrow{AB}] = [\overrightarrow{CD}].$$

Come è d'abitudine quando si lavora con classi di equivalenza, si può sempre considerare un elemento qualunque della classe stessa (un *rappresentante*) e operare direttamente con esso. Saranno dunque giustificate scritture del tipo:

$$\vec{u} = \overrightarrow{AB} = \overrightarrow{CD}, \tag{15.27}$$

nella quale $\overrightarrow{AB}$ e $\overrightarrow{CD}$ non sono pensati come segmenti orientati, ma come rappresentanti di classi di equivalenza.

Se poi si usa la notazione $B - A$ per i segmenti orientati, si può scrivere anche $\vec{u} = B - A$, da cui si trae

$$B = A + \vec{u}, \tag{15.28}$$

particolarmente significativa, in quanto precisa esattamente il significato profondo del concetto di vettore: il vettore $\vec{u}$ determina una *traslazione* dello spazio, che porta ogni punto P nel punto Q, ottenuto traslando P di un tratto uguale alla lunghezza di $\vec{u}$, nella direzione e verso di $\vec{u}$. È proprio da qui che deriva il nome: *vehere* significa infatti trasportare.

I segmenti $\overrightarrow{AA}$, $\overrightarrow{BB}$, ..., individuano il cosiddetto *vettore nullo*, indicato con $\vec{0}$ o $\mathbf{0}$.

Indichiamo con V_3 l'insieme di tutti i vettori (liberi) dello spazio, con S_3 l'insieme di tutti i punti dello spazio. È facile costruire una corrispondenza biunivoca tra V_3 e S_3: scelto un punto O dello spazio, basta associare ad ogni vettore $\vec{v}$ il punto $P = O + \vec{u}$.

Se invece di considerare i punti dello spazio, si considerano i punti di un piano π, si possono ripetere le stesse considerazioni, senza alcuna modifica. In questo caso si indicheranno con V_2 e S_2 rispettivamente i vettori e i punti del piano. Analogo discorso per i punti e i vettori di una retta, dove si useranno i simboli V_1 e S_1, con ovvio significato.

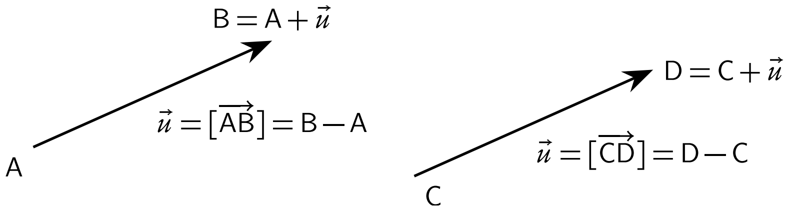


Figura 15.2.: *Vettori e traslazioni*

15.3.2. Operazioni lineari tra vettori

Nell'insieme V_3 dei vettori dello spazio si possono introdurre le operazioni di somma e prodotto per un numero. Per quanto riguarda la somma si possono dare due definizioni, perfettamente equivalenti.

Definizione 15.17 (Regola del parallelogramma)**.** *Dati due vettori $\vec{u}$ e $\vec{v}$, e considerati due rappresentanti aventi la stessa origine* A*,* $\vec{u} = B - A$ *e* $\vec{v} = D - A$*, si ha* $D + \vec{u} = B + \vec{v}$*. Posto* $C = D + \vec{u} = B + \vec{v}$*, si dice* somma *dei vettori $\vec{u}$ e $\vec{v}$ il vettore $\vec{w}$*

$$\vec{w} = \vec{u} + \vec{v} = C - A\,. \tag{15.29}$$

Definizione 15.18 (Regola del "testa-coda")**.** *Dati due vettori $\vec{u}$ e $\vec{v}$, e considerati due rappresentanti* $\vec{u} = B - A$ *e* $\vec{v} = C - B$*, si dice* somma *dei vettori $\vec{u}$ e $\vec{v}$ il vettore $\vec{w}$*

$$\vec{w} = \vec{u} + \vec{v} = C - A\,.$$

In entrambi i casi si prova che la scelta dei rappresentanti è ininfluente ai fini del risultato.

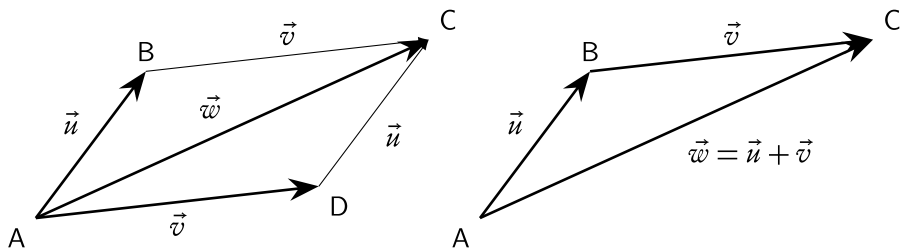


Figura 15.3.: *Somma di vettori: regole del parallelogramma e del "testa-coda"*

Definizione 15.19 (Opposto)**.** *Dato un vettore* $\vec{u} = B - A$ *si chiama suo* opposto *il vettore*

$$-\vec{u} = A - B\,, \tag{15.30}$$

cioè il vettore che ha lo stesso modulo, la stessa direzione e verso opposto.

Definizione 15.20 (Differenza di due vettori)**.** *Dati due vettori $\vec{u}$ e $\vec{v}$, si chiama loro* differenza *il vettore*

$$\vec{w} = \vec{u} + \left(-\vec{v}\right) = \vec{u} - \vec{v}\,. \tag{15.31}$$

Se $\vec{u} = \mathsf{B} - \mathsf{A}$ e $\vec{v} = \mathsf{D} - \mathsf{A}$, si ha $-\vec{v} = \mathsf{A} - \mathsf{D}$, da cui

$$\vec{u} - \vec{v} = (\mathsf{B} - \mathsf{A}) + (\mathsf{A} - \mathsf{D}) = (\mathsf{B} - \mathsf{A}) - (\mathsf{D} - \mathsf{A}) = \mathsf{B} - \mathsf{D}.$$

Anche se nella precedente uguaglianza non si devono sciogliere le parentesi applicando le usuali regole dei segni (si otterrebbe una somma di punti che non abbiamo definito), tutto funziona come se si "semplificasse il punto A".

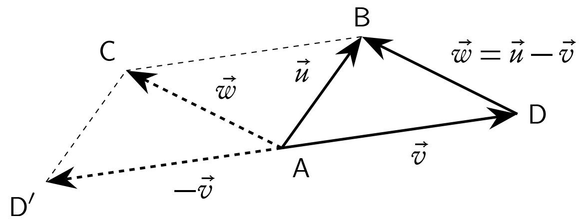


Figura 15.4.: *Differenza di vettori*

Definizione 15.21 (Prodotto di un vettore per un numero). *Dato un vettore $\vec{u}$ e un numero reale λ, si chiama* prodotto del vettore $\vec{u}$ per λ *il vettore $\vec{w}$ così definito:*

1. *se $\lambda = 0$ oppure $\vec{u} = \vec{0}$, $\vec{w} = \lambda\vec{u} = \vec{0}$;*
2. *se $\lambda \neq 0$ e $\vec{u} \neq \vec{0}$, $\vec{w} = \lambda\vec{u}$ è il vettore che ha*
   - *modulo uguale al modulo di λ per il modulo di $\vec{u}$, $|\lambda\vec{u}| = |\lambda||\vec{u}|$;*
   - *direzione uguale a quella di $\vec{u}$;*
   - *verso concorde o discorde a quello di $\vec{u}$ a seconda che λ sia positivo o negativo.*

Le operazioni di somma e prodotto per un numero appena introdotte godono delle seguenti proprietà, per ogni $\vec{u}$, $\vec{v}$, $\vec{w}$ e per ogni λ, μ:

1. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$: proprietà associativa della somma;
2. $\vec{v} + \vec{0} = \vec{v}$: esistenza dell'elemento neutro della somma;
3. $\vec{v} + (-\vec{v}) = \vec{0}$: esistenza dell'opposto;
4. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$: proprietà commutativa della somma;
5. $\lambda(\mu\vec{u}) = (\lambda\mu)\vec{u}$;
6. $(\lambda + \mu)\vec{u} = \lambda\vec{u} + \mu\vec{u}$;
7. $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$
8. $1\vec{u} = \vec{u}$;
9. $-1\vec{u} = -\vec{u}$;
10. $0\vec{u} = \vec{0}$.

Si noti che l'operazione di somma tra due vettori è un'operazione interna nell'insieme V_3: a una coppia di vettori fa corrispondere un terzo vettore; per contro l'operazione di prodotto per un numero è un'operazione esterna: ad una coppia costituita da un numero e un vettore fa corrispondere un vettore.

L'insieme V_3 con le operazioni ora introdotte è un esempio di una struttura algebrica di grande importanza in tutte le applicazioni, detta *spazio vettoriale*.

Definizione 15.22 (Combinazione lineare di vettori). *Dati n vettori $\vec{v}_1$, $\vec{v}_2$, …, $\vec{v}_n$ e n numeri reali λ_1 λ_2, …, λ_n, si chiama* combinazione lineare *dei vettori dati, con coefficienti i numeri reali dati, il vettore*

$$\vec{w} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_n \vec{v}_n \,. \tag{15.32}$$

È facile provare che, nel caso di due soli vettori $\vec{u}$ e $\vec{v}$, detta $\vec{w}$ una loro combinazione lineare arbitraria, e considerato un punto O dello spazio, il punto $\mathsf{P} = \mathsf{O} + \vec{w}$ è sempre complanare con O, $\mathsf{A} = \mathsf{O} + \vec{u}$ e $\mathsf{B} = \mathsf{O} + \vec{v}$. Se poi $\vec{u}$ e $\vec{v}$ sono paralleli, i quattro punti considerati sono addirittura allineati.

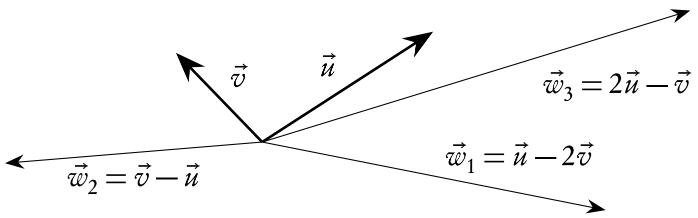


Figura 15.5.: *Combinazioni lineari di due vettori*

15.3.3. Prodotto scalare

Dati due vettori $\vec{u}$ e $\vec{v}$ e un punto O, consideriamo i punti $\mathsf{A} = \mathsf{O} + \vec{u}$ e $\mathsf{B} = \mathsf{O} + \vec{v}$. Si chiama *angolo* tra i due vettori l'angolo convesso (eventualmente piatto) individuato dalle semirette OA e OB.

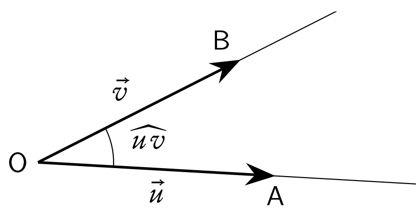


Figura 15.6.: *Angolo tra due vettori*

Consideriamo di nuovo due vettori $\vec{u}$ e $\vec{v}$, un punto O e i punti $\mathsf{A} = \mathsf{O} + \vec{u}$ e $\mathsf{B} = \mathsf{O} + \vec{v}$. Indicheremo con A′ la proiezione ortogonale di A sulla retta OB e con B′ la proiezione ortogonale di B sulla retta OA. Si può facilmente provare che

- A′ e B′ stanno sulla semiretta OB e OA, rispettivamente, se $\widehat{uv} < \pi/2$;
- A′ e B′ coincidono con O se $\widehat{uv} = \pi/2$;
- A′ e B′ stanno sulle semirette opposte a OB e OA, rispettivamente, se $\widehat{uv} > \pi/2$.

Definizione 15.23 (Proiezione di un vettore su un altro). *Le lunghezze dei segmenti $\overline{\mathsf{OA}'}$ e $\overline{\mathsf{OB}'}$, prese col segno + o − a seconda che $\widehat{uv} < \pi/2$ oppure $\pi/2 < \widehat{uv} \le \pi$ si chiamano* proiezioni *di $\vec{u}$ su $\vec{v}$ (rispettivamente di $\vec{v}$ su $\vec{u}$), e si indicano con u_v e v_u rispettivamente.*

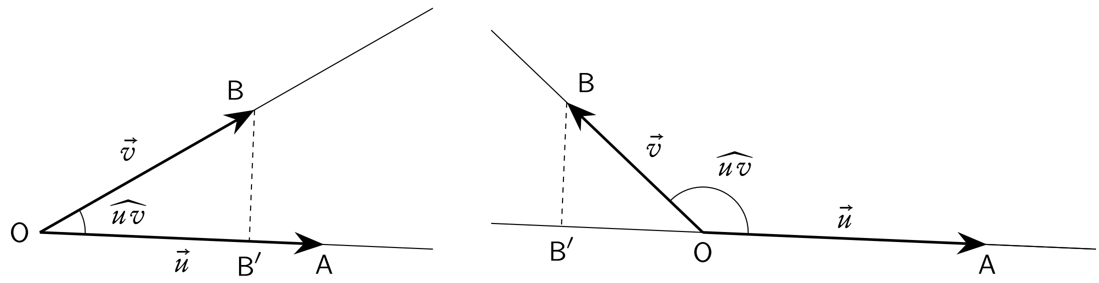


Figura 15.7.: *Proiezione di un vettore su un altro*

Definizione 15.24 (Prodotto scalare di due vettori)**.** *Dati due vettori $\vec{u}$ e $\vec{v}$, si chiama loro* prodotto scalare, *e si indica con*

$$\vec{u}\cdot\vec{v}\,, \tag{15.33}$$

il numero reale definito in uno dei seguenti tre modi equivalenti:

— $\vec{u}\cdot\vec{v}=|\vec{u}|\,|\vec{v}|\,\cos(\widehat{uv})$.
— $\vec{u}\cdot\vec{v}=|\vec{u}|\,v_u$.
— $\vec{u}\cdot\vec{v}=|\vec{v}|\,u_v$.

Il prodotto scalare di due vettori non nulli è nullo se e soltanto se i due vettori sono *ortogonali*.

Osservazione 15.25 (Notazioni sul prodotto scalare)*.* Sono in uso diverse notazioni per il prodotto scalare, oltre a quella qui adottata. Tra le altre citiamo:

$$\vec{u}\times\vec{v}\,,\quad \langle\vec{u},\vec{v}\rangle\,,\quad \langle\vec{u}|\vec{v}\rangle\,.$$

La prima di queste è da sconsigliare in quanto, in particolare secondo le norme ISO, è utilizzata per il prodotto vettoriale, che definiremo tra poco. L'ultima è particolarmente usata dai fisici nei testi di meccanica quantistica, ed è nota come notazione di Dirac: la parte sinistra del simbolo, $\langle\vec{u}|$, si chiama *vettore bra*, la parte di destra, $|\vec{v}\rangle$, *vettore ket*, il simbolo completo (che denota in generale uno stato), si chiama *braket*.

Il prodotto scalare gode delle seguenti proprietà, per ogni $\vec{u},\vec{v},\vec{w}$ e per ogni λ:

— $\vec{u}\cdot\vec{v}=\vec{v}\cdot\vec{u}$: proprietà commutativa;
— $(\lambda\vec{u})\cdot\vec{v}=\vec{u}\cdot(\lambda\vec{v})=\lambda(\vec{u}\cdot\vec{v})$
— $\vec{u}\cdot(\vec{v}+\vec{w})=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w}$: proprietà distributiva;
— $\vec{u}\cdot\vec{v}=0 \iff \vec{u}\perp\vec{v}$ (con la convenzione che un vettore nullo possa essere considerato perpendicolare a ogni altro vettore). Naturalmente parlando di perpendicolarità tra due vettori intendiamo riferirci a due rappresentanti dei vettori aventi la stessa origine.

Si noti che l'operazione di prodotto scalare è un'operazione esterna nell'insieme V_3: ad una coppia di vettori fa corrispondere un numero reale.

15.3.4. Prodotto vettoriale

La definizione di prodotto vettoriale non è così semplice come le altre tre operazioni introdotte nell'insieme dei vettori dello spazio e richiede l'uso di un concetto (quello di *verso orario* o *antiorario*) di non facile spiegazione formale, anche se intuitivamente evidente. Daremo quindi solo una definizione "elementare" di questa importante operazione.

È però molto importante segnalare subito una differenza fondamentale con le operazioni precedenti, in particolare le operazioni lineari: una combinazione lineare di due vettori paralleli è ancora un vettore parallelo ai dati, una combinazione lineare di due vettori è un vettore complanare ai vettori dati. Ciò significa, come abbiamo già osservato, che si potrebbe anche operare, senza cambiare nulla, in V_1 o V_2, anziché in V_3. Il prodotto vettoriale, come noi lo definiremo, è invece una operazione *intrinsecamente tridimensionale*, cioè non ha senso in V_1 o V_2.

Definizione 15.26 (Prodotto vettoriale o esterno). *Dati due vettori $\vec{u}$ e $\vec{v}$ si dice loro* prodotto vettoriale o esterno *il vettore $\vec{w}$, che si indica con $\vec{u} \times \vec{v}$, e si legge $\vec{u}$ vettore $\vec{v}$ o $\vec{u}$ esterno $\vec{v}$, definito come segue:*

— *se $\vec{u}$ e $\vec{v}$ sono paralleli $\vec{u} \times \vec{v} = \vec{0}$;*
— *se $\vec{u}$ e $\vec{v}$ non sono paralleli*
  - *il suo modulo è dato da $|\vec{w}| = |\vec{u} \times \vec{v}| = |\vec{u}||\vec{v}| \sin \widehat{\vec{u}\vec{v}}$*
  - *la direzione è perpendicolare sia a $\vec{u}$ che a $\vec{v}$;*
  - *il verso è quello di avanzamento di una vite detrorsa (cavatappi) che ruoti nel senso in cui $\vec{u}$ ruota per sovrapporsi a $\vec{v}$, compiendo il minimo angolo.*

È immediato che il modulo di $\vec{u} \times \vec{v}$ è uguale all'area del parallelogramma di lati consecutivi $\overline{\mathsf{AB}}$ e $\overline{\mathsf{AC}}$, dove A è un punto qualunque e $\mathsf{B} = \mathsf{A} + \vec{u}$ e $\mathsf{C} = \mathsf{A} + \vec{v}$. Per quanto riguarda il verso si può anche, in maniera equivalente (ma sempre un po' azzardata dal punto di vista del rigore), dire che il verso è quello testa-piedi di un osservatore che, posto sul piano per i punti A, B, C appena considerati, veda la minima rotazione di $\vec{u}$ per sovrapporsi a $\vec{v}$ avvenire in senso antiorario, oppure ancora è il verso indicato dal pollice della mano destra se il palmo della stessa mano compie la minima rotazione che porta $\vec{u}$ a sovrapporsi a $\vec{v}$.

Il prodotto vettoriale di due vettori gode delle seguenti proprietà, per ogni $\vec{u}, \vec{v}, \vec{w}$ e per ogni λ:

— $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$: proprietà *anticommutativa*;
— $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$: proprietà distributiva;
— $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$: proprietà distributiva;
— $(\lambda\vec{u}) \times \vec{v} = \vec{u} \times (\lambda\vec{v}) = \lambda(\vec{u} \times \vec{v})$;
— $\vec{u} \times \vec{v} = \vec{0} \Leftrightarrow \vec{u} \parallel \vec{v}$ (con la convenzione che un vettore nullo possa essere considerato parallelo a ogni altro vettore).

Si noti che l'operazione di prodotto vettoriale è un'operazione interna all'insieme V_3: ad una coppia di vettori fa corrispondere un terzo vettore.

Si tenga presente che la notazione qui adottata per il prodotto vettoriale, che, al solito, è quella prevista dalla normativa ISO, non è l'unica possibile. In particolare in molti testi in lingua italiana è più diffusa la notazione $\vec{u} \wedge \vec{v}$. In ogni caso, leggendo un testo, è sempre bene controllare le notazioni usate, o consultando l'apposita tabella (se presente), oppure controllando le convenzioni usate in occasione del primo uso di un simbolo.

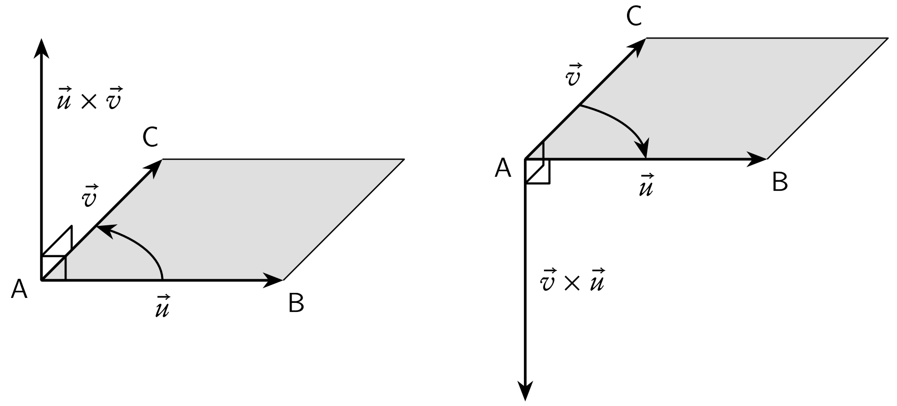


Figura 15.8.: *Prodotto vettoriale di due vettori*

È molto importante prestare attenzione al fatto che il prodotto vettoriale non gode della proprietà associativa, per cui, ad esempio, dati tre vettori $\vec{u}$, $\vec{v}$, $\vec{w}$, in generale

$$(\vec{u} \times \vec{v}) \times \vec{w} \neq \vec{u} \times (\vec{v} \times \vec{w}). \tag{15.34}$$

Un'espressione del tipo di quelle considerate si chiama un *doppio prodotto vettoriale*. La non associatività del prodotto vettoriale risulta chiaramente dalla figura 15.9, dove si è introdotto un sistema cartesiano ortogonale e sono state evidenziate anche le coordinate di alcuni punti, per rendere la figura stessa più leggibile.

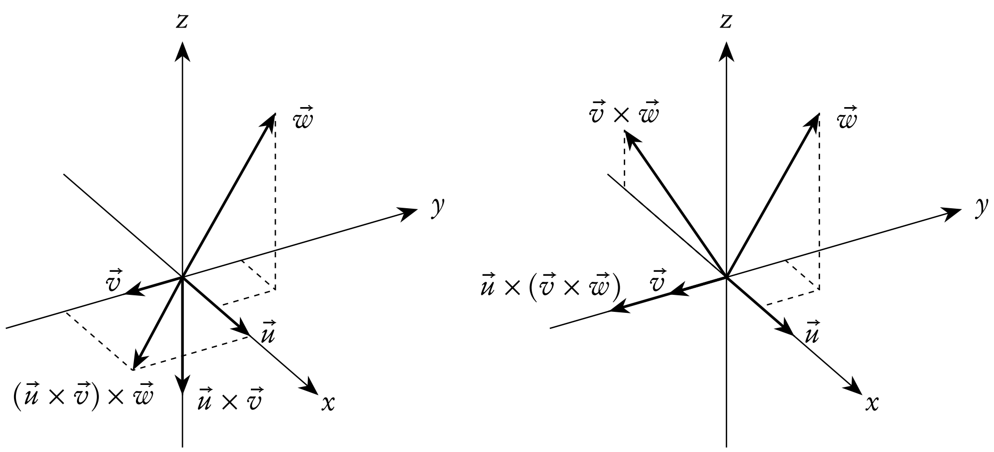


Figura 15.9.: *Doppio prodotto vettoriale*

15.3.5. Prodotto misto

Dati tre vettori $\vec{u}$, $\vec{v}$, $\vec{w}$, e considerato il prodotto vettoriale, $\vec{z}$, di due dei tre, ha senso calcolare il prodotto scalare di $\vec{z}$ con il terzo vettore, per esempio $(\vec{u} \times \vec{v}) \cdot \vec{w}$. In considerazione delle caratteristiche dei due prodotti, le parentesi sono inutili: nella scrittura $\vec{u} \times \vec{v} \cdot \vec{w}$ si deve eseguire prima il prodotto vettoriale e poi quello scalare, altrimenti la scrittura sarebbe priva di senso. Il prodotto $\vec{u} \times \vec{v} \cdot \vec{w}$ si chiama *prodotto misto*.

Si prova facilmente che il modulo del prodotto misto di tre vettori uguaglia il volume del prisma costruito sui tre vettori, come nella figura 15.10: basta solo tenere conto che il prodotto vettoriale ha per modulo l'area del parallelogramma "di base", mentre il successivo prodotto scalare (a parte il segno) rappresenta il prodotto tra questa area di base e l'altezza.

Dal fatto che il prodotto misto rappresenta, a meno del segno, il volume del prisma costruito sui tre vettori si possono dedurre le seguenti proprietà:

— il modulo del prodotto misto non dipende dall'ordine in cui i vettori sono scritti, né dall'ordine in cui si eseguono i due prodotti, ovvero

$$|\vec{u} \times \vec{v} \cdot \vec{w}| = |\vec{u} \cdot \vec{v} \times \vec{w}| = |\vec{w} \cdot \vec{u} \times \vec{v}| = \ldots;$$

— il prodotto misto è nullo se e solo se i tre vettori sono complanari, con la convenzione di considerare complanari tre vettori di cui uno o più siano nulli. Naturalmente parlando di complanarità di tre vettori intendiamo riferirci a tre rappresentanti dei vettori che abbiano la stessa origine.

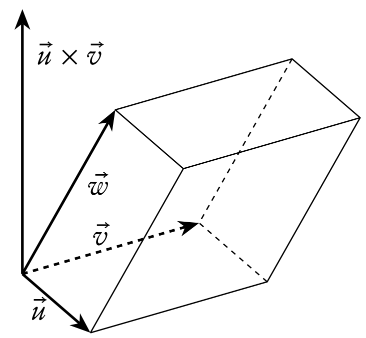


Figura 15.10.: *Prodotto misto di tre vettori*

Si può anche provare che il prodotto misto di tre vettori non cambia se si scambia il prodotto vettoriale con quello scalare oppure si permutano circolarmente i tre vettori.

$$\vec{u} \times \vec{v} \cdot \vec{w} = \vec{u} \cdot \vec{v} \times \vec{w}; \tag{15.35}$$

$$\vec{u} \times \vec{v} \cdot \vec{w} = \vec{v} \times \vec{w} \cdot \vec{u} = \vec{w} \times \vec{u} \cdot \vec{v}. \tag{15.36}$$

In sostanza quello che conta nel prodotto misto è, a meno di una permutazione circolare, l'ordine in cui i vettori sono scritti.

15.3.6. Parallelismo, perpendicolarità, complanarità

Considerata l'importanza dei concetti di parallelismo, perpendicolarità, complanarità, richiamiamo qui le relazioni, già menzionate, che intercorrono tra essi e le operazioni tra vettori.

— Due vettori sono *paralleli* se e soltanto se il loro *prodotto vettoriale* è nullo.
— Due vettori sono *perpendicolari* se e soltanto se il loro *prodotto scalare* è nullo.
— Tre vettori sono *complanari* se e soltanto se il loro *prodotto misto* è nullo.

In tutti i casi si comprende la possibilità che uno o più dei vettori sia nullo, con la convenzione che il vettore nullo sia parallelo oppure perpendicolare a ogni altro vettore, e che sia complanare a ogni altra coppia di vettori.

Vedremo come la verifica del parallelismo, perpendicolarità, o complanarità di rette e piani si possa fare proprio tenendo conto di queste proprietà.

15.4. Coordinate cartesiane, vettori e componenti

Abbiamo già trattato, nel capitolo 8, l'introduzione di un sistema cartesiano ortogonale e monometrico nel piano e nello spazio. La figura 15.11 riassume lo schema del ragionamento.

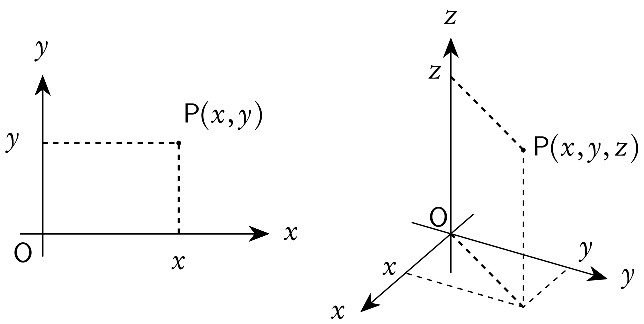


Figura 15.11.: *Coordinate cartesiane nel piano e nello spazio*

Riportiamo anche le formule per il punto medio di un segmento, il baricentro di un triangolo e la distanza tra due punti, analoghe alle (8.1), (8.2), (8.3) del capitolo 8 che si riferivano al caso del piano.

$$x_\mathsf{M}=\frac{x_1+x_2}{2},\quad y_\mathsf{M}=\frac{y_1+y_2}{2},\quad z_\mathsf{M}=\frac{z_1+z_2}{2}. \tag{15.37}$$

$$x_\mathsf{G}=\frac{x_1+x_2+x_3}{2},\quad y_\mathsf{G}=\frac{y_1+y_2+y_3}{2},\quad z_\mathsf{G}=\frac{z_1+z_2+z_3}{2}. \tag{15.38}$$

$$d(\mathsf{A},\mathsf{B})=|\overline{\mathsf{AB}}|=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}. \tag{15.39}$$

15.4.1. Componenti di vettori nel piano e nello spazio

Al sistema $\mathsf{O}xyz$ si può associare una *base* dello spazio V_3 dei vettori, nel senso e con le modalità che ora indicheremo.

Abbiamo già considerato, nella definizione 15.22, la *combinazione lineare* di più vettori. Ci poniamo ora il problema inverso: dati n vettori $\vec{v}_1$, $\vec{v}_2$, …, $\vec{v}_n$ è in generale possibile esprimere un arbitrario vettore $\vec{u}$ come loro combinazione lineare? In caso affermativo i coefficienti della combinazione sono unici o no? La risposta è negativa in entrambi i casi, come si può vedere sui due esempi che seguono.

1. Se i vettori dati sono solo due, $\vec{v}_1$ e $\vec{v}_2$, e non sono paralleli, consideriamo un punto A e il piano passante per A, $\mathsf{B} = \mathsf{A} + \vec{v}_1$, $\mathsf{C} = \mathsf{A} + \vec{v}_2$. Sappiamo già che se $\vec{u}$ è una qualunque combinazione lineare di $\vec{v}_1$ e $\vec{v}_2$, il punto $\mathsf{D} = \mathsf{A} + \vec{u}$ è complanare con A, B, C; dunque un vettore $\vec{w} = \mathsf{E} - \mathsf{A}$, con E non complanare con A, B, C, non potrà essere combinazione lineare di $\vec{v}_1$ e $\vec{v}_2$.
2. Se i vettori dati sono tre, $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, e sono complanari, è facile provare che ogni altro vettore $\vec{u}$ complanare con i tre si può esprimere in infiniti modi come combinazione lineare dei tre vettori dati. In altri termini si può dire che $\vec{u}$ si può *decomporre* in infiniti modi nella somma di tre vettori paralleli ai tre vettori dati. Si veda la figura 15.12.

Servendosi anche della griglia presente nella figura 15.12 è facile provare che si ha, per esempio:

$$\vec{u} = \frac{3}{2}\vec{v}_1 + \frac{2}{3}\vec{v}_2 + 0\vec{v}_3 = 0\vec{v}_1 + \frac{2}{3}\vec{v}_2 - \vec{v}_3 = -3\vec{v}_1 + \frac{2}{3}\vec{v}_2 - 3\vec{v}_3\,.$$

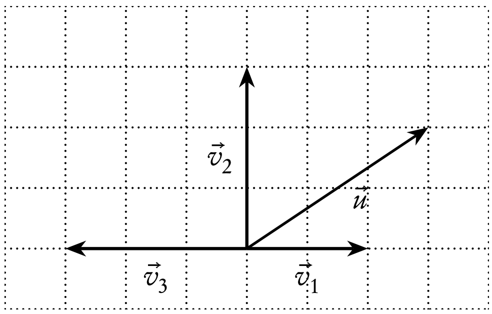


Figura 15.12.: *Non unicità della scomposizione di un vettore*

In generale non è difficile provare che se si considerano nel piano due vettori non paralleli oppure nello spazio tre vettori non complanari, ogni vettore rispettivamente del piano o dello spazio è *esprimibile in maniera unica* come loro combinazione lineare. Un tal insieme di vettori si chiama una base, rispettivamente per V_2 o per V_3. Si dà cioè la seguente

Definizione 15.27 (Base per V_2 o per V_3). *Un insieme di vettori, rispettivamente di V_2 o V_3, è una* base *per V_2 o V_3 se ogni altro vettore $\vec{u}$ di V_2 o, rispettivamente, V_3 può essere espresso come loro combinazione lineare in maniera unica.*

Si prova che una base di V_2 è necessariamente costituita da 2 vettori non paralleli, una base di V_3 è necessariamente costituita da 3 vettori non complanari.

Dunque:

— In V_2 dati 2 vettori $\vec{v}_1$ e $\vec{v}_2$ non paralleli, per ogni altro vettore $\vec{u}$ vale la scomposizione

$$\vec{u} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2\,.$$

— In V_3 dati 3 vettori $\vec{v}_1$, $\vec{v}_2$ e $\vec{v}_3$ non complanari, per ogni altro vettore $\vec{u}$ vale la scomposizione

$$\vec{u} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \lambda_3\vec{v}_3\,.$$

I numeri λ_1, λ_2, λ_3 si chiamano *coordinate* o *componenti* del vettore $\vec{u}$ rispetto alla base data; i vettori $\lambda_1\vec{v}_1$, $\lambda_2\vec{v}_2$, $\lambda_3\vec{v}_3$ si chiamano *vettori componenti* del vettore $\vec{u}$ rispetto alla base data. Si scrive anche $\vec{u} = (\lambda_1, \lambda_2, \lambda_3)$.

Se nello spazio S_3 è dato un sistema di coordinate cartesiane ortogonali monometrico, si può, in maniera naturale, associare a esso una base di V_3 scegliendo tre vettori di modulo 1 paralleli ed equiversi ai tre assi coordinati. Tre vettori come quelli indicati si indicano[6] $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$.

Facendo una scelta come questa si ha la seguente importante conseguenza, di facile verifica:

Teorema 15.28 (Componenti dei vettori e coordinate dei punti)**.** *Dato nello spazio un sistema cartesiano ortogonale monometrico* $\mathsf{O}xyz$ *e considerata la base* $\vec{e}_1, \vec{e}_2, \vec{e}_3$ *di* V_3 *naturalmente associata a esso, le componenti di ogni vettore* $\vec{u}$ *e le coordinate del punto* $\mathsf{P} = \mathsf{O} + \vec{u}$ *coincidono: se* $\mathsf{P} = (x_\mathsf{P}, y_\mathsf{P}, z_\mathsf{P})$ *anche* $\vec{u} = (x_\mathsf{P}, y_\mathsf{P}, z_\mathsf{P})$.

Considerati poi un punto A *arbitrario e il punto* $\mathsf{B} = \mathsf{A} + \vec{u}$*, si ha*

$$\vec{u} = (x_\mathsf{B} - x_\mathsf{A}, y_\mathsf{B} - y_\mathsf{A}, z_\mathsf{B} - z_\mathsf{A}), \tag{15.40}$$

con ovvio significato dei simboli.

In S_2 e V_2 valgono considerazioni identiche, con una coordinata in meno.

L'introduzione delle componenti facilita grandemente la trattazione di tutti i problemi connessi ai vettori: come vedremo, le operazioni introdotte sui vettori possono essere eseguite lavorando direttamente sulle componenti. Inoltre l'uso delle componenti consente facili generalizzazioni dei concetti introdotti nello spazio ordinario a casi di spazi con un numero arbitrario di dimensioni.

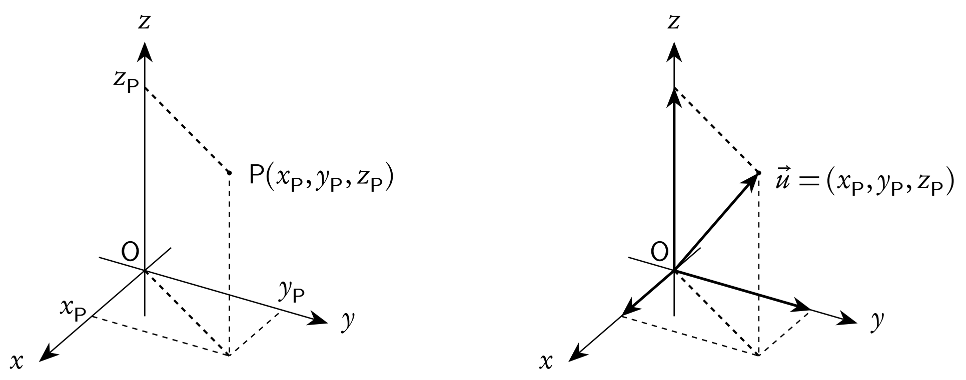


Figura 15.13.: *Coordinate cartesiane di un punto e componenti di un vettore*

Osservazione 15.29 (Coordinate, componenti, vettori riga e vettori colonna)*.* I concetti di coordinate di un punto e di componenti di un vettore ci hanno permesso di associare ai punti dello spazio e ai vettori liberi *terne* di numeri reali. Vedremo fra breve che le ordinarie operazioni tra vettori si possono eseguire mediante operazioni sulle terne, come fatto nelle operazioni tra matrici. Per uniformità di linguaggio le terne associate ai punti e ai vettori andrebbero scritte come *vettori colonna*, ma ciò comporterebbe ovvie

[6]La scrittura proposta, che è quella prevista dalle norme ISO, è utile perché facilita l'estensione a più dimensioni; si usano comunque anche i simboli $\vec{i}$, $\vec{j}$, $\vec{k}$.

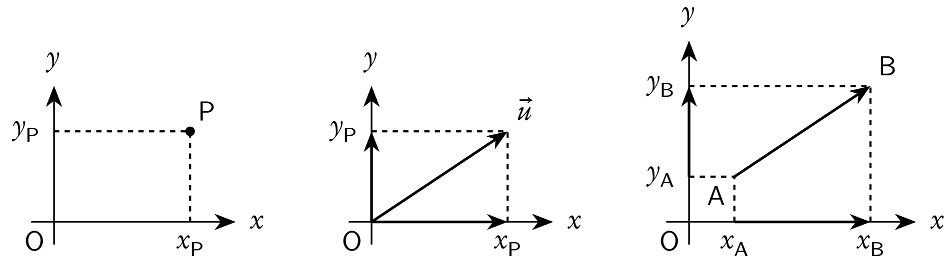


Figura 15.14.: *Coordinate cartesiane di punti e componenti di un vettore, nel piano*

difficoltà tipografiche. Si potrebbe far ricorso alla notazione con le matrici trasposte, ma si avrebbe un inutile appesantimento delle scritture. Si usa dunque rappresentare le terne di coordinate dei punti e dei vettori con *vettori riga*, tenendo però in mente che, quando si ha bisogno di usare la teoria delle matrici, queste terne sono in realtà *vettori colonna*.

15.4.2. Operazioni tra vettori, mediante le componenti

Supporremo sempre, nel seguito, di avere introdotto nello spazio un sistema cartesiano ortogonale monometrico $\mathrm{O}xyz$ e di avere associato a esso la base $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, di V_3, che chiameremo *base canonica*. Se $\vec{u}$ è un vettore di V_3 (diremo brevemente un vettore dello spazio), useremo le seguenti scritture, con ovvio significato dei simboli:

$$\vec{u} = u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3 = (u_1, u_2, u_3) = u_x\vec{e}_1 + u_y\vec{e}_2 + u_z\vec{e}_3 = (u_x, u_y, u_z) = \vec{u}_x + \vec{u}_y + \vec{u}_z\,. \tag{15.41}$$

Valgono le seguenti proprietà.

— *Somma di vettori*: $\vec{u} + \vec{v} = (u_1, u_2, u_3) + (v_1, v_2, v_3) = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$.
— *Prodotto di un vettore per un numero*: $c\vec{u} = c(u_1, u_2, u_3) = (cu_1, cu_2, cu_3)$.
— *Prodotto scalare di due vettori*: $\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2, u_3v_3$.

Si noti che il prodotto scalare si può pensare, usando il prodotto tra matrici, nel modo seguente:

$$\vec{u} \cdot \vec{v} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}^{\mathsf{T}} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1v_1 + u_2v_2 + u_3v_3 \end{pmatrix}, \tag{15.42}$$

dove l'ultima è una matrice a una sola riga e una sola colonna, ovvero un numero reale.

Un po' più complessa la situazione per quanto riguarda il prodotto vettoriale. Dati due vettori $\vec{u} = (u_1, u_2, u_3)$ e $\vec{v} = (v_1, v_2, v_3)$, si considera la seguente *matrice simbolica*:

$$\begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix}.$$

Il *prodotto vettoriale* dei due vettori si ottiene calcolando il determinante simbolico di questa matrice, effettuando lo sviluppo secondo la prima riga:

$$\vec{u}\times\vec{v}=\begin{vmatrix}\vec{e}_1 & \vec{e}_2 & \vec{e}_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\,.\end{vmatrix}=(u_2v_3-u_3v_2)\vec{e}_1-(u_1v_3-u_3v_1)\vec{e}_2+(u_1v_2-u_2v_1)\vec{e}_3=$$
$$=(u_2v_3-u_3v_2)\vec{e}_1+(u_3v_1-u_1v_3)\vec{e}_2+(u_1v_2-u_2v_1)\vec{e}_3\,. \tag{15.43}$$

È come dire che le componenti del prodotto vettoriale di due vettori sono i cofattori degli elementi della prima riga nella matrice simbolica sopra considerata.

A questo punto è facile provare che il *prodotto misto* di tre vettori è dato da

$$\vec{u}\cdot\vec{v}\times\vec{w}=\vec{u}\times\vec{v}\cdot\vec{w}=\begin{vmatrix}u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\\ w_1 & w_2 & w_3\end{vmatrix}. \tag{15.44}$$

Proponiamo ora qualche esempio applicativo.

Esempio 15.10. Supponiamo di dover calcolare l'angolo tra i due vettori $\vec{u}=(1,0,2)$ e $\vec{v}=(-1,1,1)$. Si ha (teorema di Pitagora)

$$|\vec{u}|=\sqrt{1^2+0^2+2^2}=\sqrt{5}\,,\quad |\vec{v}|=\sqrt{(-1)^2+1^2+1^2}=\sqrt{3}\,.$$

Dunque

$$\vec{u}\cdot\vec{v}=|\vec{u}||\vec{v}|\cos(\widehat{uv})=\sqrt{5}\sqrt{3}\cos(\widehat{uv})=1\cdot(-1)+0\cdot1+2\cdot1=1\,,$$

e quindi

$$\cos(\widehat{uv})=\frac{1}{\sqrt{15}}\quad\Rightarrow\quad \widehat{uv}=\arccos\left(\frac{1}{\sqrt{15}}\right)\approx 75^\circ\,.$$

Esempio 15.11. Calcoliamo i prodotti scalari e vettoriali tra le coppie di vettori della base canonica. Basta tenere conto delle condizioni di parallelismo e perpendicolarità e delle proprietà dei due prodotti per concludere che:

— $\vec{e}_1\cdot\vec{e}_1=1,\quad \vec{e}_2\cdot\vec{e}_2=1,\quad \vec{e}_3\cdot\vec{e}_3=1;$
— $\vec{e}_1\cdot\vec{e}_2=0,\quad \vec{e}_2\cdot\vec{e}_3=0,\quad \vec{e}_1\cdot\vec{e}_3=0;$
— $\vec{e}_1\times\vec{e}_1=0,\quad \vec{e}_2\times\vec{e}_2=0,\quad \vec{e}_3\times\vec{e}_3=0;$
— $\vec{e}_1\times\vec{e}_2=\vec{e}_3,\quad \vec{e}_2\times\vec{e}_3=\vec{e}_1,\quad \vec{e}_3\times\vec{e}_1=\vec{e}_2;$
— $\vec{e}_2\times\vec{e}_1=-\vec{e}_3,\quad \vec{e}_3\times\vec{e}_2=-\vec{e}_1,\quad \vec{e}_1\times\vec{e}_3=-\vec{e}_2.$

Esempio 15.12. Utilizzando i risultati dell'esempio precedente e le proprietà dei prodotti scalare e vettoriale, ritroviamo le formule per il calcolo di questi prodotti, mediante le componenti.

Dati due vettori $\vec{u}$ e $\vec{v}$ si ha:

$$\vec{u}=u_1\vec{e}_1+u_2\vec{e}_2+u_3\vec{e}_3\,,\quad \vec{v}=v_1\vec{e}_1+v_2\vec{e}_2+v_3\vec{e}_3\,.$$

Dunque

$$
\begin{aligned}
\vec{u}\cdot\vec{v} = {} & (u_1\vec{e}_1+u_2\vec{e}_2+u_3\vec{e}_3)\cdot(v_1\vec{e}_1+v_2\vec{e}_2+v_3\vec{e}_3) = \\
& = u_1v_1\vec{e}_1\cdot\vec{e}_1+u_1v_2\vec{e}_1\cdot\vec{e}_2+u_1v_3\vec{e}_1\cdot\vec{e}_3+u_2v_1\vec{e}_2\cdot\vec{e}_1+ \\
& \qquad + u_2v_2\vec{e}_2\cdot\vec{e}_2+u_2v_3\vec{e}_2\cdot\vec{e}_3+u_3v_1\vec{e}_3\cdot\vec{e}_1+u_3v_2\vec{e}_3\cdot\vec{e}_2+u_3v_3\vec{e}_3\cdot\vec{e}_3 = \\
& = u_1v_1+u_2v_2+u_3v_3 .
\end{aligned}
$$

Inoltre

$$
\begin{aligned}
\vec{u}\times\vec{v} = {} & (u_1\vec{e}_1+u_2\vec{e}_2+u_3\vec{e}_3)\times(v_1\vec{e}_1+v_2\vec{e}_2+v_3\vec{e}_3) = \\
& = u_1v_1\vec{e}_1\times\vec{e}_1+u_1v_2\vec{e}_1\times\vec{e}_2+u_1v_3\vec{e}_1\times\vec{e}_3+u_2v_1\vec{e}_2\times\vec{e}_1+ \\
& \qquad + u_2v_2\vec{e}_2\times\vec{e}_2+u_2v_3\vec{e}_2\times\vec{e}_3+u_3v_1\vec{e}_3\times\vec{e}_1+u_3v_2\vec{e}_3\times\vec{e}_2+u_3v_3\vec{e}_3\times\vec{e}_3 = \\
& = u_1v_2\vec{e}_3-u_1v_3\vec{e}_2-u_2v_1\vec{e}_3+u_2v_3\vec{e}_1+u_3v_1\vec{e}_2-u_3v_2\vec{e}_1 = \\
& = (u_2v_3-u_3v_2)\vec{e}_1-(u_1v_3-u_3v_1)\vec{e}_2+(u_1v_2-u_2v_1)\vec{e}_3 .
\end{aligned}
$$

Esempio 15.13. Usando le proprietà dei determinanti è facile verificare la proprietà anticommutativa del prodotto vettoriale, in quanto la cosa è una semplice conseguenza del fatto che, nel calcolo del determinante, uno scambio di righe produce un cambio di segno.

Esempio 15.14. Per verificare se i tre vettori $\vec{u}=(1,2,1)$, $\vec{v}=(-2,0,3)$, $\vec{w}=(-1,2,4)$ sono complanari basta calcolare il prodotto misto (in un ordine qualunque):

$$
\vec{u}\cdot\vec{v}\times\vec{w} = \begin{vmatrix} 1 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & 2 & 4 \end{vmatrix} = (-1)^{1+2}2\begin{pmatrix} -2 & 3 \\ -1 & 4 \end{pmatrix} + (-1)^{3+2}2\begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix} = (-2)(-5)+(-2)(+5) = 0 :
$$

i vettori sono complanari.

15.5. Rette nel piano, rette e piani nello spazio

Abbiamo già trattato, nel capitolo 8, la rappresentazione analitica dei luoghi geometrici nel piano in cui si sia introdotto un sistema cartesiano ortogonale monometrico Oxy. In questa sezione vediamo come l'uso dell'algebra lineare possa facilitare la trattazione di questo tipo di problemi, in particolare riproponendo con un'altra ottica lo studio delle rette nel piano. Mostreremo anche come sia possibile estendere in maniera per quasi "naturale" la trattazione allo spazio tridimensionale, dove considereremo, oltre alle rette, anche i piani.

15.5.1. Grafici non cartesiani

Prima di vedere le applicazioni dell'algebra lineare allo studio delle rette e dei piani, facciamo una breve digressione sui sistemi di coordinate in generale.

I sistemi di coordinate cartesiane ortogonali monometrici (nel piano e nello spazio) non costituiscono l'unico metodo possibile, e spesso nemmeno il più semplice, per associare coppie di reali a punti del piano, o terne di reali a punti dello spazio.

Per limitarci al piano, uno dei sistemi alternativi più importanti è quello delle coordinate *polari*, di cui abbiamo già indirettamente parlato trattando i numeri complessi nel paragrafo 14.5 del capitolo 14. Ce ne occupiamo ora un po' più estesamente, ma sempre solo nelle linee essenziali.

Fissato nel piano un punto O, che sarà detto *polo*, e una semiretta r per O, che sarà l'origine degli angoli, ad ogni punto P del piano si può far corrispondere la coppia di reali (ϱ, ϑ), data dalla distanza ϱ di P da O, e dall'angolo, misurato "in senso antiorario" e *obbligatoriamente* in radianti, tra le semirette r e OP. Diremo che abbiamo introdotto nel piano un *sistema di coordinate polari*.

Fissati O ed r si può considerare il sistema di coordinate cartesiane che ha come semiasse positivo delle ascisse la semiretta r, e come semiasse positivo delle ordinate la semiretta OP, con $P = (1, \pi/2)$. I due sistemi, polare e cartesiano, si dicono anche associati. Quando si usano, come spesso accade, contemporaneamente i due sistemi di coordinate è possibile che si ingeneri confusione tra le coppie di reali che rappresentano lo stesso punto P nei due sistemi. Tuttavia nella quasi totalità dei casi il contesto chiarisce di che cosa si sta parlando. Se proprio occorre fare una distinzione, si possono seguire le convenzioni adottate da alcuni software, convenzioni di cui abbiamo già parlato trattando dei numeri complessi, e usare il punto e virgola o i due punti al posto della virgola nel caso delle coordinate polari.

Il sistema polare ha alcuni inconvenienti, tra cui importanti:

— al punto O risulta associato il numero $\varrho = 0$, ma nessun angolo;
— l'angolo ϑ può assumere valori solo tra nell'intervallo $[0, 2\pi[$ (o altro analogo di ampiezza 2π).

Il secondo inconveniente è particolarmente fastidioso, perché se consideriamo, per esempio, un punto che si muova in senso antiorario sulla circonferenza di centro O e raggio 1, a partire dal punto con $\varrho = 1$ e $\vartheta = 0$, troveremo che quando "riattraversa" la semiretta origine, la sua seconda coordinata passa bruscamente da valori vicini a 2π a 0. Si può risolvere questo inconveniente introducendo angoli generalizzati, cioè maggiori di 2π, ma questo comporta una perdita della biunivocità della corrispondenza tra punti del piano e coppie di numeri reali (biunivocità che comunque già manca per il punto O). Nonostante questi fatti, il sistema di coordinate polari è molto usato nelle applicazioni (si pensi per esempio alla descrizione dei moti circolari).

La cosa che qui ci preme segnalare è comunque che, quando si parla di grafico delle soluzioni di un'equazione, è *assolutamente indispensabile* precisare il sistema di coordinate che si vuole usare. Consideriamo per esempio l'equazione

$$y = x \text{ con le condizioni } x \geq 0, y \geq 0,$$

e rappresentiamone graficamente le soluzioni: in un sistema cartesiano ortogonale monometrico, pensando la x come ascissa e la y come ordinata, si avrà una semiretta (la bisettrice del primo quadrante); in un sistema polare, pensando la x come distanza dall'origine e la y come angolo (generalizzato), avremo una spirale, come mostra la figura 15.15.

È per questo motivo che, quando si chiede di fare il grafico delle soluzioni di un'equazione o di un sistema, è *indispensabile* precisare il sistema di coordinate scelto.

Si deve in particolare tenere presente che affermazioni del tipo: "un'equazione di primo grado in due incognite ha come grafico una retta", sottintende che si considerino sistemi cartesiani e non polari. Detto in altri termini, l'insieme delle soluzioni di un'equazione in due incognite è un ben determinato sottoinsieme di $\mathbb{R}^2$, la sua rappresentazione grafica dipende invece dal sistema di coordinate scelto.

Tuttavia, quando non si fa esplicita menzione del contrario, si sottintende che il sistema di coordinate usato sia un sistema cartesiano (ortogonale e monometrico).

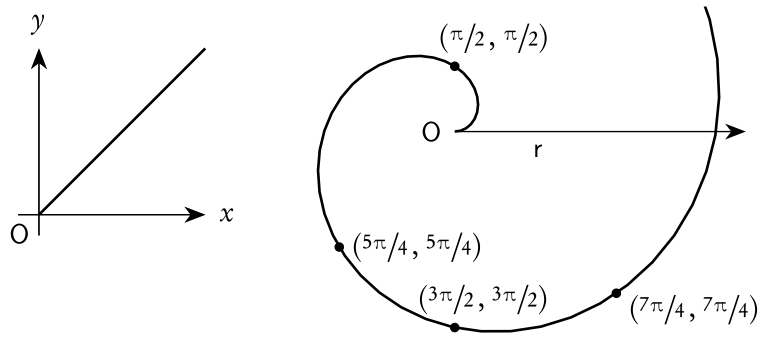


Figura 15.15.: *Grafici di $y = x$ in coordinate cartesiane e di $\varrho = \vartheta$ in coordinate polari*

15.5.2. La retta nel piano cartesiano

Consideriamo la più generale equazione di primo grado in due incognite:

$$ax + by + c = 0. \tag{15.45}$$

Perché sia effettivamente un'equazione di primo grado occorre che i coefficienti a e b non siano contemporaneamente nulli. In formule questa condizione si può scrivere, per esempio, $a^2 + b^2 > 0$ oppure $|a| + |b| > 0$.

Poiché vogliamo usare la teoria delle matrici, consideriamo la matrice incompleta e la matrice completa relativa al "sistema" costituito dalla sola equazione (15.45), dopo aver scritto l'equazione stessa nella forma, tipica dei sistemi, $ax + by = -c$, cioè con il termine noto a secondo membro:

$$A|b = \left(\begin{array}{cc|c} a & b & -c \end{array} \right). \tag{15.46}$$

Poiché, come già osservato, i coefficienti a e b non devono essere contemporaneamente nulli, la condizione sopradetta coincide allora con la condizione di risolubilità del "sistema" (costituito da una sola equazione)

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -c \end{pmatrix}, \tag{15.47}$$

ovvero con la condizione

$$\mathrm{rg}(A) = \mathrm{rg}(A|b) = 1. \tag{15.48}$$

Tenendo conto della teoria generale dei sistemi lineari, possiamo affermare che in queste condizioni l'equazione ha ∞^1 soluzioni, dipendenti da un parametro. Precisamente

— se $a \neq 0$ si ottiene $ax = -by - c$, da cui, ponendo $y = t$,

$$\begin{cases} x = -\dfrac{b}{a}t - \dfrac{c}{a} \\ y = t \end{cases};$$

— se $b \neq 0$ si ottiene $by = -ax - c$, da cui, ponendo $x = t$,

$$\begin{cases} x = t \\ y = -\dfrac{a}{b}t - \dfrac{c}{b} \end{cases};$$

— se $a \neq 0 \wedge b \neq 0$, si può scegliere indifferentemente una o l'altra forma.

Volendo considerare una formulazione generale che comprenda i casi visti, si usa scrivere la soluzione nella forma:

$$\begin{cases} x = \lambda t + \alpha \\ y = \mu t + \beta \end{cases}; \tag{15.49}$$

In un sistema cartesiano ortogonale monometrico Oxy, l'insieme delle soluzioni di un'equazione di primo grado in due incognite ha *sempre* come grafico una retta r e, viceversa, ad ogni retta r del piano corrisponde una equazione di primo grado in due incognite le cui soluzioni sono proprio tutti e soli i punti della retta.

Ebbene

1. un'equazione del tipo 15.45 si dice *equazione implicita della retta*;
2. un sistema di equazioni del tipo 15.49 si dice *(sistema di) equazioni parametriche della retta*, o semplicemente *equazione parametrica della retta*.

Se $b \neq 0$ l'equazione 15.45 si può anche scrivere nella forma

$$y = -\frac{a}{b}x - \frac{c}{b} = mx + q, \tag{15.50}$$

che viene detta *equazione esplicita della retta*. Si noti che la condizione $b \neq 0$ implica che la retta non sia "verticale", cioè parallela all'asse y.

È ovvio che l'equazione esplicita si può sempre scrivere in forma implicita ($mx - y + q = 0$) oppure parametrica ($x = t\,; y = mt + q$).

Valgono le seguenti proprietà:

— *equazione implicita*: il vettore $\vec{v} = (a, b)$ è perpendicolare alla retta;
— *equazione parametrica*: il vettore $\vec{u} = (\lambda, \mu)$ è parallelo alla retta, esso è anche detto *vettore direttore* della retta; il punto $P(\alpha, \beta)$ è un punto della retta;
— *equazione esplicita*: il numero m, detto *coefficiente angolare*, è la tangente trigonometrica dell'angolo φ formato tra la semiretta appartenente al semipiano delle y positive e la direzione positiva dell'asse delle ascisse, angolo considerato nullo se la retta è parallela all'asse delle ascisse. In considerazione di quanto sopra detto possiamo affermare che il vettore $\vec{v} = (m, -1)$ è perpendicolare alla retta, mentre il vettore $\vec{u} = (1, m)$ le è parallelo (e in effetti si vede subito che $\vec{v} \cdot \vec{u} = 0$, in accordo con la condizione di perpendicolarità tra due vettori). Il numero q è anche detto *ordinata all'origine*, in quanto rappresenta l'ordinata del punto di ascissa 0.

Il coefficiente angolare di una retta (non verticale!) ha, come è noto, un importante significato geometrico: se $A(x_A, y_A)$ e $B(x_B, y_B)$ dono due punti della retta si ha

$$m = \frac{y_B - y_A}{x_B - x_A} = \frac{\Delta y}{\Delta x}. \tag{15.51}$$

In sostanza il coefficiente angolare dà la variazione verticale (cioè di quota) in rapporto alla variazione orizzontale: è per questo motivo che si chiama anche *pendenza*.

L'ultima forma dell'equazione di una retta, utile in alcune circostanze, è la cosiddetta *equazione segmentaria*, che si può scrivere nel caso in cui sia a che b che c siano diversi da zero (retta non parallela a nessuno dei due assi e non passante per l'origine). In questo caso si ottiene:

$$ax+by=-c \Rightarrow \frac{x}{-c/a}+\frac{y}{-c/b}=1 \Rightarrow \frac{x}{p}+\frac{y}{q}=1\,. \tag{15.52}$$

I numeri p e q rappresentano l'ascissa e, rispettivamente, l'ordinata dei punti di intersezione della retta con gli assi.

Osservazione 15.30. Quando si parla di equazioni parametriche di una retta non si dovrebbe mai dire "l'equazione parametrica", ma "una equazione parametrica". Infatti la stessa retta può essere rappresentata in forma parametrica in infiniti modi diversi: l'unica cosa che conta è che il vettore (λ,μ) sia parallelo alla retta e che il punto (α,β) sia un punto della retta. Per esempio le seguenti equazioni sono tutte equazioni della stessa retta.

$$\begin{cases} x=t+2 \\ y=-2t+1 \end{cases},\quad \begin{cases} x=-t+3 \\ y=2t-1 \end{cases},\quad \begin{cases} x=(1/2)\,t+2 \\ y=-t+1 \end{cases}.$$

La domanda che può sorgere spontanea è: date due equazioni parametriche, individuate da (λ_1,μ_1), (α_1,β_1) e (λ_2,μ_2), (α_2,β_2) rispettivamente, come si può controllare se sono o no le equazioni di una stessa retta? Il modo di più semplice è quello di controllare intanto se (λ_1,μ_1) è parallelo a (λ_2,μ_2), e poi se il punto (α_1,β_1) verifica la seconda coppia di equazioni, mentre il punto (α_2,β_2) verifica la prima coppia. Lo si può fare come utile esercizio sull'esempio sopra proposto.

Una osservazione simile vale anche per l'equazione implicita (detta anche cartesiana): $x+2y-1=0$ e $2x+4y-2=0$ sono equazioni che hanno, ovviamente, le stesse soluzioni e quindi rappresentano la stessa retta. È molto importante tenere conto di questa osservazione, in particolare quando viene richiesto di trovare l'equazione implicita di una retta: anche se i parametri da determinare sono tre (a, b, c) le condizioni da considerare sono solo due.

15.5.3. Applicazioni

Utilizzando la teoria dei vettori e le osservazioni sopra riportate è facile ricavare le più importanti formule riguardanti la geometria analitica della retta nel piano. In sostanza tutte le formule che otterremo, e di cui abbiamo già parlato nel capitolo 8, possono essere considerate degli esercizi di calcolo vettoriale: per questo svolgeremo quasi sempre i calcoli in dettaglio.

In quanto segue r ed s denoteranno due rette, rispettivamente di equazioni

- implicite: $a_1x+b_1y+c_1=0$ e $a_2x+b_2y+c_2=0$;
- parametriche: $x=\lambda_1t+\alpha_1\,;\ y=\mu_1t+\beta_1$ e $x=\lambda_2t+\alpha_2\,;\ y=\mu_2t+\beta_2$;
- esplicite: $y=m_1x+q_1$ e $y=m_2x+q_2$.

Condizioni di parallelismo

Due rette r ed s sono parallele se e solo se lo sono due loro vettori normali o due loro vettori direttori. Dunque

$$r \parallel s \Leftrightarrow \vec{n}_1 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \parallel \vec{n}_2 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix},$$

oppure

$$r \parallel s \Leftrightarrow \begin{pmatrix} \lambda_1 \\ \mu_1 \end{pmatrix} \parallel \begin{pmatrix} \lambda_2 \\ \mu_2 \end{pmatrix}.$$

Si può usare la condizione di parallelismo espressa tramite prodotto vettoriale e si ottiene, nel caso dell'equazione implicita,

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & b_1 & 0 \\ a_2 & b_2 & 0 \end{vmatrix} = \vec{0} \Leftrightarrow \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 0 \Leftrightarrow a_1 b_2 - a_2 b_1 = 0.$$

Operando nello stesso modo con le equazioni parametriche si ottiene:

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \lambda_1 & \mu_1 & 0 \\ \lambda_2 & \mu_2 & 0 \end{vmatrix} = \vec{0} \Leftrightarrow \begin{vmatrix} \lambda_1 & \mu_1 \\ \lambda_2 & \mu_2 \end{vmatrix} = 0 \Leftrightarrow \lambda_1 \mu_2 - \lambda_2 \mu_1 = 0.$$

Ancora più semplice la situazione nel caso dell'equazione esplicita:

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ m_1 & -1 & 0 \\ m_2 & -1 & 0 \end{vmatrix} = \vec{0} \Leftrightarrow \begin{vmatrix} m_1 & -1 \\ m_2 & -1 \end{vmatrix} = 0 \Leftrightarrow -m_1 + m_2 \Leftrightarrow m_1 = m_2.$$

Condizione di perpendicolarità

In questo caso si tratta di utilizzare la condizione di perpendicolarità esprimibile mediante l'annullamento del prodotto scalare. Si ottiene subito:

- equazione implicita: $a_1 a_2 + b_1 b_2 = 0$;
- equazione parametrica: $\lambda_1 \lambda_2 + \mu_1 \mu_2 = 0$;
- equazione esplicita: $m_1 m_2 + 1 = 0, \Leftrightarrow, m_1 m_2 = -1$ (in questo caso occorre supporre che le due rette non siano né verticali né orizzontali).

Retta per due punti

Dati due punti distinti $\mathsf{A}(x_\mathsf{A}, y_\mathsf{A})$ e $\mathsf{B}(x_\mathsf{B}, y_\mathsf{B})$, vogliamo trovare l'equazione della retta da essi univocamente individuata.

Cominciamo con l'osservare che il vettore $\vec{u} = (x_\mathsf{B} - x_\mathsf{A}, y_\mathsf{B} - y_\mathsf{A})$ è parallelo alla retta. Questo ci consente di scrivere immediatamente le equazioni parametriche della retta cercata:

$$\begin{cases} x = (x_\mathsf{B} - x_\mathsf{A})t + x_\mathsf{A} \\ y = (y_\mathsf{B} - y_\mathsf{A})t + y_\mathsf{A} \end{cases},$$

che sono della forma 15.49.

Per scrivere l'equazione implicita basta osservare che un punto $\mathsf{P}(x,y)$ del piano $\mathsf{O}xy$ appartiene alla retta se e soltanto se il vettore $\overrightarrow{\mathsf{AP}}$ risulta parallelo al vettore $\vec{u}$. Scriviamo la condizione di parallelismo (direttamente col prodotto vettoriale, senza tentare di memorizzare la formula che pure abbiamo appena ricavato!):

$$\begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ x_\mathsf{B}-x_\mathsf{A} & y_\mathsf{B}-y_\mathsf{A} & 0 \\ x-x_\mathsf{A} & y-y_\mathsf{A} & 0 \end{vmatrix} = 0 \Leftrightarrow \begin{vmatrix} x_\mathsf{B}-x_\mathsf{A} & y_\mathsf{B}-y_\mathsf{A} \\ x-x_\mathsf{A} & y-y_\mathsf{A} \end{vmatrix} = 0,$$

condizione che si può scrivere anche

$$(x-x_\mathsf{A})(y_\mathsf{B}-y_\mathsf{A}) = (y-y_\mathsf{A})(x_\mathsf{B}-x_\mathsf{A}).$$

Spesso nei testi questa formula si trova scritta nella forma

$$\frac{x-x_\mathsf{A}}{x_\mathsf{B}-x_\mathsf{A}} = \frac{y-y_\mathsf{A}}{y_\mathsf{B}-y_\mathsf{A}}.$$

Ne sconsigliamo l'uso, in quanto in quest'ultima forma occorre che i denominatori siano diversi da zero, ovvero che la retta non sia parallela a nessuno dei due assi, la qual cosa non è sempre verificata nelle applicazioni che interessano.

Retta per un punto e parallela (o perpendicolare) a una retta data

Se la retta data è non verticale, si può, come già noto, scriverla in forma esplicita: $y = mx + q$, cioè $mx - y + q = 0$, con vettore perpendicolare $\vec{v} = (m, -1)$, o parametrica $x = t\,; y = mt + q$, con vettore parallelo $\vec{u} = (1, m)$. Procediamo intanto a trovare la parallela a una retta data. Detto $\mathsf{A}(x_\mathsf{A}, y_\mathsf{A})$ il punto assegnato, un punto $\mathsf{P}(x,y)$ appartiene alla retta se e solo se $\overrightarrow{\mathsf{AP}}$ è parallelo a $\vec{u}$.

— Usando la condizione di parallelismo si ottiene:

$$\begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ 1 & m & 0 \\ x-x_\mathsf{A} & y-y_\mathsf{A} & 0 \end{vmatrix} = 0 \Leftrightarrow \begin{vmatrix} 1 & m \\ x-x_\mathsf{A} & y-y_\mathsf{A} \end{vmatrix} = 0 \Leftrightarrow y-y_\mathsf{A} = m(x-x_\mathsf{A}).$$

— Usando la condizione di perpendicolarità si ottiene:

$$\begin{pmatrix} m \\ -1 \end{pmatrix} \cdot \begin{pmatrix} x-x_\mathsf{A} \\ y-y_\mathsf{A} \end{pmatrix} = 0 \Leftrightarrow y-y_\mathsf{A} = m(x-x_\mathsf{A}),$$

cioè lo stesso risultato di prima.

Se non è possibile, o conveniente, scrivere la retta data in forma esplicita e si vuole usare la forma implicita, $ax + by + c = 0$, un vettore perpendicolare alla retta è immediato ($\vec{v} = (a, b)$), un vettore parallelo si può trovare o passando alle equazioni parametriche o, semplicemente, prendendo due punti P e Q arbitrari della retta e considerando il vettore $u = \mathsf{P} - \mathsf{Q}$. Dopodiché si procede esattamente come prima.

Distanza di un punto da una retta

Siano $\mathsf{A}(x_\mathsf{A}, y_\mathsf{A})$ un punto e r una retta di equazione implicita $ax+by+c=0$ (in questo caso l'equazione implicita è la più conveniente). Detto $\mathsf{P}(x_\mathsf{P}, y_\mathsf{P})$ un generico punto della retta, la distanza richiesta, $d(\mathsf{A}, \mathsf{r})$, non è altro che il modulo della proiezione di $\overrightarrow{\mathsf{AP}}$ sulla perpendicolare alla retta stessa.

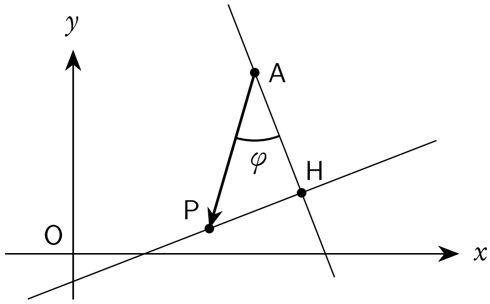


Figura 15.16.: *Distanza di un punto da una retta*

Con riferimento alla figura 15.16, si ha: $|\overline{\mathsf{AH}}| = |\overrightarrow{\mathsf{AP}}| \cdot |\cos(\varphi)|$. Teniamo ora conto che $\vec{v} = (a, b)$ è *un* vettore perpendicolare alla retta, cioè parallelo ad $\overline{\mathsf{AH}}$. Il vettore $\vec{v}$ potrebbe avere il verso di $\overrightarrow{\mathsf{AH}}$ oppure quello di $\overrightarrow{\mathsf{HA}}$. Allora l'angolo tra $\vec{v}$ e $\overrightarrow{\mathsf{AP}}$ è φ, oppure $\pi - \varphi$, ma questi due angoli hanno coseno che differisce solo per il segno. Se ne deduce che

$$\left|\vec{v} \cdot \overrightarrow{\mathsf{AP}}\right| = |a(x_\mathsf{P} - x_\mathsf{A}) + b(y_\mathsf{P} - y_\mathsf{A})| = |\vec{v}| \left|\overrightarrow{\mathsf{AP}}\right| |\cos(\varphi)|,$$

da cui

$$|\cos(\varphi)| = \frac{|a(x_\mathsf{P} - x_\mathsf{A}) + b(y_\mathsf{P} - y_\mathsf{A})|}{|\vec{v}| \left|\overrightarrow{\mathsf{AP}}\right|}.$$

Teniamo ora conto che il punto A appartiene alla retta, per cui $ax_\mathsf{A} + by_\mathsf{A} = -c$, e che

$$|\vec{v}| = \sqrt{a^2 + b^2}.$$

Allora

$$|\cos(\varphi)| = \frac{|ax_\mathsf{P} + by_\mathsf{P} - (ax_\mathsf{A} + by_\mathsf{A})|}{|\vec{v}| \left|\overrightarrow{\mathsf{AP}}\right|} = \frac{|ax_\mathsf{P} + by_\mathsf{P} + c|}{|\vec{v}| \left|\overrightarrow{\mathsf{AP}}\right|}.$$

Si conclude che

$$d(\mathsf{A}, r) = |\overline{\mathsf{AH}}| = |\overrightarrow{\mathsf{AP}}| \cdot |\cos(\varphi)| = |\overrightarrow{\mathsf{AP}}| \frac{|ax_\mathsf{P} + by_\mathsf{P} + c|}{\sqrt{a^2 + b^2} \left|\overrightarrow{\mathsf{AP}}\right|} = \frac{|ax_\mathsf{P} + by_\mathsf{P} + c|}{\sqrt{a^2 + b^2}}.$$

15.5.4. Intersezioni di rette nel piano

Date due o più rette nel piano ci possiamo chiedere se esse hanno o no punti in comune. Dal punto di vista algebrico il problema si traduce nella risoluzione di un sistema di equazioni in due incognite, a cui si potranno applicare tutte le tecniche già viste.

È particolarmente importante, per il suo significato geometrico, il caso di due rette: $a_1x+b_1y+c_1=0$ e $a_2x+b_2y+c_2=0$. Il sistema formato dalle due equazioni è

$$\begin{cases} a_1x+b_1y & = & -c_1 \\ a_2x+b_2y & = & -c_2 \end{cases},$$

con le matrici completa e incompleta seguenti

$$A|b=\left(\begin{array}{cc|c} a_1 & b_1 & -c_1 \\ a_2 & b_2 & -c_2 \end{array}\right).$$

Si possono presentare tre possibilità.

— $\operatorname{rg}(A)=2$ (e quindi a fortiori $\operatorname{rg}(A|b)=2$), ovvero

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0.$$

In questo caso il sistema ha una sola soluzione, il punto di intersezione delle due rette. Si può notare che la condizione appena scritta non è altro che la condizione che le due rette non siano parallele, dunque appartengono a un *fascio proprio* di rette.

— $\operatorname{rg}(A)=1$ e $\operatorname{rg}(A|b)=2$, ovvero

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=0,\ \text{ma}\ \begin{vmatrix} a_1 & -c_1 \\ a_2 & -c_2 \end{vmatrix} \neq 0 \vee \begin{vmatrix} b_1 & -c_1 \\ b_2 & -c_2 \end{vmatrix} \neq 0.$$

In questo caso il sistema non ha soluzioni. Le condizioni appena scritte esprimono il fatto che le due rette sono parallele, ma distinte, dunque appartengon a un *fascio improprio* di rette.

— $\operatorname{rg}(A)=1$ e $\operatorname{rg}(A|b)=1$, ovvero

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=\begin{vmatrix} a_1 & -c_1 \\ a_2 & -c_2 \end{vmatrix}=\begin{vmatrix} b_1 & -c_1 \\ b_2 & -c_2 \end{vmatrix}=0.$$

In questo caso il sistema ha infinite (∞^1) soluzioni. Le condizioni appena scritte esprimono il fatto che le due rette sono parallele e coincidenti.

15.5.5. Piani nello spazio cartesiano

La trattazione dell'equazione di un piano nello spazio cartesiano avviene in maniera sostanzialmente identica a quella delle rette nello spazio; naturalmente, avendo ora una variabile in più, la situazione potrà essere più complessa (e quindi più interessante). Lo si può constatare confrontando le righe che seguono con le definizioni date nel caso della retta nel piano.

Consideriamo la più generale equazione di primo grado in tre incognite:

$$ax+by+cz+d=0. \tag{15.53}$$

Perché sia effettivamente un'equazione di primo grado occorre che i coefficienti a, b e c non siano contemporaneamente nulli. In formule questa condizione si può scrivere, per esempio, $a^2+b^2+c^2>0$ oppure $|a|+|b|+|c|>0$.

Poiché vogliamo usare la teoria delle matrici, consideriamo la matrice incompleta e la matrice completa, scrivendo l'equazione nella forma, tipica dei sistemi, $ax+by+cz=-d$, cioè con il termine noto a secondo membro:

$$A|b=\left(\begin{array}{ccc|c} a & b & c & -d \end{array}\right). \tag{15.54}$$

Poiché, come già osservato, i coefficienti a, b e c non devono essere contemporaneamente nulli, la condizione sopradetta coincide allora con la condizione di risolubilità del "sistema" (costituito da una sola equazione)

$$\begin{pmatrix} a & b & c \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} -d \end{pmatrix}, \tag{15.55}$$

che è la condizione

$$\mathrm{rg}(A)=\mathrm{rg}(A|b)=1. \tag{15.56}$$

Tenendo conto della teoria generale dei sistemi lineari, possiamo affermare che in queste condizioni l'equazione ha ∞^2 soluzioni, dipendenti da due parametri. Precisamente

— se $a\neq 0$, allora si scrive $ax=-by-cz-d$, da cui, ponendo $y=u$ e $z=v$,

$$\begin{cases} x &= -\dfrac{b}{a}u-\dfrac{c}{a}v-\dfrac{d}{a} \\ y &= u \\ z &= v \end{cases};$$

— se $b\neq 0$, allora si scrive $by=-ax-cz-d$, da cui, ponendo $x=u$ e $z=v$,

$$\begin{cases} x &= u \\ y &= -\dfrac{a}{b}u-\dfrac{c}{b}v-\dfrac{d}{b} \\ z &= v \end{cases};$$

— se $c\neq 0$, allora si scrive $cz=-ax-by-d$, da cui, ponendo $x=u$ e $y=v$,

$$\begin{cases} x &= u \\ y &= v \\ y &= -\dfrac{a}{c}u-\dfrac{b}{c}v-\dfrac{d}{c} \end{cases};$$

— se $a\neq 0\land b\neq 0\land c\neq 0$, si può scegliere indifferentemente una o l'altra forma.

Volendo considerare una formulazione generale che comprenda i casi visti, si usa scrivere la soluzione nella forma:

$$\begin{cases} x &= \lambda_1 u+\lambda_2 v+\alpha \\ y &= \mu_1 u+\mu_2 v+\beta \\ z &= \nu_1 u+\nu_2 v+\gamma \end{cases}; \tag{15.57}$$

In un sistema cartesiano ortogonale monometrico $Oxyz$, l'insieme delle soluzioni di un'equazione di primo grado in tre incognite ha *sempre* come grafico un piano π e, viceversa, ad ogni piano π dello spazio corrisponde una equazione di primo grado in due incognite le cui soluzioni sono proprio tutti e soli i punti del piano.

Ebbene

1. un'equazione del tipo 15.53 si dice *equazione implicita del piano*;
2. un sistema di equazioni del tipo 15.57 si dice *(sistema di) equazioni parametriche del piano*, o, semplicemente *equazione parametrica del piano*.

Se $c \neq 0$ l'equazione 15.53 si può anche scrivere nella forma

$$z = -\frac{a}{c}x - \frac{b}{c}y - \frac{d}{c} = mx + ny + q\,, \tag{15.58}$$

che viene detta *equazione esplicita del piano*. Si noti che la condizione $c \neq 0$ implica che il piano non sia "verticale", cioè parallelo all'asse z, o, il che è lo stesso, perpendicolare al piano xy.

È ovvio che l'equazione esplicita si può sempre scrivere in forma implicita ($mx + ny - z + q = 0$) oppure parametrica ($x = u\,; y = v\,; z = mu + nv + q$).

Valgono le seguenti proprietà:

— *equazione implicita*: il vettore $\vec{v} = (a, b, c)$ è perpendicolare al piano;
— *equazione parametrica*: i vettori $\vec{u}_1 = (\lambda_1, \mu_1, \nu_1)$ e $\vec{u}_2 = (\lambda_2, \mu_2, \nu_2)$ sono paralleli al piano, e non paralleli tra di loro, e si dicono anche *vettori di giacitura* del piano; il punto $P(\alpha, \beta, \gamma)$ è un punto del piano;
— *equazione esplicita*: in considerazione di quanto appena detto possiamo affermare che il vettore $\vec{v} = (m, n, -1)$ è perpendicolare al piano, mentre i vettori $\vec{u}_1 = (1, 0, m)$ e $\vec{u}_2 = (0, 1, n)$ sono paralleli al piano (e in effetti si può facilmente controllare che i vettori $\vec{v} = (m, n, -1)$ e $\vec{u} = \vec{u}_1 \times \vec{u}_2$ sono tra di loro paralleli).

L'ultima forma dell'equazione di un piano, utile in alcune circostanze, è la cosiddetta *equazione segmentaria*, che si può scrivere nel caso in cui a, b, c e d sono tutti diversi da zero (piano non parallelo a nessuno dei tre assi e non passante per l'origine). In questo caso si ottiene:

$$ax + by + cz = -d \;\Rightarrow\; \frac{x}{-d/a} + \frac{y}{-d/b} + \frac{z}{-d/c} = 1 \;\Rightarrow\; \frac{x}{p} + \frac{y}{q} + \frac{z}{r} = 1\,. \tag{15.59}$$

I numeri p, q ed r rappresentano la coordinata non nulla dei punti di intersezione del piano con gli assi, come mostrato nella figura 15.17.

15.5.6. Applicazioni

Impiegando quasi le stesse parole già usate nel caso della retta, utilizzando la teoria dei vettori si possono ottenere tutte le più importanti formule della geometria analitica del piano nello spazio. Anche in questo caso le formule che via via otterremo possono essere considerate esercizi di calcolo vettoriale. Come nel caso della retta, anche qui useremo notazioni il cui significato sarà ovvio dal contesto.

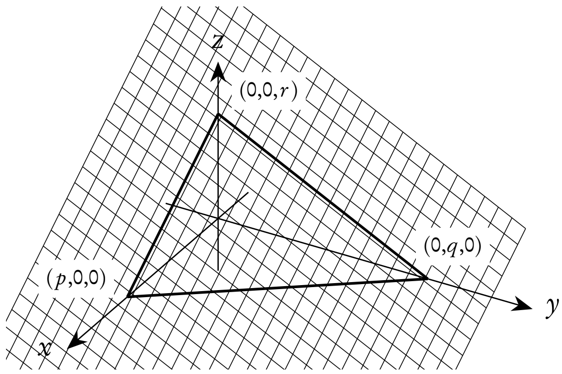


Figura 15.17.: *Piano e significato della equazione segmentaria*

Condizioni di parallelismo e perpendicolarità

Due piani α e β sono paralleli se e solo se due loro vettori normali sono paralleli. Se i piani sono dati in forma implicita, due vettori normali sono semplicemente dati da (a_1, b_1, c_1) e (a_2, b_2, c_2). Se i piani sono dati in forma parametrica, per trovare due vettori normali basta fare il prodotto vettoriale di due vettori di giacitura. Nel primo caso, per esempio, la condizione di parallelismo si potrà scrivere

$$\begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = \vec{0} \Leftrightarrow \operatorname{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 1\,.$$

Due piani α e β sono perpendicolari se e solo se due loro vettori normali sono perpendicolari. Analogamente alla condizione di parallelismo, la perpendicolarità tra due piani in equazione implicita si potrà scrivere, usando il prodotto scalare,

$$a_1a_2 + b_1b_2 + c_1c_2 = 0$$

Piano per tre punti non allineati

Dati tre punti non allineati $\mathsf{A}(x_\mathsf{A}, y_\mathsf{A}, z_\mathsf{A})$, $\mathsf{B}(x_\mathsf{B}, y_\mathsf{B}, z_\mathsf{B})$ e $\mathsf{C}(x_\mathsf{C}, y_\mathsf{C}, z_\mathsf{C})$, vogliamo trovare l'equazione de piano da essi univocamente individuato.

Cominciamo con l'osservare che i vettori $\overrightarrow{\mathsf{AB}}$ e $\overrightarrow{\mathsf{AC}}$ sono due vettori non paralleli, che possono essere usati come vettori di giacitura del piano. Le equazioni parametriche sono ora immediate:

$$\begin{cases} x = (x_\mathsf{B} - x_\mathsf{A})u + (x_\mathsf{C} - x_\mathsf{A})v + x_\mathsf{A} \\ y = (y_\mathsf{B} - y_\mathsf{A})u + (y_\mathsf{C} - y_\mathsf{A})v + y_\mathsf{A} \\ z = (z_\mathsf{B} - z_\mathsf{A})u + (z_\mathsf{C} - z_\mathsf{A})v + z_\mathsf{A} \end{cases}.$$

Per scrivere un'equazione implicita si può procedere in vari modi.

Se si calcola il prodotto vettoriale $\vec{v} = \overrightarrow{\mathsf{AB}} \times \overrightarrow{\mathsf{AC}}$, si ottiene un vettore perpendicolare al piano: le sue componenti, che possiamo indicare con a, b, c, si possono assumere come primi tre parametri nell'equazione $ax+by+cz+d=0$. Il parametro restante si può trovare imponendo la condizione di passaggio per uno dei tre punti dati.

Si può in alternativa osservare che un punto $\mathsf{P}(x,y,z)$ appartiene al piano se, e solo se, i tre vettori $\overrightarrow{\mathsf{AP}}$, $\overrightarrow{\mathsf{AB}}$ e $\overrightarrow{\mathsf{AC}}$ sono paralleli. La condizione di complanarità scritta con il prodotto misto, o direttamente con il determinante diventa:

$$\begin{vmatrix} x-x_\mathsf{A} & y-y_\mathsf{A} & z-z_\mathsf{A} \\ x_\mathsf{B}-x_\mathsf{A} & y_\mathsf{B}-y_\mathsf{A} & z_\mathsf{B}-z_\mathsf{A} \\ x_\mathsf{C}-x_\mathsf{A} & y_\mathsf{C}-y_\mathsf{A} & z_\mathsf{C}-z_\mathsf{A} \end{vmatrix}=0$$

Piano per un punto e parallelo a un piano dato

Si può semplicemente osservare che se $ax+by+cz+d=0$ è il piano dato, quello cercato deve essere del tipo $ax+by+cz+\delta=0$ e la determinazione di δ richiede solo di scrivere la condizione di passaggio per il punto assegnato.

Distanza di un punto da un piano

È un utile esercizio ripetere, quasi con le stesse parole, quanto già detto per la distanza di un punto da una retta.

Detto $\mathsf{A}(x_\mathsf{A},y_\mathsf{A},z_\mathsf{A})$ un punto e π un piano di equazione implicita $ax+by+cz+d=0$ (anche in questo caso l'equazione implicita è la più conveniente) si ottiene facilmente:

$$d(\mathsf{A},\pi)=\frac{|ax_\mathsf{A}+by_\mathsf{A}+cz_\mathsf{A}+d|}{\sqrt{a^2+b^2+c^2}}.$$

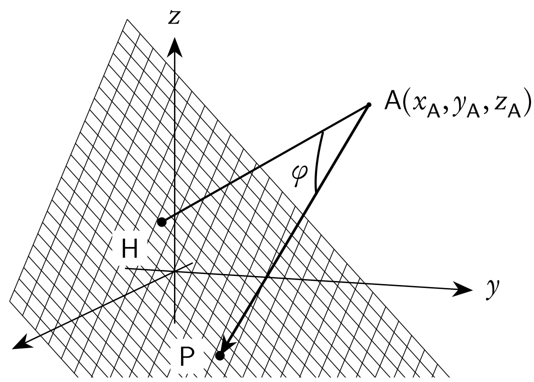


Figura 15.18.: *Distanza di un punto da un piano*

15.5.7. Intersezione di piani nello spazio

Dati due o più piani nello spazio ci possiamo chiedere se essi hanno o no punti in comune. Dal punto di vista algebrico il problema si traduce nella risoluzione di un sistema di equazioni in tre incognite, a cui si potranno applicare tutte le tecniche già viste.

Sono particolarmente importanti, per il loro significato geometrico, il caso di due e quello di tre piani: $a_1x+b_1y+c_1z+d_1=0$, $a_2x+b_2y+c_2z+d_2=0$ e $a_3x+b_3y+c_3z+d_3=0$.

Nel caso di tre piani, il sistema formato dalle loro equazioni è

$$\begin{cases} a_1x+b_1y+c_1z & = & -d_1 \\ a_2x+b_2y+c_2z & = & -d_2 \\ a_3x+b_3y+c_3z & = & -d_3 \end{cases},$$

con le matrici completa e incompleta seguenti

$$A|b=\left(\begin{array}{ccc|c} a_1 & b_1 & c_1 & -d_1 \\ a_2 & b_2 & c_2 & -d_2 \\ a_3 & b_3 & c_3 & -d_3 \end{array}\right).$$

Se la matrice A ha rango massimo, 3, il sistema è sicuramente compatibile e ha una e una sola soluzione: i tre piani passano per uno stesso punto, ovvero appartengono a quella che sia chiama una *stella di piani*.

Se la matrice A ha rango minore di 3 allora il sistema può essere compatibile o incompatibile, a seconda che $A|b$ abbia o no lo stesso rango di A. Se il sistema è compatibile e A ha rango 2, allora una delle tre equazioni è superflua e il sistema formato dalle altre due ha infinite (∞^1) soluzioni, dipendenti da un parametro t: si tratta di tre piani che hanno in comune una retta, cioè appartengono ad un *fascio di piani*. Se il sistema è compatibile e A ha rango 1, allora due delle tre equazioni sono superflue e il sistema ha infinite (∞^2) soluzioni, dipendenti da due parametri u e v: si tratta di tre piani coincidenti. Se il sistema è incompatibile si possono presentare varie situazioni. Per esempio se i piani sono distinti possono essere tutti tra di loro paralleli, due tra di loro paralleli e uno no, a due a due incidenti in rette parallele. È un utile esercizio discutere le varie possibilità alla luce dei ranghi rispettivi di A e $A|b$.

Consideriamo ora un sistema formato da due sole equazioni.

$$\begin{cases} a_1x+b_1y+c_1z & = & -d_1 \\ a_2x+b_2y+c_2z & = & -d_2 \end{cases},$$

con le matrici completa e incompleta seguenti

$$A|b=\left(\begin{array}{ccc|c} a_1 & b_1 & c_1 & -d_1 \\ a_2 & b_2 & c_2 & -d_2 \end{array}\right).$$

Se il rango di A, e quindi a fortiori di $A|b$, è 2, il sistema ha infinite (∞^1) soluzioni dipendenti da un parametro t: si tratta, come già più sopra osservato, dei punti appartenenti alla retta intersezione dei due piani, che, in ragione del valore del rango della matrice incompleta, non sono paralleli. Anzi, nello spazio, è proprio come intersezione di due piani non paralleli che si scrive una retta in *equazione cartesiana*. Se il rango di A è 1 e il sistema è compatibile, il sistema ha ∞^2 soluzioni, cioè i due piani coincidono; se il sistema non è compatibile, i due piani sono paralleli. Il fatto che il rango di A sia 1 è difatti espresso proprio dalla condizione di parallelismo dei due vettori (a_1,b_1,c_1) e (a_2,b_2,c_2).

15.5.8. Rette nello spazio

Una retta nello spazio può essere rappresentata sostanzialmente in due modi: mediante *equazioni cartesiane*, ovvero come intersezione di due piani non paralleli (come già osservato), oppure mediante *equazioni parametriche*.

Riconsideriamo allora il sistema di 2 equazioni in tre incognite già visto sopra:

(15.60)
$$\begin{cases} a_1x + b_1y + c_1z + d_1 & = & 0 \\ a_2x + b_2y + c_2z + d_2 & = & 0 \end{cases},$$

che, al solito, scriviamo con i termini noti a secondo membro

(15.61)
$$\begin{cases} a_1x + b_1y + c_1z & = & -d_1 \\ a_2x + b_2y + c_2z & = & -d_2 \end{cases},$$

con le matrici completa e incompleta seguenti

(15.62)
$$A|b = \left(\begin{array}{ccc|c} a_1 & b_1 & c_1 & -d_1 \\ a_2 & b_2 & c_2 & -d_2 \end{array}\right),$$

e supponiamo che il rango di A sia 2, condizione che traduce in formule, come già osservato, il fatto che i due piani non sono paralleli

Almeno uno dei tre minori della matrice A ha dunque determinante diverso da zero, e si può applicare il noto metodo di risoluzione dei sistemi. Se, per esempio,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0,$$

il sistema si può scrivere nella forma:

$$\begin{cases} a_1x + b_1y & = & -c_1z - d_1 \\ a_2x + b_2y & = & -c_2z - d_2 \end{cases}.$$

Ponendo $z = t$ il sistema si risolve con il metodo di Cramer e la soluzione si può scrivere nella forma:

$$\begin{cases} x = \lambda t + \alpha \\ y = \mu t + \beta \\ z = t \end{cases},$$

con opportuni valori dei parametri.

In generale, considerando le varie possibilità, la soluzione potrà essere scritta nella forma:

(15.63)
$$\begin{cases} x = \lambda t + \alpha \\ y = \mu t + \beta \\ z = \nu t + \gamma \end{cases},$$

detta *(sistema di) equazioni parametriche* di una retta nello spazio. Il sistema 15.60, per contro, si chiama *equazione cartesiana* di una retta nello spazio. Come si noterà, l'equazione 15.63 è praticamente la stessa della 15.49, che rappresentava una retta nel piano: l'unica differenza è l'aggiunta della variabile z. Valgono anche le stesse proprietà. In particolare il vettore $\vec{u} = (\lambda, \mu, \nu)$ è parallelo alla retta, esso è anche detto *vettore direttore* della retta; il punto $\mathrm{P}(\alpha, \beta, \gamma)$ è un punto della retta.

Il passaggio da una forma all'altra dell'equazione della retta si può fare in vari modi. Per esempio dall'equazione cartesiana si passa a quella parametrica seguendo il procedimento (basato sulla regola di Cramer) delineato sopra. Si può, in alternativa, osservare che il vettore ottenuto come prodotto vettoriale tra $\vec{v}_1 = (a_1, b_1, c_1)$ e $\vec{v}_2 = (a_2, b_2, c_2)$ è un vettore direttore della retta e quindi basta successivamente trovare un punto della retta. Il passaggio dall'equazione parametrica a quella cartesiana è poi immediato: basta, per esempio, ricavare la t in una delle tre equazioni e sostituirla nelle altre due.

La cosa che si deve ricordare, sia per il piano che per lo spazio, è che nell'equazione cartesiana (o implicita) si dispone immediatamente di uno (nel piano) o due (nello spazio) vettori perpendicolari alla retta, nell'equazione parametrica si dispone invece immediatamente di un vettore parallelo alla retta. Usando opportunamente le proprietà dei vettori si potranno risolvere tutti i problemi che riguardano rette e piani anche nello spazio.

15.5.9. Esempi e applicazioni

Proponiamo ora alcuni problemi tipici che concernono rette e piani nello spazio. In tutti i casi daremo solo indicazioni di massima su uno dei metodi applicabili: è un utile esercizio tradurre le indicazioni in procedimenti operativi, magari esplorando le (spesso numerose) alternative possibili.

Piano parallelo a due rette non parallele e passante per un punto

Le due rette individuano due vettori $\vec{u}$ e $\vec{v}$ da cui si deduce la giacitura del piano. La conoscenza del punto rende poi unico il piano cercato: le equazioni parametriche si possono scrivere facilmente. Osserviamo che se le due rette fossero parallele, il problema avrebbe infinite soluzioni, e precisamente tutto il fascio di piani individuato dalla retta r passante per il punto dato e parallela alle due rette date.

Retta per due punti

Detti A e B i due punti, basta osservare che il vettore $\overrightarrow{AB}$ è un vettore direttore della retta: questo fatto, unitamente al passaggio per uno dei punti consente di scrivere subito le equazioni parametriche.

Piano per un punto e una retta che non si appartengono

Tra i procedimenti possibili c'è quello di trovare un punto A sulla retta. Detto poi P il punto dato e $\vec{u}$ un vettore direttore della retta, il piano è individuato dai vettori di giacitura $\overrightarrow{AP}$ e $\vec{u}$ e dal punto P.

Piano individuato da due rette incidenti

I vettori direttori delle due rette danno due vettori di giacitura del piano, il loro punto di intersezione (o un qualunque altro punto su una delle due rette) individua un punto del piano.

Retta per un punto e parallela a due piani non paralleli

Basta trovare la retta intersezione dei due piani: un suo vettore direttore è anche vettore direttore della retta cercata, che sarà individuata dall'ulteriore condizione di passare per il punto dato.

Retta per un punto e perpendicolare a una retta esterna al punto

Detti $\mathsf{A}(x_\mathsf{A}, y_\mathsf{A}, z_\mathsf{A})$ il punto e $x = \lambda t + \alpha$; $y = \mu t + \beta$; $z = \nu t + \gamma$ l'equazione parametrica della retta r, il punto $\mathsf{Q}(x, y, z)$, proiezione di A su r, è individuato dalla condizione che $\overrightarrow{\mathsf{QA}}$ sia perpendicolare a (λ, μ, ν). La condizione, usando il prodotto scalare, si scrive:

$$(\lambda t + \alpha - x_\mathsf{A})\lambda + (\mu t + \beta - y_\mathsf{A})\mu + (\nu t + \gamma - z_\mathsf{A})\nu = 0\,.$$

Da qui si trova il punto Q: l'equazione della retta cercata è ora quella della retta per 2 punti.

16. Cenno alle affinità: la via analitica

In questo capitolo sono presentate le principali caratteristiche delle affinità nel piano, facendo riferimento esclusivamente alla via analitica, piuttosto che a quella sintetica. Spesso si fa uso del calcolo matriciale, che facilita di molto la trattazione. In ogni caso il contenuto è fruibile anche da parte di coloro che non conoscono i pochi rudimenti di algebra lineare richiesti.

16.1. Trasformazioni del piano in sè

Definizione 16.1 (Trasformazioni del piano in sè)**.** *Sia dato, in un piano π, un riferimento cartesiano ortonormale. Una qualunque funzione f del piano in sè, cioè una funzione che associa ad ogni punto del piano un altro punto del piano, si chiama una* trasformazione del piano in sè.

Poiché abbiamo fissato un riferimento cartesiano, per assegnare la funzione basterà assegnare una regola per calcolare le coordinate di $\mathsf{P}' = f(\mathsf{P})$, a partire da quelle di P, ovvero considerare una funzione $f: \mathbb{R}^2 \to \mathbb{R}^2$.

Esempio 16.1. Un esempio di funzione di questo genere è il seguente, dove abbiamo indicato con (x', y') le coordinate di P' (punto immagine) e con (x, y) quelle di P, notazioni che useremo regolarmente in seguito.

$$\begin{cases} x' &= x^2 - y + 2 \\ y' &= \mathrm{e}^x + y^2 \end{cases} . \tag{16.1}$$

Potremo anche scrivere

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = f \begin{pmatrix} x \\ y \end{pmatrix}, \tag{16.2}$$

ove abbiamo preferito la scrittura delle coppie di numeri reali in colonna anziché in riga, in accordo con le notazioni comuni nell'algebra lineare. Si noti che, nella sostanza, per assegnare la funzione f abbiamo dovuto assegnare 2 funzioni: la prima alla coppia (x, y) fa corrispondere la prima coordinata, x', del punto trasformato, la seconda fa corrispondere alla coppia (x, y) la seconda coordinata, y', del punto trasformato. Come d'abitudine, chiameremo queste due funzioni le due *componenti* della funzione f.

La figura 16.1 illustra, su alcuni punti, come opera questa funzione. Sono rappresentati 3 punti, P_1, P_2, P_3 e le loro immagini P'_1, P'_2, P'_3, collegati da una linea tratteggiata al solo scopo di rendere evidente l'azione della funzione.

Nell'esempio considerato abbiamo rappresentato solo tre punti e le loro immagini: ovviamente non sarebbe stato possibile rappresentare tutti i punti e le loro immagini. Nelle applicazioni ha molto interesse valutare le relazioni che intercorrono tra certi insiemi di punti del piano e i loro trasformati mediante funzioni come la (16.1), in particolare quando questi insiemi di punti sono figure geometriche

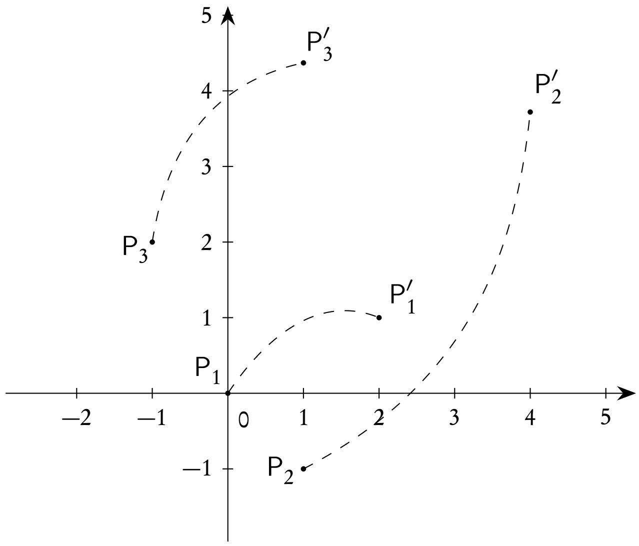


Figura 16.1.: *Alcuni punti e le loro immagini, tramite la trasformazione dell'esempio* 16.1

come rette, circonferenze, ecc. Per trasformazioni del tutto generali come la (16.1) le figure trasformate non mantengono nessuna delle caratteristiche delle figure originarie.

Ci occuperemo in questo capitolo delle trasformazioni in cui entrambe le funzioni che compongono la f sono di primo grado in x e y, con qualche ulteriore condizione: le chiameremo *affinità*.

Prima di procedere vogliamo però porre l'attenzione su un fatto molto importante. Trattando delle trasformazioni del piano in sè, cioè di funzioni di $\mathbb{R}^2$ in $\mathbb{R}^2$, *non* siamo interessati a trattare il grafico della funzione f: questo grafico[1] non può avere alcuna rappresentazione nel senso tradizionale del termine, in quanto servirebbe un sistema di 4 assi cartesiani (magari ortogonali), cosa chiaramente impossibile nello spazio tridimensionale. Quello a cui siamo interessati è piuttosto l'aspetto geometrico di particolari sottoinsiemi del dominio e dei loro insiemi immagini.

16.2. Affinità nel piano

La più generica trasformazione del piano in sè, rappresentabile mediante funzioni di primo grado nelle variabili x e y si può scrivere nella forma

$$\left\{\begin{array}{lcl} x' & = & ax+by+p \\ y' & = & cx+dy+q \end{array}\right. . \tag{16.3}$$

Queste funzioni sono spesso chiamate *lineari*, anche se il termine è improprio, in quanto una funzione lineare di $\mathbb{R}^2$ in $\mathbb{R}^2$ è del tipo

$$\left\{\begin{array}{lcl} x' & = & ax+by \\ y' & = & cx+dy \end{array}\right. , \tag{16.4}$$

[1] Il grafico può essere definito analiticamente come l'insieme delle coppie $\big((x,y),(x',y')\big)$, ovvero delle quaterne (x,y,x',y').

cioè, rispetto alle (16.3), è priva dei *termini noti*. Meglio sarebbe chiamarla, anche in vista di quanto diremo, *funzione affine* e noi ci atterremo a questa nomenclatura.

Posto

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix},\quad \vec{u}=\begin{pmatrix} x \\ y \end{pmatrix},\quad \vec{u}'=\begin{pmatrix} x' \\ y' \end{pmatrix},\quad \vec{\tau}=\begin{pmatrix} p \\ q \end{pmatrix}, \tag{16.5}$$

le (16.3) si possono scrivere, in forma matriciale compatta,

$$\vec{u}'=A\vec{u}+\vec{\tau}. \tag{16.6}$$

Definizione 16.2 (Affinità)**.** *Si dicono* affinità *(nel piano) le trasformazioni del tipo* (16.3)*, oppure del tipo* (16.6)*, che godano dell'ulteriore importante proprietà*

$$\det(A)\neq 0,\quad \textit{ovvero}\quad ad-bc\neq 0. \tag{16.7}$$

La proprietà (16.7) garantisce che la trasformazione è biunivoca e quindi *invertibile*, ovvero che è sempre possibile ricavare (x,y) da (x',y'). Se si usa la scrittura matriciale questo calcolo è (almeno formalmente) immediato:

$$\vec{u}'=A\vec{u}+\vec{\tau}\ \Rightarrow\ A\vec{u}=\vec{u}'-\vec{\tau}\ \Rightarrow\ A^{-1}A\vec{u}=A^{-1}\big(\vec{u}'-\vec{\tau}\big),$$

ovvero

$$\vec{u}=A^{-1}\big(\vec{u}'-\vec{\tau}\big). \tag{16.8}$$

Comunque anche procedendo direttamente sulle (16.3), per esempio (con qualche cautela) per sostituzione, si ricavano facilmente le formule inverse, che sono dello stesso tipo:

$$\begin{cases} x &= a'x+b'y+p' \\ y &= c'x+d'y+q' \end{cases}, \tag{16.9}$$

con

$$a'd'-b'c'=\frac{1}{ad-bc}\neq 0,\quad \text{ovvero}\quad \det\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}=\det\big(A^{-1}\big)=\frac{1}{\det(A)}\neq 0. \tag{16.10}$$

Definizione 16.3 (Inversa di un'affinità)**.** *L'affinità espressa dalle* (16.8) *oppure dalle* (16.9) *si chiama* inversa dell'affinità (16.3)*, e si indica con* f^{-1}*, se la* (16.3) *era indicata con* f.

Osservazione 16.4. È molto importante osservare che le formule (16.3) consentono di ottenere le coordinate dei punti trasformati (o immagine) a partire dalle coordinate dei punti originari, mentre le formule (16.9) consentono, oltre naturalmente a ricavare le coordinate dei punti originari a partire da quelle dei punti trasformati, *soprattutto* di ottenere le trasformate delle equazioni di luoghi geometrici del piano. L'esempio che segue chiarisce questa osservazione.

Esempio 16.2. Sia data la trasformazione

$$\left\{\begin{array}{lcl} x' & = & 2x-y+1 \\ y' & = & x+y \end{array}\right. , \tag{16.11}$$

che risulta essere un'affinità perché $2\times 1-(-1)\times 1=3\neq 0$. Con calcoli standard si ottiene

$$\left\{\begin{array}{lcl} x & = & \dfrac{1}{3}x'+\dfrac{1}{3}y'-\dfrac{1}{3} \\ y & = & -\dfrac{1}{3}x'+\dfrac{2}{3}y'+\dfrac{1}{3} \end{array}\right. . \tag{16.12}$$

Utilizzando le (16.11) si ottiene, per esempio, che l'immagine del punto $\mathsf{P}=(-1,2)$ è il punto $\mathsf{P}'=(-3,1)$. Utilizzando invece le (16.12) si ottiene, intanto, che per esempio il punto $\mathsf{P}'=(0,0)$ è immagine del punto $\mathsf{P}=(-1/3,1/3)$, ma, *soprattutto* che se considero un'equazione come $x^2+y^2-1=0$ (circonferenza di centro l'origine e raggio 1), essa si trasforma nell'equazione

$$\left(\frac{1}{3}x'+\frac{1}{3}y'-\frac{1}{3}\right)^2+\left(-\frac{1}{3}x'+\frac{2}{3}y'+\frac{1}{3}\right)^2-1=0,$$

ovvero

$$2x'^2+5y'^2-2x'y'-4x'+2y'-7=0,$$

che rappresenta un'ellisse, mostrata nella figura 16.2.

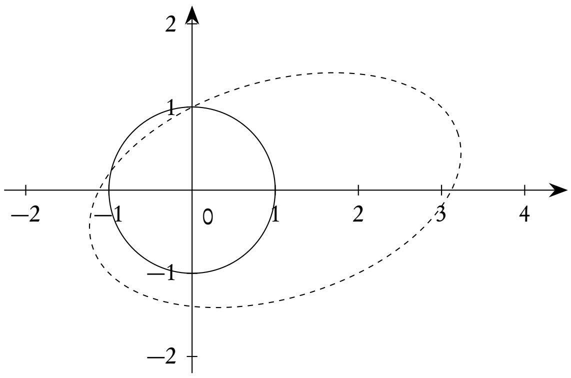


Figura 16.2.: *Circonferenza e sua immagine tramite la trasformazione dell'esempio* 16.2

Per le affinità si può dimostrare il seguente teorema, che elenca le proprietà di cui godono queste trasformazioni del piano in sè.

Teorema 16.5.

1. *L'immagine di una retta è sempre una retta.*
2. *L'immagine di una conica è sempre una conica.*
3. *A rette parallele corrispondono rette parallele.*
4. *A rette incidenti corrispondono rette incidenti.*

5. *Il rapporto delle aree di regioni corrispondenti è costante, si chiama* rapporto di affinità *e vale*

$$r_a = |ad - bc| = |\det(A)|.$$

6. *Le affinità conservano il punto medio dei segmenti, nel senso che se* M *è il punto medio di* $\overline{\mathsf{AB}}$, M′ *è il punto medio di* $\overline{\mathsf{A'B'}}$.

La dimostrazione delle prime due proprietà è immediata, in quanto le (16.9) sono funzioni di primo grado.

Definizione 16.6 (Affinità dirette e inverse). *Se* $\det(A) > 0$ *l'affinità si dice* diretta, *se* $\det(A) < 0$ *l'affinità si dice* inversa.

Dalle proprietà sopra enunciate, e tenendo conto che l'inversa di un'affinità è ancora un'affinità, si deduce subito che un triangolo non degenere si trasforma in un triangolo non degenere. Vediamo su due esempi uno dei motivi della nomenclatura utilizzata nella definizione 16.6.

Esempio 16.3. L'affinità, già considerata nell'esempio 16.2,

$$\begin{cases} x' &= 2x - y + 1 \\ y' &= x + y \end{cases},$$

è diretta, in quanto $\det(A) = 3$. Il triangolo di vertici A $= (-1, 0)$, B $= (2, 0)$, C $= (0, -2)$ e il suo trasformato di vertici A′ $= (-1, -1)$, B′ $= (5, 2)$, C′ $= (3, -2)$ hanno lo stesso orientamento (antiorario) del bordo.

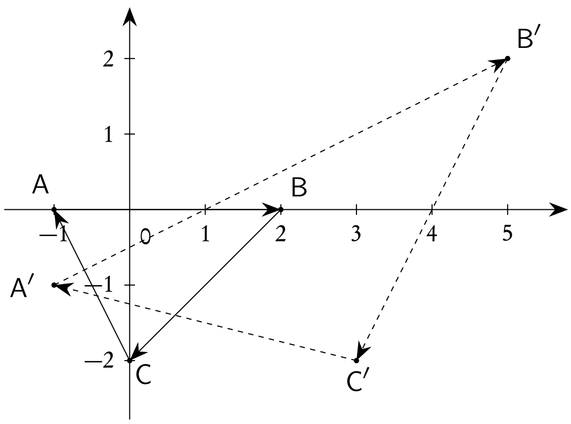


Figura 16.3.: *Affinità diretta*

Esempio 16.4. L'affinità

$$\begin{cases} x' &= 2x - y + 1 \\ y' &= x - y \end{cases},$$

è inversa, in quanto $\det(A) = -1$. Il triangolo di vertici A $= (0, 0)$, B $= (1, 0)$, C $= (0, 1)$ e il suo trasformato di vertici A′ $= (1, 0)$, B′ $= (3, 1)$, C′ $= (0, 1)$ hanno orientamento del bordo opposto.

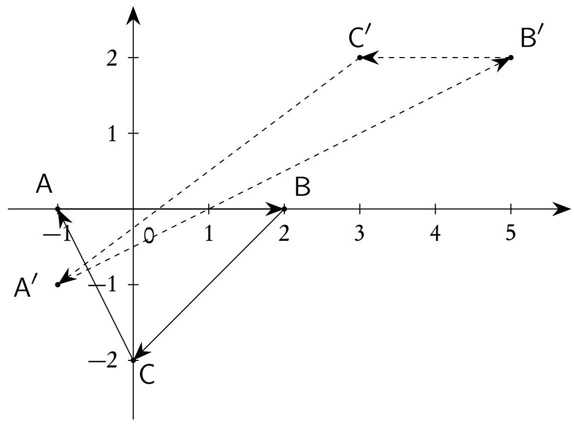


Figura 16.4.: *Affinità inversa*

16.3. Il gruppo delle affinità

Date due affinità

$$f\colon \begin{cases} x' = a_1x + b_1y + p_1 \\ y' = c_1x + d_1y + q_1 \end{cases} \quad \text{e} \quad g\colon \begin{cases} x' = a_2x + b_2y + p_2 \\ y' = c_2x + d_2y + q_2 \end{cases}, \tag{16.13}$$

si possono considerare le affinità *composte* $h = f \circ g$ e $k = g \circ f$. L'affinità h si ottiene sostituendo in f ad x e y rispettivamente $a_2x + b_2y + p_2$ e $c_2x + d_2y + q_2$. L'affinità k si ottiene invece sostituendo in g ad x e y rispettivamente $a_1x + b_1y + p_1$ e $c_1x + d_1y + q_1$. Come risulta ovvio, e come di solito succede quando si considera la composta di due funzioni, non vale la proprietà commutativa: $f \circ g \neq g \circ f$.

Con la scrittura matriciale tutto diventa più semplice. Date

$$f\colon \quad \vec{u}' = A_1\vec{u} + \vec{\tau}_1 \quad \text{e} \quad g\colon \quad \vec{u}' = A_2\vec{u} + \vec{\tau}_2, \tag{16.14}$$

si ha

$$f \circ g\colon \quad \vec{u}' = A_1(A_2\vec{u} + \vec{\tau}_2) + \vec{\tau}_1 = A_1A_2\vec{u} + A_1\vec{\tau}_2 + \vec{\tau}_1, \tag{16.15}$$

e

$$g \circ f\colon \quad \vec{u}' = A_2(A_1\vec{u} + \vec{\tau}_1) + \vec{\tau}_2 = A_2A_1\vec{u} + A_2\vec{\tau}_1 + \vec{\tau}_2, \tag{16.16}$$

espressioni che rendono subito evidente la non commutatività (legata, tra l'altro, alla non commutatività del prodotto matriciale) e che inoltre mostrano come la matrice della trasformazione composta sia il prodotto delle matrici delle due trasformazioni.

Naturalmente l'operazione di composizione gode della proprietà associativa, come sempre succede nella composizione di funzioni e come si può verificare facilmente per calcolo diretto (molto più semplice con la scrittura in forma matriciale). Questo significa che, date tre affinità, f, g, h si ha

$$(f \circ g) \circ h = f \circ (g \circ h). \tag{16.17}$$

Nell'insieme delle affinità esiste anche l'elemento neutro, dato dalla *trasformazione identica*:

$$\text{id:}\left\{\begin{array}{lcl} x' & = & x \\ y' & = & y \end{array}\right. , \quad \text{ovvero} \quad \vec{u}' = \boldsymbol{I}\vec{u} + \vec{0}, \tag{16.18}$$

dove $\boldsymbol{I}$ è la matrice unità e $\vec{0}$ è il vettore nullo:

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{16.19}$$

Se si tiene conto che, come già detto, ogni affinità ha un'inversa (in ragione del fatto che $\det(\boldsymbol{A}) \neq 0$), si conclude subito che l'insieme delle affinità nel piano costituisce un *gruppo non abeliano*.

16.4. Elementi uniti

Definizione 16.7 (Punti uniti). *Data una funzione $f : A \to A$ si chiama* punto fisso *o* punto unito *di f, un elemento x_0 di A che coincide con la sua immagine: $f(x_0) = x_0$.*

Questo concetto ha un notevole interesse in molti campi, tra cui quello delle affinità. Citiamo, perché molto famoso e importante, un teorema relativo ai punti fissi, in un caso molto particolare, quello delle funzioni reali da $[0,1]$ in $[0,1]$:

Teorema 16.8 (di Brouwer, *caso particolare*). *Ogni funzione continua*

$$f : [0,1] \to [0,1]$$

ha almeno un punto fisso.

In questo caso elementare il teorema è graficamente evidente e si può enunciare così: una qualunque funzione continua il cui grafico parta da un punto dell'asse y, con $0 \le y \le 1$, e arrivi su un punto della retta $x = 1$, sempre con $0 \le y \le 1$, incontra almeno una volta la bisettrice del primo e terzo quadrante.

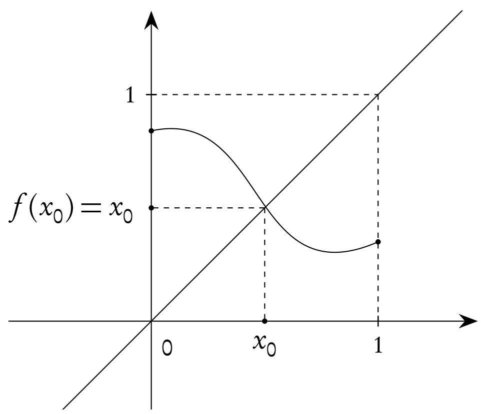


Figura 16.5.: *Punto fisso per una funzione continua $f : [0,1] \to [0,1]$*

È chiaro che la determinazione degli eventuali punti fissi di una funzione si fa risolvendo l'equazione

$$f(x) = x\,. \tag{16.20}$$

Nel caso delle affinità, dove non è affatto detto che esistano punti uniti, si procederà esattamente allo stesso modo: basterà imporre che

$$\begin{cases} ax + by + p &= x \\ cx + dy + q &= y \end{cases}\,. \tag{16.21}$$

Esempio 16.5. Si determinino gli eventuali punti uniti della affinità seguente.

$$\begin{cases} x' &= x + y + 1 \\ y' &= 3x - y \end{cases}\,.$$

Procedendo come indicato, si dovrà risolvere il sistema

$$\begin{cases} x &= x + y + 1 \\ y &= 3x - y \end{cases}\,.$$

Si trova facilmente l'unico punto $\mathrm{P} = (-2/3, -1)$.

Esempio 16.6. Si determinino gli eventuali punti uniti della affinità seguente.

$$\begin{cases} x' &= -y + 1 \\ y' &= -x + 1 \end{cases}\,.$$

Procedendo come indicato, si dovrà risolvere il sistema

$$\begin{cases} x &= -y + 1 \\ y &= -x + 1 \end{cases}\,.$$

Si vede subito che sono uniti tutti i punti della retta $x + y - 1 = 0$: in questo caso dunque i punti uniti sono infiniti. Naturalmente questa retta si potrà anche chiamare *retta* (puntualmente) *unita*.

Oltra agli elementi uniti, nelle applicazioni hanno anche interesse le figure, in particolare le rette, i cui punti vengono trasformati in altri punti della stessa figura. Precisamente si dà la seguente definizione.

Definizione 16.9 (Rette globalmente unite)**.** *Si chiama* retta globalmente unita *una retta i cui punti hanno immagine appartenente alla retta stessa.*

In sostanza data una retta $ax + by + c = 0$ si tratta di verificare se, applicando una data affinità, cioè sostituendo ad x e y le loro espressioni in funzione di x' e y' (affinità inversa), si ottiene ancora la stessa retta, ovvero $kax' + kay' + kb = 0$, con k costante non nulla. Di solito si preferisce lavorare con le rette in forma esplicita, per cui, data la retta $y = mx + q$ e sostituendo in essa ad x e y le loro espressioni in funzione di x' e y', si tratta di verificare che si ottiene una retta con lo stesso coefficiente angolare e con la stessa ordinata all'origine. In questo caso bisognerà però verificare a parte se una retta parallela all'asse y ($x = h$) è o no globalmente unita, cioè si trasforma nella retta $x' = h$. Naturalmente si potrebbe anche utilizzare il procedimento inverso, cioè data la retta $y' = mx' + q$ e sostituendo in essa ad x' e y' le loro espressioni in funzione di x e y si tratta di verificare che si ottiene una retta con lo stesso coefficiente angolare e con la stessa ordinata all'origine.

Esempio 16.7. Si determinino le eventuali rette globalmente unite nell'affinità seguente, già considerata nell'esempio 16.5.

$$\begin{cases} x' &= x+y+1 \\ y' &= 3x-y \end{cases}.$$

Determiniamo intanto l'affinità inversa, procedendo come più sopra indicato. Si trova:

$$\begin{cases} x &= \dfrac{1}{4}x' + \dfrac{1}{4}y' - \dfrac{1}{4} \\ y &= \dfrac{3}{4}x' - \dfrac{1}{4}y' - \dfrac{3}{4} \end{cases}.$$

Una retta del tipo $x = h$ viene mutata in

$$\frac{1}{4}x' + \frac{1}{4}y' - \frac{1}{4} = h,$$

che non è del tipo $x' = h$. Prendiamo allora una retta non verticale $y = mx + q$ e sostituiamo i valori di x e y appena trovati:

$$\frac{3}{4}x' - \frac{1}{4}y' - \frac{3}{4} = m\left(\frac{1}{4}x' + \frac{1}{4}y' - \frac{1}{4}\right) + q.$$

Eseguendo i calcoli e semplificando si trova:

$$(3-m)x' - (1+m)y' - 3 - 4q + m = 0.$$

Il valore $m = -1$ produce una retta trasformata verticale, che dunque non soddisfa le ipotesi (visto che la retta originaria non lo era); per $m \neq -1$ si trova

$$y' = \frac{3-m}{1+m}x' + \frac{m-3-4q}{1+m}.$$

Si deve dunque avere

$$m = \frac{3-m}{1+m} \quad \wedge \quad \frac{m-3-4q}{1+m} = q.$$

Si trovano facilmente le due soluzioni

$$m = -3,\ q = -3 \quad \vee \quad m = 1, q = -\frac{1}{3}.$$

Si osservi che il punto $\mathsf{P} = (-2/3, -1)$, che nell'esempio 16.5 si è dimostrato essere unito per l'affinità che stiamo studiando, è punto di intersezione delle due rette globalmente unite

$$r: \quad y = -3x - 3 \quad \text{ed} \quad s: \quad y = x - \frac{1}{3}$$

appena trovate. La cosa è ovvia: se ogni punto della retta r viene trasformato in un altro punto della stessa retta r, e lo stesso succede per la retta s, il punto P, comune ad r e s, deve avere immagine contemporaneamente su r e su s, cioè deve essere unito.

Procedendo in senso inverso, data la retta non verticale $y' = mx' + q$, sostituiamo i valori di x' e y' con le loro espressioni date in funzione di x ed y. Procedendo come sopra, dopo semplificazione si ottiene:

$$y = \frac{3-m}{1+m}x - \frac{m+q}{1+m},$$

da cui

$$\frac{3-m}{1+m} = m \quad \wedge \quad -\frac{m+q}{1+m} = q,$$

che produce gli stessi valori di prima per m e q.

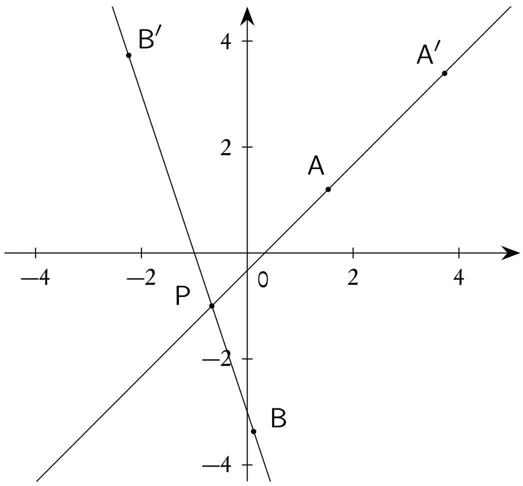


Figura 16.6.: *Elementi uniti per l'affinità dell'esempio 16.7:* A *e* A′, B *e* B′ *sono punti corrispondenti,* P *è punto unito. Le due rette indicate sono globalmente unite*

16.5. Similitudini

Occupiamoci ora di un particolare tipo di affinità, che gode di qualche ulteriore proprietà rispetto alle generiche affinità.

Definizione 16.10 (Similitudini)**.** *Una affinità*

(16.22)
$$\begin{cases} x' = ax + by + p \\ y' = cx + dy + q \end{cases}$$

in cui

(16.23)
$$|a| = |d| \quad \wedge \quad |b| = |c|$$

si chiama una similitudine.

Se teniamo conto che $\det(A) = ad - bc \neq 0$, si può dedurre da (16.23) che le similitudini sono del tipo

$$\begin{cases} x' = ax - by + p \\ y' = bx + ay + q \end{cases}, \tag{16.24}$$

cioè con gli elementi della diagonale principale uguali e quella della diagonale secondaria opposti, oppure del tipo

$$\begin{cases} x' = ax + by + p \\ y' = bx - ay + q \end{cases}, \tag{16.25}$$

cioè con gli elementi della diagonale principale opposti e quella della diagonale secondaria uguali. Il rapporto di affinità è, in entrambi i casi, $r_a = a^2 + b^2$. Naturalmente una similitudine del tipo (16.24) si dirà *diretta*, una del tipo (16.25) si dirà *indiretta*.

Definizione 16.11. *Il numero* $k = \sqrt{a^2 + b^2}$ *si chiama* rapporto di similitudine.

In forma matriciale le (16.24) e (16.25) si scrivono

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}, \quad A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{oppure} \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}. \tag{16.26}$$

Vale il seguente teorema che descrive le proprietà delle similitudini, in aggiunta a quelle delle generiche affinità.

Teorema 16.12.

1. *Il rapporto delle lunghezze di segmenti corrispondenti è costante ed uguale al rapporto di similitudine* k.
2. *Le circonferenze sono mutate in circonferenze.*
3. *Le similitudini conservano gli angoli, nel senso che per ogni terna di punti* A, B, C *l'ampiezza, non necessariamente il verso, dell'angolo* $\mathrm{A\widehat{B}C}$ *è uguale all'ampiezza dell'angolo* $\mathrm{A'\widehat{B'}C'}$.

Esempio 16.8. Consideriamo la similitudine

$$\begin{cases} x' = x - 2y + 1 \\ y' = 2x + y \end{cases}. \tag{16.27}$$

Si tratta di una similitudine diretta di rapporto di similitudine $k = \sqrt{5}$. Poiché, come detto, si conservano gli angoli e i rapporti tra i segmenti e i loro trasformati sono costanti, è ovvio che ogni triangolo viene mutato in un triangolo simile (eventualmente ruotato, traslato e con cambiamento di orientamento sul perimetro). La figura 16.7 illustra questa situazione.

Si osservi (torneremo in seguito su questo fatto) che questa similitudine ha solo il punto $(0, 1/2)$ come punto fisso.

Per evidenziare il significato geometrico delle similitudini è utile riscrivere le equazioni (16.24) e (16.25), oppure (16.26), in un altro modo. Procederemo direttamente sulla matrice A, per comodità di scrittura. Poiché $k \neq 0$, si ha:

$$A = k \begin{pmatrix} \dfrac{a}{k} & -\dfrac{b}{k} \\ \dfrac{b}{k} & \dfrac{a}{k} \end{pmatrix} \quad \text{oppure} \quad A = k \begin{pmatrix} \dfrac{a}{k} & \dfrac{b}{k} \\ \dfrac{b}{k} & -\dfrac{a}{k} \end{pmatrix}$$

Essendo

$$\left(\frac{a}{k}\right)^2 + \left(\frac{b}{k}\right)^2 = \frac{a^2+b^2}{k^2} = 1,$$

se ne deduce che $\exists \alpha$, con $0 \leq \alpha < 2\pi$ tale che

$$\cos\alpha = \frac{a}{k} \quad \text{e} \quad \sin\alpha = \frac{b}{k}. \tag{16.28}$$

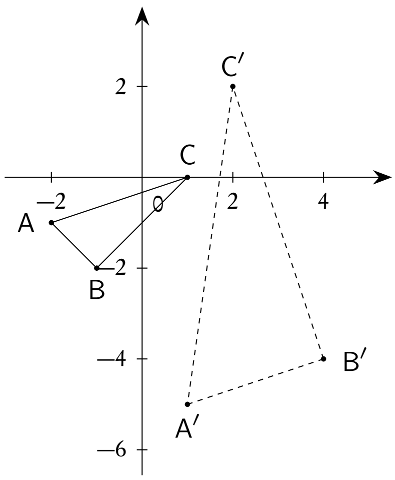


Figura 16.7.: *Le similitudini mutano figure in figure simili*

Dunque le equazioni di una generica similitudine possono scriversi in uno dei due modi seguenti:

$$\begin{cases} x' = k(x\cos\alpha - y\sin\alpha) + p \\ y' = k(x\sin\alpha + y\cos\alpha) + q \end{cases}, \tag{16.29}$$

oppure

$$\begin{cases} x' = k(x\cos\alpha + y\sin\alpha) + p \\ y' = k(x\sin\alpha - y\cos\alpha) + q \end{cases}, \tag{16.30}$$

rispettivamente per le similitudini dirette e inverse e ove $k > 0$ è il rapporto di similitudine.

Esempio 16.9. Sia data la similitudine diretta

$$\begin{cases} x' & = & x+2y+1 \\ y' & = & -2x+y+3 \end{cases} = \begin{cases} x' & = & x-(-2)y+1 \\ y' & = & -2x+y+3 \end{cases},$$

avente $k=\sqrt{5}$. Si noti che abbiamo riscritto la similitudine in modo da uniformarla al modello (16.24): ciò serve per l'interpretazione geometrica che ne ricaveremo. Per arrivare ad una scrittura del tipo (16.29) basterà trovare α tale che

$$\cos\alpha = \frac{1}{\sqrt{5}} \quad \wedge \quad \sin\alpha = \frac{-2}{\sqrt{5}},$$

ovvero

$$\alpha = 2\pi + \arcsin\frac{-2}{\sqrt{5}} \simeq 5.18 \quad (\simeq 297°).$$

Si sarebbe anche potuto prendere

$$\alpha = \arcsin\frac{-2}{\sqrt{5}} \simeq -1.11 \quad (\simeq -63°),$$

se avessimo convenuto, per esempio, che $-\pi \le \alpha < \pi$.

Per concludere questa introduzione alle similitudini segnaliamo che si può dimostrare che vale il seguente teorema, che mostra la struttura dell'insieme delle similitudini.

Teorema 16.13. *L'insieme delle similitudini del piano costituisce un sottogruppo del gruppo delle affinità.*

16.6. Particolari similitudini: le omotetie

Definizione 16.14 (Omotetia). *Una similitudine diretta in cui $b=0$ si chiama una* omotetia*, di cui a si chiama* rapporto di omotetia*. Usando le* (16.29) *si può dire, in maniera equivalente, che un'omotetia è una similitudine diretta con $\alpha=0 \vee \alpha=\pi$.*

Dunque le equazioni di una omotetia sono del tipo:

(16.31)
$$\begin{cases} x' & = & ax+p \\ y' & = & ay+q \end{cases}, \quad a \neq 0,$$

ovvero, in forma matriciale,

(16.32)
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}, \quad A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

Si noti che il rapporto di similitudine è il modulo del rapporto di omotetia: $k=|a|$.

L'inversa di un'omotetia si trova immediatamente. Si ha

(16.33)
$$\begin{cases} x & = & \frac{1}{a}x' - \frac{p}{a} \\ y & = & \frac{1}{a}y - \frac{q}{a} \end{cases}.$$

Anche la ricerca degli elementi uniti di una omotetia è immediata:

$$\begin{cases} x &= ax+p \\ y &= ay+q \end{cases} \quad \Rightarrow \quad \begin{cases} x(1-a) &= p \\ y(1-a) &= q \end{cases},$$

che, se $a \neq 1$, ha una e una sola soluzione

$$C = \left(\frac{p}{1-a}, \frac{q}{1-a}\right), \quad \text{se} \quad a \neq 1. \tag{16.34}$$

Definizione 16.15 (Centro di omotetia)**.** *Il punto* C *dato dalle* (16.34)*, per le omotetie di rapporto diverso da* 1*, si chiama* centro di omotetia*.*

Le omotetie con $a = 1$ rientrano nelle *traslazioni* di cui parleremo in seguito. Le omotetie con $a = -1$ rientrano nelle *rotazioni* (in particolare nei mezzigiri o simmetrie centrali), di cui parleremo in seguito.

Esempio 16.10. Sia data l'omotetia

$$\begin{cases} x' &= 2x+1 \\ y' &= 2y-2 \end{cases}.$$

L'unico punto unito è $\mathsf{C} = (-1, 2)$. Cerchiamo le rette globalmente unite, cominciando da quelle verticali, $x' = h$. Sostituendo si trova

$$2x+1 = h \quad \Rightarrow \quad x = \frac{h-1}{2} \quad \Rightarrow \quad \frac{h-1}{2} = h \quad \Rightarrow \quad h = -1.$$

Passiamo alle rette non verticali $y' = mx' + q$. Sostituendo si trova

$$2y-2 = m(2x+1)+q \quad \Rightarrow \quad y = mx + \frac{2+m+q}{2},$$

da cui si deduce che m può assumere qualunque valore, mentre $q = m+2$. Si tratta delle rette $y = m(x+1)+2$, ovvero di tutte le rette (tranne la verticale) del fascio di centro $\mathsf{C} = (-1, 2)$. Se aggiungiamo la retta verticale precedentemente trovata, troviamo che sono globalmente unite tutte le rette del fascio di centro C.

Vediamo ora come cambiano alcune figure (per esempio i triangoli) del piano: la figura 16.8 mostra, senza bisogno di commenti, che cosa succede.

Le proprietà viste nell'esempio 16.10 sono comuni a tutte le omotetie, come afferma il seguente teorema.

Teorema 16.16. *Le omotetie di rapporto $a \neq 1$ hanno un unico punto unito,* C*, che è il centro dell'omotetia; tutte le rette per* C *sono globalmente unite. Le omotetie conservano le direzioni, nel senso che se* r *è una retta qualsiasi, la sua trasformata è una retta parallela ad* r*. Se $a > 0$ un punto* P *e il suo trasformato* P′ *si trovano sulla stessa semiretta di origine* C*, se $a < 0$ un punto* P *e il suo trasformato* P′ *si trovano sulle semirette opposte di origine* C*.*

In sostanza le omotetie dilatano o contraggono le figure rispetto al centro di omotetia, mantenendone la forma. Tenendo anche conto delle proprietà delle similitudini, che naturalmente sono conservate dalle omotetie, si deduce facilmente che se $a = -1$ le omotetie sono semplicemente delle simmetrie centrali rispetto al centro di omotetia, argomento che tratteremo in seguito, se $|a| > 1$ dilatano la figura, se $(0 <) |a| < 1$ contraggono la figura.

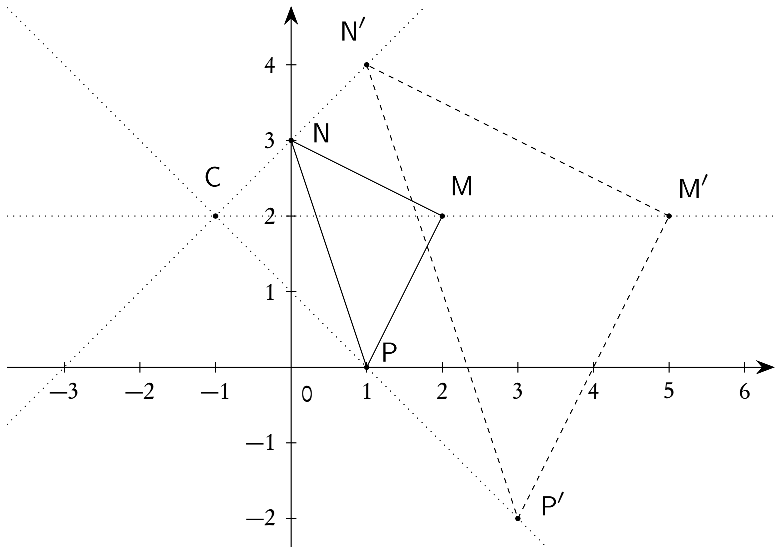


Figura 16.8.: *L'omotetia dell'esempio* 16.10

16.7. Particolari similitudini: le isometrie

Definizione 16.17. *Le similitudini con rapporto di similitudine* $k = 1$ *si chiamano* isometrie. *In particolare quelle con* $\det(A) = a^2 + b^2 = 1$ *si dicono isometrie dirette, quelle con* $\det(A) = -a^2 - b^2 = -1$ *si dicono isometrie inverse.*

Il nome è chiaramente legato al fatto che, essendo $k = 1$, le lunghezze dei segmenti si mantengono nella trasformazione, e quindi le figure geometriche restano inalterate sia nella forma che nelle misure per effetto di una isometria.

Si noti che una similitudine con $\det(A) = \pm 1$ è un'isometria, ma che, invece, una generica affinità con $\det(A) = \pm 1$ non è detto che sia un'isometria. Un esempio è fornito dalla trasformazione

$$\begin{cases} x' &= 2x + y \\ y' &= x + y \end{cases},$$

che ha determinante 1, ma che non è una similitudine e quindi nemmeno una isometria. Lo mostra chiaramente la figura 16.9 dove abbiamo rappresentato un triangolo e il suo trasformato: le aree sono uguali ($r_a = 1$), ma non si tratta chiaramente di una isometria.

Tenendo conto del fatto che $k = 1$, le equazioni delle isometrie si possono scrivere nella forma

$$\begin{cases} x' &= x\cos\alpha - y\sin\alpha + p \\ y' &= x\sin\alpha + y\cos\alpha + q \end{cases}, \tag{16.35}$$

se sono isometrie dirette, e nella forma

$$\begin{cases} x' &= x\cos\alpha + y\sin\alpha + p \\ y' &= x\sin\alpha - y\cos\alpha + q \end{cases}, \tag{16.36}$$

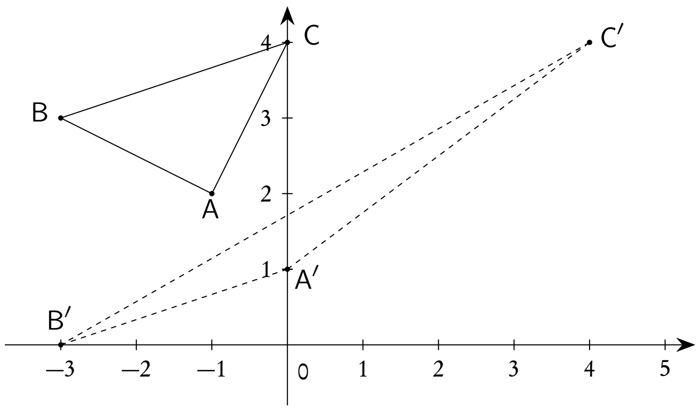


Figura 16.9.: *Affinità che conserva le aree*

se sono isometrie inverse.

Per le isometrie vale il seguente teorema, che mostra la struttura dell'insieme delle isometrie.

Teorema 16.18. *L'insieme delle isometrie del piano costituisce un sottogruppo del gruppo delle similitudini.*

Esaminiamo ora in dettaglio le isometrie, a seconda delle caratteristiche della matrice A.

16.7.1. Traslazioni

Se la matrice A è la matrice identica, ovvero se $\alpha = 0$ con riferimento alle (16.35), l'isometria si scrive

$$\left\{\begin{array}{rcl} x' & = & x+p \\ y' & = & y+q \end{array}\right. , \quad \text{ovvero} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}. \tag{16.37}$$

Come già osservato, questa trasformazione è anche un'omotetia, priva di punti uniti e con rapporto di omotetia 1. Essa si chiama una *traslazione*: le (16.37) mostrano che il punto P' è ottenuto dal punto P mediante la traslazione individuata dal vettore

$$\vec{\tau} = \begin{pmatrix} p \\ q \end{pmatrix}. \tag{16.38}$$

Nel caso particolare in cui $\vec{\tau} = \vec{0}$ la trasformazione stessa si riduce all'identità. Come è ovvio dal significato, e come si può verificare immediatamente, le traslazioni non hanno alcun punto unito. Naturalmente, invece, tutte le rette parallele al vettore $\vec{\tau}$ sono globalmente unite ("scorrono su se stesse").

Esempio 16.11. L'affinità

$$\left\{\begin{array}{rcl} x' & = & x+2 \\ y' & = & y-1 \end{array}\right. , \quad \text{ove} \quad \vec{\tau} = \begin{pmatrix} 2 \\ -1 \end{pmatrix},$$

è una traslazione di vettore $\vec{\tau}$. Nella figura 16.10 è illustrata la traslazione di un triangolo.

Ovviamente l'inversa di una traslazione è ancora una traslazione, di vettore $-\vec{\tau}$. È immediato che l'insieme delle traslazioni è un gruppo (sottogruppo del gruppo delle isometrie).

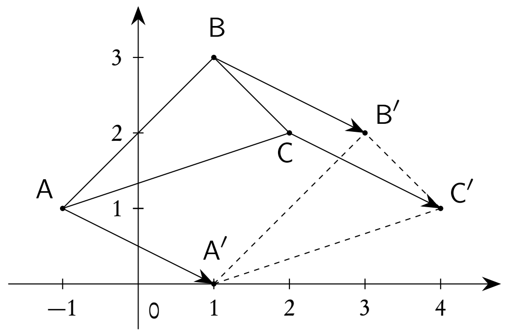


Figura 16.10.: *La traslazione dell'esempio 16.11*

16.7.2. Rotazioni

Consideriamo ora una isometria diretta con matrice diversa dalla matrice unità.

$$\begin{cases} x' &= ax-by+p \\ y' &= bx+ay+q \end{cases}, \quad \text{con} \quad \det(A)=a^2+b^2=1 \tag{16.39}$$

In ragione delle osservazioni precedenti se ne deduce che questo tipo di isometrie si possono sempre scrivere nella forma:

$$\begin{cases} x' &= x\cos\alpha-y\sin\alpha+p \\ y' &= x\sin\alpha+y\cos\alpha+q \end{cases}, \quad \text{ovvero} \quad \begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}+\begin{pmatrix} p \\ q \end{pmatrix}, \tag{16.40}$$

con $0<\alpha<2\pi$: se $\alpha=0$ la matrice $\boldsymbol{A}$ si riduce alla matrice unità e si ricade nel caso delle traslazioni.

Vale il seguente teorema.

Teorema 16.19. *Le isometrie del tipo (16.40), con $\alpha\neq 0$, hanno sempre un unico punto unito.*

Proponiamo la dimostrazione per coloro che conoscono la teoria dei sistemi lineari. Si deve avere

$$\begin{cases} x &= x\cos\alpha-y\sin\alpha+p \\ y &= x\sin\alpha+y\cos\alpha+q \end{cases} \Rightarrow \begin{cases} x(1-\cos\alpha)+y\sin\alpha &= p \\ -x\sin\alpha+y(1-\cos\alpha) &= +q \end{cases}. \tag{16.41}$$

Si tratta di un sistema lineare di 2 equazioni in 2 incognite, con matrice dei coefficienti

$$\boldsymbol{M}=\begin{pmatrix} 1-\cos\alpha & \sin\alpha \\ -\sin\alpha & 1-\cos\alpha \end{pmatrix},$$

avente determinante $2-2\cos\alpha\neq 0$, perché $\alpha\neq 0$. Il sistema ha dunque un'unica soluzione data da

$$x=\frac{\det\begin{pmatrix} p & \sin\alpha \\ q & 1-\cos\alpha \end{pmatrix}}{2-2\cos\alpha}, \quad y=\frac{\det\begin{pmatrix} 1-\cos\alpha & p \\ -\sin\alpha & q \end{pmatrix}}{2-2\cos\alpha}. \tag{16.42}$$

Non è difficile provare, a questo punto, che le trasformazioni del tipo che stiamo considerando sono delle *rotazioni* di un angolo α attorno al punto dato dalle (16.42)[(2)], detto *centro della rotazione*.

[2]Nelle applicazioni è importante determinare il punto unito, ma sconsigliamo dal memorizzare le formule relative: è molto meglio risolvere direttamente il sistema (16.41)

Nel caso particolare $\alpha = \pi$, le (16.40) si riducono alle

$$\begin{cases} x' = -x + p \\ y' = -y + q \end{cases}, \quad \text{ovvero} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}. \tag{16.43}$$

In questo caso la rotazione prende il nome di *mezzogiro* attorno al centro di rotazione o, più comunemente, di *simmetria centrale* di centro il centro di rotazione. Questa trasformazione è anche una omotetia di centro il centro di rotazione e rapporto di omotetia -1.

Le circonferenze con centro nel centro di rotazione sono globalmente unite. Se $\alpha \neq \pi$ non esistono rette globalmente unite, se $\alpha = \pi$ tutte le rette passanti per il centro di rotazione sono globalmente unite (ovvio trattandosi di una particolare omotetia).

Per quanto riguarda la struttura dell'insieme delle rotazioni citiamo (senza ulteriori commenti) solo il fatto che le rotazioni attorno ad uno stesso punto formano gruppo, mentre in generale le rotazioni non formano gruppo.

In ogni caso, se teniamo conto che il prodotto di due affinità ha come matrice il prodotto delle matrici, possiamo osservare che se facciamo il prodotto di due rotazioni di angolo α e β rispettivamente, otteniamo come matrice del prodotto la matrice:

$$\begin{aligned} &\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} = \\ &\qquad = \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{pmatrix} = \\ &\qquad\qquad = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & +\cos(\alpha+\beta) \end{pmatrix}, \end{aligned} \tag{16.44}$$

ovvero che il prodotto di due rotazioni, anche con centro diverso, produce una rotazione di un angolo che è somma degli angoli delle due rotazioni, se questa somma è diversa da 2π, altrimenti produce una traslazione.

Si tenga però ben presente che, in generale, non è possibile invertire l'ordine nel prodotto delle due rotazioni: se si esegue una rotazione di centro C_1 e angolo α e successivamente una rotazione di centro C_2 e angolo β, si ottiene ancora una rotazione, di centro C_3 e angolo $\alpha + \beta$; se si inverte l'ordine delle rotazioni si ottiene una rotazione ancora di angolo $\alpha + \beta$, ma in generale di centro $C_4 \neq C_3$. Solo per le rotazioni attorno allo stesso centro è possibile invertire l'ordine nella composizione.

Esempio 16.12. La trasformazione

$$f: \begin{cases} x' = \dfrac{3}{5}x - \dfrac{4}{5}y + 1 \\[2mm] y' = \dfrac{4}{5}x + \dfrac{3}{5}y \end{cases}$$

è una isometria diretta in quanto

$$\det(A) = \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = 1 .$$

Si ha
$$\alpha = \arcsin\frac{4}{5} \simeq 0.93 \quad (\simeq 53.13°).$$
Determiniamo il punto unito:
$$\begin{cases} x = \dfrac{3}{5}x - \dfrac{4}{5}y + 1 \\ y = \dfrac{4}{5}x + \dfrac{3}{5}y \end{cases},$$
sistema di facile risoluzione che porge $C_1 = (1/2, 1)$. Si tratta dunque di una rotazione dell'angolo α indicato, con centro il punto C_1. La figura 16.11 ne mostra l'effetto, al solito, su un triangolo.

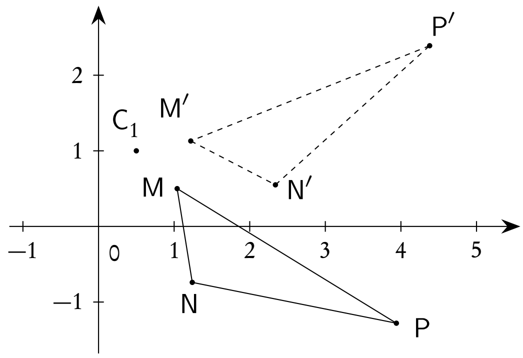


Figura 16.11.: *Rotazione attorno ad un punto*

Esempio 16.13. La trasformazione
$$g: \begin{cases} x' = \dfrac{3}{5}x - \dfrac{4}{5}y \\ y' = \dfrac{4}{5}x + \dfrac{3}{5}y + 1 \end{cases}$$
è una isometria diretta in quanto
$$\det(A) = \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = 1.$$
Si ha
$$\alpha = \arcsin\frac{4}{5} \simeq 0.93 \quad (\simeq 53.13°).$$
Determiniamo il punto unito:
$$\begin{cases} x = \dfrac{3}{5}x - \dfrac{4}{5}y \\ y = \dfrac{4}{5}x + \dfrac{3}{5}y + 1 \end{cases},$$
sistema di facile risoluzione che porge $C_2 = (-1, 1/2)$. Si tratta dunque di una rotazione dell'angolo α indicato, con centro il punto C_2.

Esempio 16.14. Consideriamo ora le composizioni $f\circ g$ e $g\circ f$, dove f e g sono le rotazioni degli esercizi 16.12 e 16.13 rispettivamente. Con facili calcoli si trova:

$$f\circ g:\begin{cases} x' &= -\dfrac{7}{25}x-\dfrac{24}{25}y+\dfrac{1}{5}\\ y' &= \dfrac{24}{25}x-\dfrac{7}{25}y+\dfrac{3}{5}\end{cases},$$

e

$$g\circ f:\begin{cases} x' &= -\dfrac{7}{25}x-\dfrac{24}{25}y+\dfrac{3}{5}\\ y' &= \dfrac{24}{25}x-\dfrac{7}{25}y+\dfrac{9}{5}\end{cases}.$$

Si può constatare che:

— $f\circ g\neq g\circ f$;
— l'angolo di rotazione del prodotto è 2α (le due rotazioni avevano lo stesso α);
— il centro di rotazione di $f\circ g$ è $C_3=(-1/8\,,\,3/8)$;
— il centro di rotazione di $g\circ f$ è $C_4=(-3/8\,,\,9/8)$.

16.7.3. Simmetrie assiali (o riflessioni) e glissoriflessioni

Consideriamo ora una isometria inversa.

(16.45) $$\begin{cases} x' &= ax+by+p\\ y' &= bx-ay+q\end{cases},\quad \text{con}\quad \det(A)=-a^2-b^2=-1$$

che si può scrivere

(16.46) $$\begin{cases} x' &= x\cos\alpha+y\sin\alpha+p\\ y' &= x\sin\alpha-y\cos\alpha+q\end{cases},\quad \text{ovvero}\quad \begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}\cos\alpha & \sin\alpha\\ \sin\alpha & -\cos\alpha\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}p\\q\end{pmatrix},$$

Per quanto riguarda gli elementi uniti, per queste trasformazioni vale il seguente teorema.

Teorema 16.20. *Le trasformazioni del tipo* (16.46) *si dividono in due sole classi:*

1. *quelle che hanno una retta* r *di punti uniti e in questo caso ogni retta perpendicolare ad* r *è globalmente unita;*
2. *quelle che non hanno punti punti e in questo caso hanno una sola retta* s *globalmente unita.*

La retta r *delle isometrie del primo tipo e la retta* s *delle isometrie del secondo tipo hanno coefficiente angolare* $\tan \alpha/2$.

Per coloro che conoscono la teoria dei sistemi lineari proponiamo solo un inizio di dimostrazione di questo teorema. La ricerca dei punti uniti richiede la risoluzione del sistema

$$\begin{cases} x(1-\cos\alpha)-y\sin\alpha &= p\\ -x\sin\alpha+y(1+\cos\alpha) &= q\end{cases}.$$

e questo sistema ha matrice dei coefficienti con determinante nullo: per questo motivo il sistema può avere infinite soluzioni o nessuna soluzione e da qui seguono i due casi citati nel teorema.

In base al teorema 16.20 e ad un esame dettagliato di ulteriori proprietà delle isometrie inverse, che non proponiamo, si può concludere con le seguenti definizioni.

Definizione 16.21 (Simmetria assiale o riflessione)**.** *Una isometria inversa avente infiniti punti uniti è una* riflessione *o* simmetria assiale *rispetto alla retta* r *dei punti uniti, che ha coefficiente angolare* $\tan{}^{\alpha}/_{2}$. *La retta dei punti uniti si chiama* asse della simmetria*. Tutte le rette perpendicolari all'asse di simmetria sono globalmente unite.*

Definizione 16.22 (Glissoriflessione)**.** *Una isometria inversa non avente punti uniti è la composizione di una riflessione e una traslazione di un vettore parallelo all'asse della riflessione (in un ordine qualunque: in questo caso riflessione e traslazione di un vettore parallelo all'asse della riflessione commutano): la trasformazione si chiama una* glissoriflessione*. Esiste una sola retta globalmente unita, di coefficiente angolare* $\tan{}^{\alpha}/_{2}$*, che si può chiamare* asse della glissoriflessione*: questa retta viene fatta scorrere su se stessa dalla trasformazione.*

Esempio 16.15. La trasformazione

$$\begin{cases} x' = y-1 \\ y' = x+1 \end{cases}, \quad \text{ovvero} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

è un'isometria inversa con $\alpha = {}^{\pi}/_{2}$. Andiamo alla ricerca dei punti uniti, risolvendo il sistema

$$\begin{cases} x = y-1 \\ y = x+1 \end{cases}.$$

È immediato che esso ha per soluzioni tutti i punti della retta r: $x-y+1=0$. Si tratta dunque di una simmetria assiale avente per asse la retta r, che ha coefficiente angolare 1, ovvero proprio $\tan{}^{\alpha}/_{2} = \tan{}^{\pi}/_{4}$. È altresì immediato che tutte le rette $y=-x+q$ sono globalmente unite.

Esempio 16.16. La trasformazione

$$\begin{cases} x' = y \\ y' = x+1 \end{cases}, \quad \text{ovvero} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

è un'isometria inversa con $\alpha = {}^{\pi}/_{2}$. Andiamo alla ricerca dei punti uniti, risolvendo il sistema

$$\begin{cases} x = y \\ y = x+1 \end{cases},$$

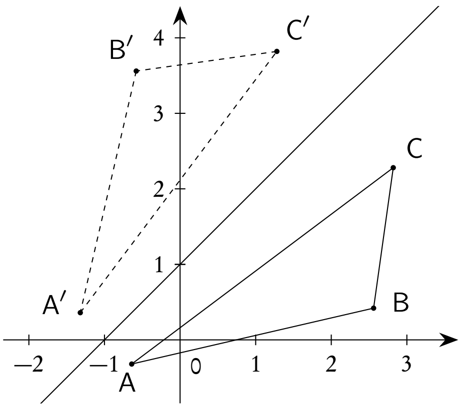
sistema che, palesemente, non ha soluzioni. Andiamo allora alla ricerca delle rette globalmente unite $y' = mx' + q$. Sostituendo i valori di x' e y' in funzione di x e y si trova $x = m(y-1)+q$, ovvero $my = x+m-q$. Scartato il valore $m=0$, che si verifica subito non essere accettabile, si trova

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

da cui

$$\frac{1}{m} = m \wedge \frac{m-q}{m} = q \quad \Rightarrow \quad m=1,\ q=\frac{1}{2}.$$

Si ha dunque la retta globalmente unita $y = x + {}^{1}/_{2}$ che è l'asse della glissoriflessione e ha coefficiente angolare proprio $1 = \tan{}^{\alpha}/_{2} = \tan{}^{\pi}/_{4}$.



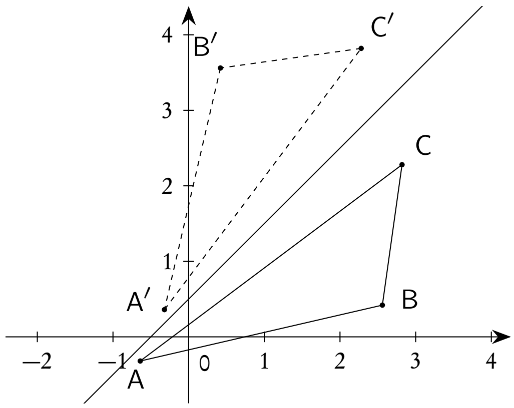


Figura 16.12.: *La simmetria assiale e la glissoriflessione degli esempi* 16.15 *e* 16.16

16.7.4. Simmetrie particolari

Alcuni casi particolari di simmetrie assiali sono importanti nelle applicazioni e meritano un cenno a parte.

Simmetria rispetto all'asse x

È una trasformazione del tipo (16.46), con $\alpha = 0$ e $\vec{\tau} = \vec{0}$. La trasformazione si riduce allora alla seguente:

$$\begin{cases} x' = x \\ y' = -y \end{cases} \tag{16.47}$$

Simmetria rispetto alla retta $y = y_0$

È una trasformazione del tipo (16.46), con $\alpha = 0$ e $\vec{\tau}$ da determinare, tenendo conto che la retta dei punti uniti deve essere $y = y_0$. Si trova subito che deve essere

$$\vec{\tau} = \begin{pmatrix} 0 \\ 2y_0 \end{pmatrix}$$

La trasformazione si riduce allora alla seguente:

$$\begin{cases} x' = x \\ y' = -y + 2y_0 \end{cases} \tag{16.48}$$

Simmetria rispetto all'asse y

È una trasformazione del tipo (16.46), con $\alpha = \pi$ (si ricordi che l'asse di simmetria forma un angolo pari ad $\alpha/2$ con l'asse x) e $\vec{\tau} = \vec{0}$. La trasformazione si riduce allora alla seguente:

$$\begin{cases} x' = -x \\ y' = y \end{cases} \tag{16.49}$$

Simmetria rispetto alla retta $x = x_0$

È una trasformazione del tipo (16.46), con $\alpha = \pi$ e $\vec{\tau}$ da determinare tenendo conto che la retta dei punti uniti deve essere $x = x_0$. Si trova subito che deve essere:

$$\vec{\tau} = \begin{pmatrix} 2x_0 \\ 0 \end{pmatrix}$$

La trasformazione si riduce allora alla seguente:

$$\begin{cases} x' = -x + 2x_0 \\ y' = y \end{cases} \tag{16.50}$$

Simmetria rispetto alla bisettrice $y = x$

È una trasformazione del tipo (16.46), con $\alpha = \pi/2$ (la bisettrice forma un angolo pari a $\pi/4$ con l'asse x) e $\vec{\tau} = \vec{0}$. La trasformazione si riduce allora alla seguente:

$$\begin{cases} x' = y \\ y' = x \end{cases} \tag{16.51}$$

Simmetria rispetto alla bisettrice $y = -x$

È una trasformazione del tipo (16.46), con $\alpha = 3\pi/2$ (la bisettrice forma un angolo pari a $3\pi/4$ con l'asse x) e $\vec{\tau} = \vec{0}$. La trasformazione si riduce allora alla seguente:

$$\begin{cases} x' = -y \\ y' = -x \end{cases} \tag{16.52}$$

Anche le simmetrie centrali, o mezzigiri, di cui abbiamo già parlato nella pagina 528, sono importanti nelle applicazioni e meritano una trattazione più dettagliata.

Simmetria rispetto all'origine

È una trasformazione del tipo (16.43) con $\vec{\tau} = \vec{0}$. Si ottiene dunque:

$$\begin{cases} x' = -x \\ y' = -y \end{cases} \tag{16.53}$$

Simmetria rispetto al punto (x_0, y_0)

È una trasformazione del tipo (16.43) con $\vec{\tau}$ da determinarsi, tenendo conto che il punto unito deve essere (x_0, y_0). Si ottiene facilmente:

$$\begin{cases} x' = -x + 2x_0 \\ y' = -y + 2y_0 \end{cases} \tag{16.54}$$

16.7.5. Osservazioni conclusive sulle isometrie

Gli esempi considerati di isometria esauriscono tutti i casi possibili. Per comodità riportiamo qui uno schema riassuntivo.

Sia f una isometria diversa dall'identità. Allora:

1. se è diretta e non ha punti uniti è una traslazione;
2. se è diretta e ha un solo punto unito è una rotazione attorno al punto unito;
3. se è inversa e ha punti uniti (che sono necessariamente infiniti) è una simmetria assiale con asse di simmetria la retta dei punti uniti;
4. se è inversa e non ha punti uniti è una glissoriflessione e ha una retta globalmente unita che è l'asse della glissoriflessione.

Per le isometrie vale anche il seguente, importante, teorema.

Teorema 16.23. *Ogni isometria del piano è ottenibile come composizione di n simmetrie assiali (o riflessioni), con $n \leq 3$. In particolare:*

- *se $n = 1$ si ha (ovviamente) una simmetria assiale;*
- *se $n = 2$ e i due assi r ed s delle simmetrie f_{r} ed f_{s} sono paralleli, allora la composizione $f_{\mathsf{r}} \circ f_{\mathsf{s}}$ è una traslazione di vettore $\vec{\tau}$ perpendicolare ai due assi e modulo doppio della distanza tra r ed s; il verso di $\vec{\tau}$ è tale che, se si sceglie il suo primo estremo su s (asse della prima simmetria), il secondo estremo appartiene al semipiano individuato da r (asse della seconda simmetria) non contenente s;*
- *se $n = 2$ e i due assi r ed s delle simmetrie sono incidenti in un punto C, allora la composizione delle due simmetrie è una rotazione di centro C e di angolo α doppio dell'angolo, orientato, individuato da s ed r;*
- *se $n = 3$ e l'isometria non ha punti fissi, allora è una glissoriflessione.*

16.8. Ancora sulle similitudini

Come utile approfondimento ed esercizio, enunciamo alcuni ulteriori risultati generali sulle similitudini, che non hanno trovato posto nel paragrafo relativo, in quanto richiedono i concetti relativi alle omotetie e alle isometrie.

Teorema 16.24. *Ogni similitudine f è composizione di una omotetia (di centro qualunque) g e di una isometria h. La scomposizione si può scrivere sia come $f = h \circ g$, sia come $f = g \circ h$, ma l'omotetia e l'isometria che compaiono nelle due scomposizioni saranno, in generale, diverse.*

La dimostrazione di questo teorema è elementare e la proponiamo come utile esercizio. Sia k il rapporto di similitudine. Se $k = 1$ non c'è nulla da dimostrare: la similitudine è essa stessa una isometria, e dunque la composizione di un'isometria con la trasformazione identica (che è un'omotetia). Se $k \neq 1$, consideriamo una omotetia g di centro arbitrario e rapporto $1/k$. La trasformazione composta $h_1 = f \circ g$ è ancora una similitudine (perché composizione di similitudini, che formano un sottogruppo del gruppo delle affinità) ed ha rapporto $k \cdot 1/k = 1$, dunque è un'isometria. Moltiplicando a destra per g^{-1}, che è ancora un'omotetia, si trova:

$$f = h_1 \circ g^{-1}. \tag{16.55}$$

Anche la trasformazione $h_2 = g \circ f$ è un'isometria e moltiplicando questa volta a sinistra per g^{-1} si ottiene:

$$f = g^{-1} \circ h_2. \tag{16.56}$$

Si è dunque ottenuta la decomposizione di f come richiesto. Si tenga presente che $h_2 \neq h_1$, in generale, in quanto il prodotto di affinità non è commutativo.

Si noti anche che, se f è diretta, allora anche h_1 lo e h_2 lo sono (in quanto g è diretta), se f è inversa, allora anche h_1 lo e h_2 lo sono.

Teorema 16.25. *Ogni similitudine che non sia un'isometria ha un unico punto unito.*

La dimostrazione di questo teorema, per chi conosce la teoria dei sistemi lineari, è particolarmente significativa e la proponiamo come utile esercizio. Sia

$$\begin{cases} x' = k(x\cos\alpha - y\sin\alpha) + p \\ y' = k(x\sin\alpha + y\cos\alpha) + q \end{cases},$$

una similitudine diretta e cerchiamo i punti uniti. Si deve risolvere il sistema

$$\begin{cases} x = k(x\cos\alpha - y\sin\alpha) + p \\ y = k(x\sin\alpha + y\cos\alpha) + q \end{cases}, \quad \text{ovvero} \quad \begin{cases} (1 - k\cos\alpha)x + (k\sin\alpha)y = p \\ (-k\sin\alpha)x + (1 - k\cos\alpha)y = q \end{cases}.$$

Il determinante della matrice incompleta A (matrice dei coefficienti) è

$$\det(A) = k^2 + 1 - 2k\cos\alpha.$$

Si ha dunque

$$\det(A) = 0 \Leftrightarrow \frac{k^2+1}{2k} = \cos\alpha, \quad \text{ove} \quad k > 0.$$

Non è difficile provare che

$$\frac{k^2+1}{2k} > 1, \quad \text{se} \quad (0 <)\, k \neq 1.$$

Infatti

$$\frac{k^2+1}{2k} > 1 \Leftrightarrow k^2 + 1 > 2k \Leftrightarrow (k-1)^2\, 0,$$

che è banalmente vera se $k \neq 1$.

Un ragionamento simile si può fare per le similitudini inverse.

Come conseguenza di questi due teoremi e delle proprietà delle omotetie e isometrie, si può concludere con i seguenti risultati generali che evidenziano il significato delle similitudini.

Teorema 16.26. *Ogni similitudine diretta che non sia un'isometria si può scomporre in infiniti modi nel prodotto tra una rotazione e un'omotetia con rapporto di omotetia uguale al rapporto di similitudine. È sempre possibile scegliere come centro sia dell'omotetia che della rotazione l'unico punto unito della similitudine; in questo caso l'ordine in cui si compongono le due trasformazioni non conta. Si esprime brevemente questo fatto dicendo che le similitudini dirette che non sono isometrie sono* rotoomotetie.

Teorema 16.27. *Ogni similitudine inversa che non sia un'isometria si può scomporre in infiniti modi nel prodotto di una simmetria assiale rispetto a una retta e di un'omotetia, di rapporto uguale al rapporto di similitudine. È sempre possibile scegliere come centro dell'omotetia l'unico punto unito,* C*, della similitudine e in questo caso l'asse* r *della simmetria assiale è la retta di coefficiente angolare* $\tan \alpha/2$ *passante per il punto unito. Da qui si deduce che una similitudine inversa ha due rette globalmente unite: la retta* r *e la perpendicolare ad* r *per* C*.*

Esempio 16.17. Se riprendiamo in esame la similitudine dell'esempio 16.8, troviamo che essa ha il punto $\mathsf{M} = (0, 1/2)$ come punto unito. Essa ha inoltre $k = \sqrt{5}$ come rapporto di similitudine. Le equazioni (16.27) della similitudine possono essere riscritte come segue:

$$\begin{cases} x' &= x - 2y + 1 \\ y' &= 2x + y \end{cases} \Rightarrow \begin{cases} x' &= \sqrt{5}\left(\dfrac{1}{\sqrt{5}}x - \dfrac{2}{\sqrt{5}}y\right) + 1 \\ y' &= \sqrt{5}\left(\dfrac{2}{\sqrt{5}}x + \dfrac{1}{\sqrt{5}}y\right) \end{cases},$$

ovvero, posto

$$\alpha = \arcsin\left(\frac{2}{\sqrt{5}}\right),$$

come

$$\begin{cases} x' &= \sqrt{5}(x\cos\alpha - y\sin\alpha) + 1 \\ y' &= \sqrt{5}(x\sin\alpha + y\cos\alpha) \end{cases}.$$

Ebbene la similitudine data può essere scomposta nel prodotto tra un'omotetia di centro M e rapporto $\sqrt{5}$ e una rotazione, attorno a M, di un angolo α, in uno dei due ordini possibili. Si vedano le figure 16.13 e 16.14.

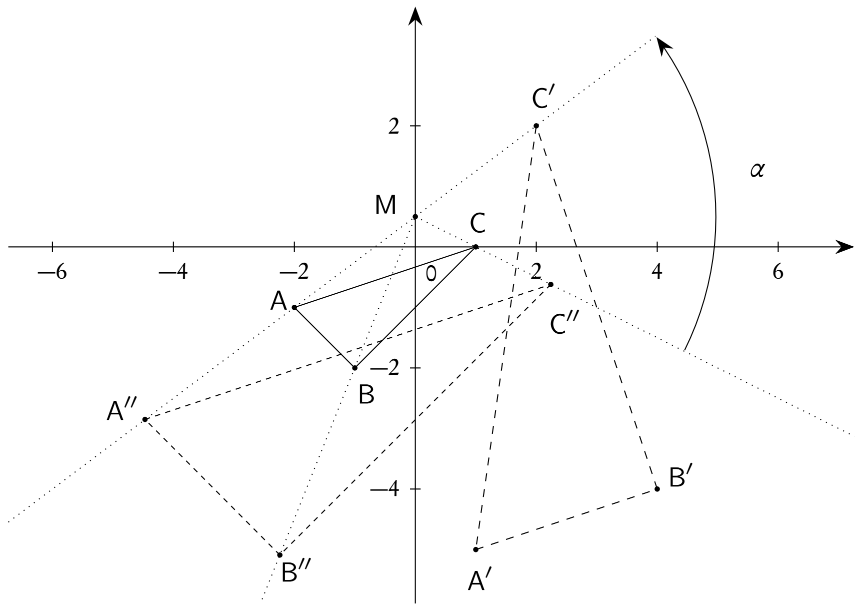


Figura 16.13.: *Similitudine ottenuta applicando prima una omotetia e poi una rotazione di stesso centro*

Esempio 16.18. Riprendiamo in esame nuovamente la similitudine dell'esempio 16.8, che sappiamo avere $k = \sqrt{5}$ come rapporto di similitudine. Indichiamo con f questa similitudine:

$$f: \begin{cases} x' &= x - 2y + 1 \\ y' &= 2x + y \end{cases}.$$

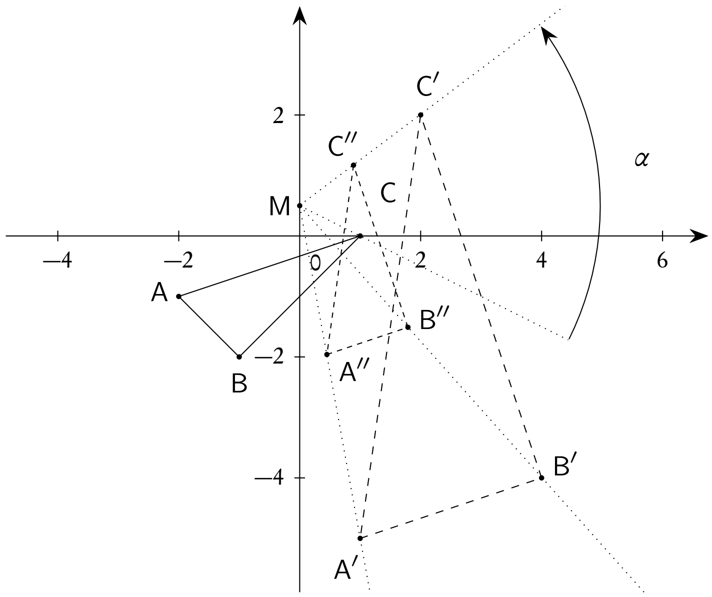


Figura 16.14.: *Similitudine ottenuta applicando prima una rotazione e poi una omotetia di stesso centro*

Vogliamo provare che essa può essere scritta come prodotto di una rotazione e un'omotetia, oppure di un'omotetia e una rotazione, senza usare come centro il punto unito della similitudine. In questo caso non varrà la proprietà commutativa nell'applicazione delle due trasformazioni, come invece visto nell'esempio 16.17.

Per questo consideriamo l'omotetia di rapporto di omotetia $1/\sqrt{5}$ e centro, per esempio, l'origine, omotetia che indichiamo con g:

$$g\colon \begin{cases} x' &= \dfrac{1}{\sqrt{5}}x \\ y' &= \dfrac{1}{\sqrt{5}}y \end{cases} .$$

L'omotetia inversa è

$$g^{-1}\colon \begin{cases} x' &= \sqrt{5}x \\ y' &= \sqrt{5}y \end{cases} .$$

È chiaro che sia $f \circ g$ che $g \circ f$ sono delle isometrie, in quanto in una composizione di similitudini il rapporto di similitudine si ottiene moltiplicando i due rapporti di similitudine[3], e in questo caso si ottiene, per costruzione

$$1 = \sqrt{5} \times \frac{1}{\sqrt{5}} = \frac{1}{\sqrt{5}} \times \sqrt{5}.$$

Poniamo

$$h_1 = f \circ g\,, \quad h_2 = g \circ f\,.$$

[3]Basta ricordare, vedi il paragrafo 16.3, che la matrice della composta di due affinità è il prodotto delle matrici delle due componenti e che il determinante del prodotto è il prodotto dei determinanti.

Se si moltiplica la prima uguaglianza a destra per g^{-1} e la seconda a sinistra sempre per g^{-1}, si ottiene

$$f = h_1 \circ g^{-1}, \quad f = g^{-1} \circ h_2,$$

ottenendo così la decomposizione di f, in due modi diversi, come prodotto di una omotetia e una isometria, o una isometria e una omotetia.

Facendo i conti si trova, facilmente,

$$h_1 = f \circ g = \begin{cases} x' &= \dfrac{1}{\sqrt{5}}x - \dfrac{2}{\sqrt{5}}y + 1 \\ y' &= \dfrac{2}{\sqrt{5}}x + \dfrac{1}{\sqrt{5}}y \end{cases}.$$

Si tratta di una rotazione di centro M_1 e angolo α dati da

$$\mathsf{M}_1 = \left(\frac{1}{2}, \frac{\sqrt{5}+1}{4}\right), \quad \alpha = \arcsin\frac{2}{\sqrt{5}}.$$

Questa decomposizione di f in $h_1 \circ g^{-1}$ è illustrata nella figura 16.15.

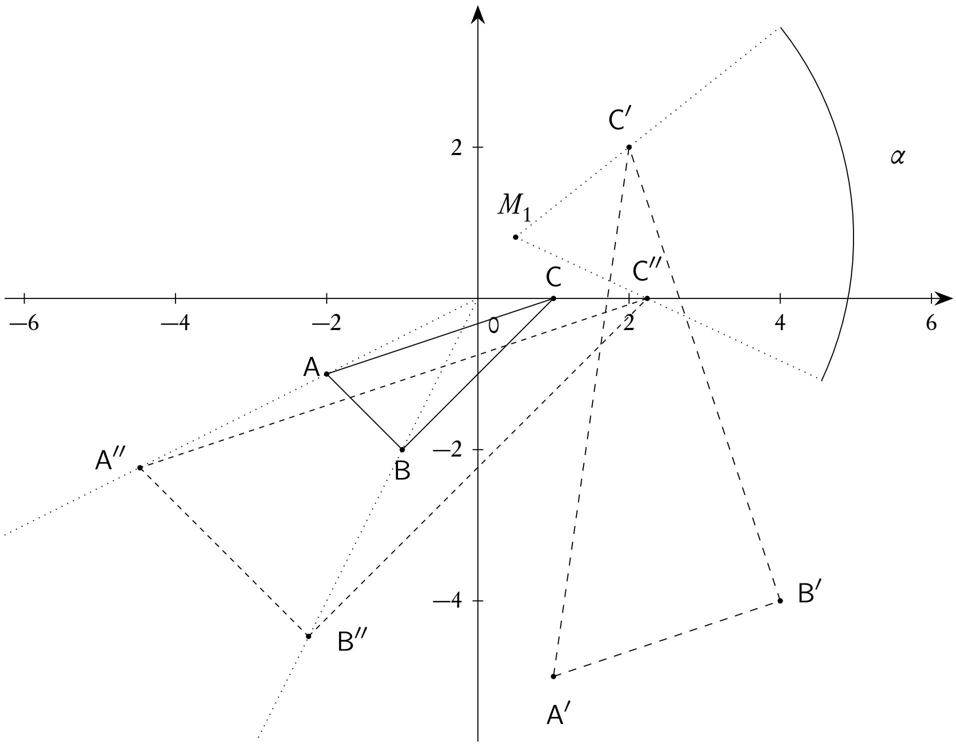


Figura 16.15.: *Decomposizione di una similitudine in omotetia e rotazione con centri diversi*

Analogamente si trova

$$h_2 = g \circ f = \begin{cases} x' & = & \dfrac{1}{\sqrt{5}}x - \dfrac{2}{\sqrt{5}}y + \dfrac{1}{\sqrt{5}} \\ y' & = & \dfrac{2}{\sqrt{5}}x + \dfrac{1}{\sqrt{5}}y \end{cases} .$$

Questa isometria, h_2, è una rotazione di centro M_2 e angolo α dati da

$$\mathsf{M}_2 = \left(\frac{\sqrt{5}}{10}, \frac{5+\sqrt{5}}{20}\right), \quad \alpha = \arcsin \frac{2}{\sqrt{5}}.$$

La decomposizione di f in $g^{-1} \circ h_2$ è rappresentata nella figura 16.16.

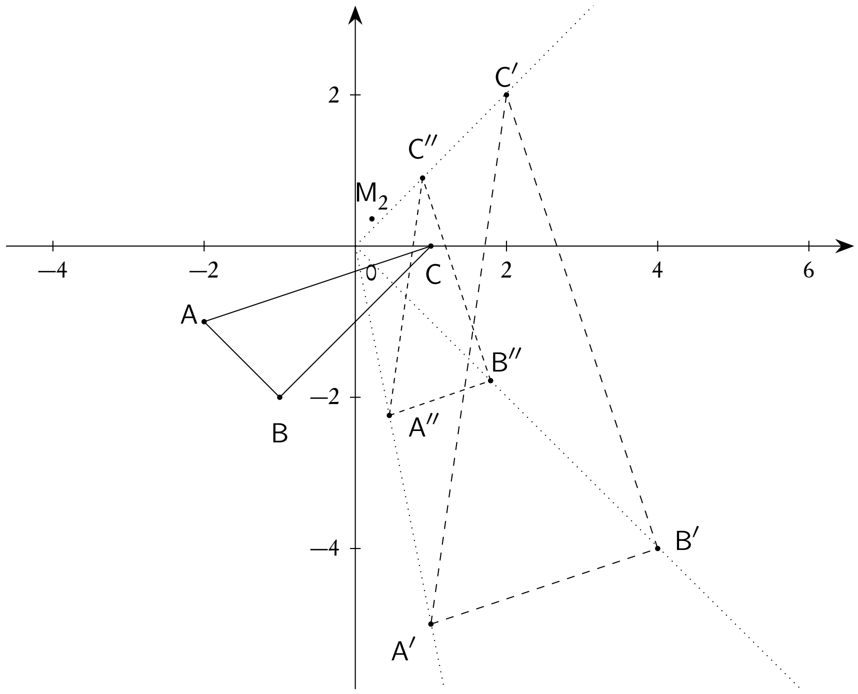


Figura 16.16.: *Decomposizione di una similitudine in rotazione e omotetia con centri diversi*

Esempio 16.19. Siano date la simmetria assiale

$$f : \begin{cases} x' & = & y \\ y' & = & x \end{cases} ,$$

e l'omotetia

$$g : \begin{cases} x' & = & 2x - 2 \\ y' & = & 2y \end{cases} ,$$

di centro $\mathsf{M}=(1,0)$ e rapporto 2. Se componiamo le due trasformazioni otterremo una similitudine inversa. Si ha:

$$s_1=f\circ g:\begin{cases} x' &= 2y\\ y' &= 2x-2\end{cases},$$

e

$$s_2=g\circ f:\begin{cases} x' &= 2y-2\\ y' &= 2x\end{cases}.$$

Calcoli ormai standard provano quanto segue.

La similitudine $s_1=f\circ g$ ha il punto $\mathsf{M}_1=(4/3\,,\,2/3)$ come punto unito; essa ha inoltre come rette unite

$$\mathsf{r}\colon y=x-\frac{2}{3}\quad \text{e}\quad \mathsf{s}\colon y=-x+2,$$

che sono tra di loro perpendicolari e si incontrano proprio nel punto unito. Essendo $\alpha=\pi/2$, la retta r ha coefficiente angolare proprio $\tan\alpha/2=\pi/4=1$. Dunque s_1 si può decomporre nel prodotto (in un ordine qualunque) tra una simmetria assiale, σ_1, rispetto alla retta r e un'omotetia, ω_1, di centro M_1 e coefficiente 2.

La similitudine $s_2=g\circ f$ ha il punto $\mathsf{M}_2=(2/3\,,\,4/3)$ come punto unito; essa ha inoltre come rette unite

$$\mathsf{t}\colon y=x+\frac{2}{3}\quad \text{e}\quad \mathsf{u}\colon y=-x+2,$$

che sono tra di loro perpendicolari e si incontrano proprio nel punto unito. Essendo $\alpha=\pi/2$, la retta t ha coefficiente angolare proprio $\tan\alpha/2=\pi/4=1$. Dunque s_2 si può decomporre nel prodotto (in un ordine qualunque) tra una simmetria assiale, σ_2, rispetto alla retta t e un'omotetia, ω_2, di centro M_2 e coefficiente 2.

L'esempio mostra che s_1, per esempio, si può scrivere come prodotto tra una simmetria assiale e una omotetia in due modi diversi:

$$s_1=f\circ g\quad\vee\quad s_1=\sigma_1\circ\omega_1\,(=\omega_1\circ\sigma_1).$$

La seconda decomposizione è particolarmente significativa, in quanto la simmetria σ_1 e l'omotetia ω_1 commutano, e inoltre l'omotetia trasforma l'asse della simmetria (che era una retta di punti uniti) in una retta globalmente unita, senza spostarla.

16.9. Schema logico per trattare le affinità

A conclusione di questa breve introduzione sulle affinità proponiamo uno schema logico che propone un procedimento standard per esaminare le proprietà di una affinità. Lo schema è costruito utilizzando una specie di chiave politomica di identificazione, adattata a questo speciale problema.

Sia data una trasformazione geometrica affine:

$$\begin{cases} x' &= ax+by+p\\ y' &= bx-ay+q\end{cases},\quad\text{ovvero}\quad \begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}+\begin{pmatrix}p\\q\end{pmatrix}.$$

Poniamo

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\quad\text{e quindi}\quad \det(A)=ad-bc\,.$$

Si può procedere con i passi indicati nella chiave seguente.

$1a$ Se $\det(A) = 0$ la trasformazione *non è un'affinità*.
$1b$ Se $\det(A) \neq 0$ vai al punto 2.

$2a$ Se $A \neq \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \wedge A \neq \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ si tratta di una *generica affinità*.
$2b$ Se $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \vee A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ vai al punto 3.

$3a$ Se $\det(A) = 1$ vai al punto 4.

$4a$ Se non ha punti uniti è una *traslazione*.
$4b$ Se ha un solo punto unito è una *rotazione*.

$3b$ Se $\det(A) = -1$ vai al punto 5.

$5a$ Se ha infiniti punti uniti è una *simmetria assiale*.
$5b$ Se non ha punti uniti è una *glissoriflessione*.

$3c$ Se $\det(A) > 0$, ma $\neq 1$, vai al punto 6.

$6a$ Se $b = 0$ è una *omotetia*.
$6b$ Se $b \neq 0$ è una similitudine *prodotto tra una rotazione e una omotetia*.

$3d$ Se $\det(A) < 0$, ma $\neq -1$, è una similitudine inversa, *prodotto tra una simmetria assiale e una omotetia*.

16.9.1. Qualche indicazione tecnica

Nel caso $2a$ è utile determinare il rapporto di affinità che fornisce il rapporto tra le aree di regioni corrispondenti (l'area della regione trasformata fratto l'area della regione originaria).

Nel caso $4a$ Il vettore

$$\vec{\tau} = \begin{pmatrix} p \\ q \end{pmatrix}$$

è il vettore traslazione. Ogni retta ad esso parallela è globalmente unita.

Nel caso $4b$ il punto unito è il centro della rotazione. Per trovare l'angolo conviene scrivere l'affinità nella forma

$$\begin{cases} x' = x\cos\alpha - y\sin\alpha + p \\ y' = x\sin\alpha + y\cos\alpha + q \end{cases} . \tag{16.57}$$

Non esistono rette globalmente unite.

Nel caso $5a$ la retta dei punti uniti è l'asse della simmetria assiale. Scritta l'equazione nella forma

$$\begin{cases} x' = x\cos\alpha + y\sin\alpha + p \\ y' = x\sin\alpha - y\cos\alpha + q \end{cases} , \tag{16.58}$$

l'asse della simmetria assiale ha coefficiente angolare $\tan \alpha/2$. Tutte le rette perpendicolari all'asse di simmetria sono globalmente unite.

Nel caso $5b$ esiste una retta globalmente unita, che è l'asse della glissoriflessione. La trasformazione è composizione di una simmetria rispetto a quest'asse e una traslazione parallela all'asse.

Nel caso $6a$ l'unico punto unito della trasformazione è il centro della omotetia. Ogni retta per il centro è globalmente unita.

Nel caso $6b$ la trasformazione è una similitudine diretta che si può ottenere come prodotto tra un'omotetia di rapporto uguale al rapporto di similitudine e una rotazione, aventi entrambe centro nell'unico punto unito della trasformazione. Per trovare l'angolo della rotazione conviene scrivere la trasformazione nella forma

$$\left\{ \begin{array}{rcl} x' & = & k(x\cos\alpha - y\sin\alpha) + p \\ y' & = & k(x\sin\alpha + y\cos\alpha) + q \end{array} \right. . \tag{16.59}$$

Nel caso $3d$ la trasformazione è una similitudine inversa che si può ottenere come prodotto tra una simmetria assiale di asse r e un'omotetia di centro $\mathsf{C} \in \mathsf{r}$, che è l'unico punto unito della trasformazione. Una trasformazione di questo tipo ha due rette globalmente unite, tra di loro perpendicolari, di cui una è l'asse della simmetria assiale; queste rette si incontrano nel punto unito. Il coefficiente angolare della retta r è dato da $\tan \alpha/2$, dove α si trova scrivendo la trasformazione nel modo

$$\left\{ \begin{array}{rcl} x' & = & k(x\cos\alpha + y\sin\alpha) + p \\ y' & = & k(x\sin\alpha + y\cos\alpha) + q \end{array} \right. . \tag{16.60}$$

16.10. Isometrie e omotetie: dalla definizione geometrica alle equazioni

In molti casi è abbastanza agevole ricavare le equazioni di una affinità, a partire dalla sua definizione geometrica. Ci occuperemo qui del caso particolare delle isometrie e delle omotetie, perché molto importanti nelle applicazioni e perché i procedimenti relativi sono particolarmente significativi.

Ognuno dei casi trattati si concluderà con una formula specifica: invitiamo il lettore a *non* memorizzare la formula, quanto piuttosto il procedimento applicato (che è sempre elementare) e che può essere replicato senza difficoltà nei casi concreti.

16.10.1. Traslazioni

Una traslazione è individuata da un vettore

$$\vec{\tau} = \begin{pmatrix} p \\ q \end{pmatrix} : \tag{16.61}$$

il punto P' è sempre ottenuto dal punto P per traslazione del vettore $\vec{\tau}$.

La figura 16.17 rende evidente la validità delle formule seguenti per la traslazione di vettore $\vec{\tau}$:

$$\left\{ \begin{array}{rcl} x' & = & x + p \\ y' & = & y + q \end{array} \right. \tag{16.62}$$

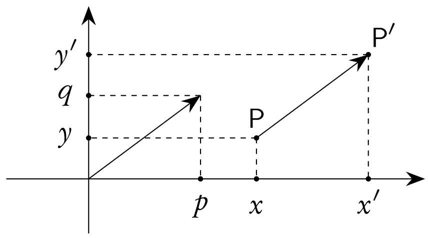


Figura 16.17.: *Traslazione di vettore* $\vec{\tau}$

16.10.2. Rotazioni

Sia $\mathsf{C} = (x_0, y_0)$ il centro della rotazione e α l'angolo di rotazione. Si esamini la figura 16.18.

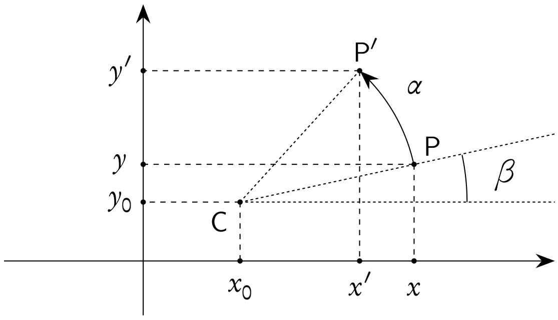


Figura 16.18.: *Rotazione di centro* $\mathsf{C} = (x_0, y_0)$ *e angolo* α

Si ottiene facilmente, posto $r = |\overline{\mathsf{CP}}| = |\overline{\mathsf{CP}'}|$:

$$\begin{cases} x - x_0 & = & r\cos\beta \\ y - y_0 & = & r\sin\beta \end{cases}$$

e

$$\begin{cases} x' - x_0 & = & r\cos(\alpha+\beta) = r\cos\alpha\cos\beta - r\sin\alpha\sin\beta = (x-x_0)\cos\alpha - (y-y_0)\sin\alpha \\ y' - y_0 & = & r\sin(\alpha+\beta) = r\sin\alpha\cos\beta + r\cos\alpha\sin\beta = (x-x_0)\sin\alpha + (y-y_0)\cos\alpha \end{cases}.$$

Da qui si ottengono le formule per la rotazione:

$$\begin{cases} x' & = & x_0 + (x-x_0)\cos\alpha - (y-y_0)\sin\alpha \\ y' & = & y_0 + (x-x_0)\sin\alpha + (y-y_0)\cos\alpha \end{cases}. \tag{16.63}$$

16.10.3. Simmetrie assiali

Le simmetrie rispetto a rette parallele agli assi e rispetto alle bisettrici $y = x$ e $y = -x$ sono elementari.

- La simmetria rispetto all'asse x lascia invariata la x e cambia la y in $-y$: $x' = x$, $y' = -y$.
- La simmetria rispetto all'asse y lascia invariata la y e cambia la x in $-x$: $x' = -x$, $y' = y$.

— La simmetria rispetto alla retta orizzontale $y = y_0$ si trova tenendo conto che il punto P e il punto P′ hanno la stessa ascissa e che l'ordinata del punto medio tra P e P' deve essere y_0: $x' = x, (y+y')/2 = y_0$, da cui $x' = x$, $y' = -y + 2y_0$.
— La simmetria rispetto alla retta verticale $x = x_0$ si trova tenendo conto che il punto P e il punto P′ hanno la stessa ordinata e che l'ascissa del punto medio tra P e P′ deve essere x_0: $y' = y$, $(x+x')/2 = x_0$, da cui $x' = -x + 2x_0$, $y' = y$.
— La simmetria rispetto alla bisettrice $y = x$ scambia la x con la y: $x' = y$, $y' = x$.
— La simmetria rispetto alla bisettrice $y = -x$ scambia la x con la y e inverte anche i segni: $x' = -y$, $y' = -x$.

Per quanto riguarda la simmetria rispetto a una retta r: $y = mx + q$, non orizzontale né verticale, essa si ricava tenendo conto che il punto medio tra P e il trasformato P′ deve stare sulla retta r e che inoltre il coefficiente angolare della retta per P e P′ deve essere $1/m$ (r e $\overline{\mathrm{PP}'}$ sono ortogonali).

$$\frac{y'+y}{2} = m\frac{x'+x}{2} + q \quad \land \quad \frac{y'-y}{x'-x} = \frac{1}{m}.$$

Svolgendo i calcoli si trova:

$$\left\{\begin{array}{lcl} x' & = & \dfrac{1}{1+m^2}\Big((1-m^2)x + 2my - 2mq\Big) \\ y' & = & \dfrac{1}{1+m^2}\Big(2mx + (1-m^2)y + 2q\Big) \end{array}\right. . \tag{16.64}$$

16.10.4. Omotetie

Sia $\mathrm{C} = (x_0, y_0)$ il centro dell'omotetia, e k il rapporto di omotetia. Basterà tenere conto che il vettore $\overrightarrow{\mathrm{CP}'}$, che ha componenti $(x' - x_0, y' - y_0)$ deve essere multiplo, secondo il fattore k (positivo o negativo) del vettore $\overrightarrow{\mathrm{CP}}$, che ha componenti $(x - x_0, y - y_0)$. Si ottiene:

$$\left\{\begin{array}{lcl} x' & = & x_0 + k(x - x_0) \\ y' & = & y_0 + k(y - y_0) \end{array}\right. . \tag{16.65}$$

16.11. Esercizi

Esercizio 16.1. *Si provi esplicitamente che le affinità conservano il punto medio dei segmenti, nel senso che se* M *è il punto medio di* $\overline{\mathrm{AB}}$, M′ *è il punto medio di* $\overline{\mathrm{A'B'}}$, *dove* M′, A′, B′ *sono i trasformati di* A, B, C.

Risoluzione. Sia

$$\left\{\begin{array}{lcl} x' & = & ax + by + p \\ y' & = & cx + dy + q \end{array}\right.$$

un'affinità e $\mathrm{A} = (x_\mathrm{A}, y_\mathrm{A})$, $\mathrm{B} = (x_\mathrm{B}, y_\mathrm{B})$ due punti. Siano inoltre $\mathrm{A}' = (x'_{\mathrm{A}'}, y'_{\mathrm{A}'})$, $\mathrm{B} = (x'_{\mathrm{B}'}, y'_{\mathrm{B}'})$ i loro trasformati. Si ha:

$$\mathrm{M} = \left(\frac{x_\mathrm{A} + x_\mathrm{B}}{2}, \frac{y_\mathrm{A} + y_\mathrm{B}}{2}\right) \quad \text{e} \quad \mathrm{M}' = \left(\frac{x'_{\mathrm{A}'} + x'_{\mathrm{B}'}}{2}, \frac{y'_{\mathrm{A}'} + y'_{\mathrm{B}'}}{2}\right).$$

Utilizzando le formule della trasformazione si ottiene

$$\frac{x'_{A'}+x'_{B'}}{2}=\frac{ax_A+by_A+p+ax_B+by_B+p}{2}=a\frac{x_A+x_B}{2}+b\frac{y_A+y_B}{2}+p=ax_M+by_M+p\,,$$

da cui si conclude che l'ascissa di M' è l'immagine dell'ascissa di M. Analogo calcolo per l'ordinata. □

Esercizio 16.2. *Si provi, seguendo la strategia dell'esercizio 16.1, che le affinità conservano anche il baricentro dei triangoli, nel senso che se* G *è il baricentro di* ABC, G' *è il baricentro di* $A'B'C'$, *con le stesse notazioni dell'esercizio* 16.1.

Esercizio 16.3. *Determinare gli elementi uniti della trasformazione*

$$\begin{cases} x' &= 2x \\ y' &= x+y \end{cases}.$$

Risoluzione. Per i punti uniti si deve risolvere il sistema:

$$\begin{cases} x &= 2x \\ y &= x+y \end{cases},$$

che ha come soluzioni tutti i punti con $x=0$, ovvero l'asse y. Cerchiamo ora le rette unite. Per quelle verticali, $x'=h$[4], si deve avere $2x=h$, ovvero $x=h/2$; da qui si deduce $h=h/2$, ovvero $h=0$: l'unica retta globalmente unita è l'asse y, che avevamo già trovato come retta puntualmente unita. Esaminiamo le rette non verticali, $y'=mx'+q$. Sostituendo si trova $x+y=m(2x)+q$, ovvero $y=(2m-1)x+q$. Si trova $m=1$ e q arbitrario. Sono globalmente unite le rette $y=x+q$. □

Esercizio 16.4. *Determinare l'omotetia che muta la parabola* $y=x^2$ *nella parabola* $y=x^2/2+x+1/2$

Risoluzione. Si ha

$$\begin{cases} x' &= ax+p \\ y' &= ay+q \end{cases}.$$

Come al solito, anziché cercare l'inversa della omotetia, partiamo dall'equazione trasformata e otteniamo:

$$ay+q=\frac{(ax+p)^2}{2}+ax+p+\frac{1}{2},\quad \text{ovvero}\quad y=\frac{a}{2}x^2+(p+1)x+\frac{p^2}{2a}+\frac{p}{a}+\frac{1}{2a}-\frac{q}{a}\,.$$

Si deve dunque avere

$$\frac{a}{2}=1\quad\wedge\quad p+1=0\quad\wedge\quad \frac{p^2}{2a}+\frac{p}{a}+\frac{1}{2a}-\frac{q}{a}=0,$$

da cui si ricava $a=2$, $p=-1$, $q=0$. □

[4] Per evitare di dover trovare l'affinità inversa, si può partire dalle rette trasformate, anziché dalle rette originarie, come abbiamo già osservato.

Esercizio 16.5. *Si determini la trasformazione del tipo*

$$\begin{cases} x' &= ax+by \\ y' &= cx+dy \end{cases}$$

che muta i punti $\mathsf{A}(1,1)$ *e* $\mathsf{B}(1,0)$ *rispettivamente nei punti* $\mathsf{A}'(0,2)$ *e* $\mathsf{B}'(1,0)$*. Si determinino le caratteristiche della trasformazione e i punti uniti.*

Risoluzione. Si deve avere

$$\begin{cases} 0 &= a+b \\ 2 &= c+d \\ 1 &= a \\ 0 &= c \end{cases}, \quad \text{ovvero} \quad a=1,\, b=-1,\, c=0,\, d=2\,:$$

$$\begin{cases} x' &= x-y \\ y' &= 2y \end{cases}$$

Per i punti uniti si deve risolvere il sistema

$$\begin{cases} x &= x-y \\ y &= 2y \end{cases}$$

che ha come soluzioni i punti con $y=0$ e x qualunque, ovvero i punti dell'asse x, che risulta puntualmente unita.

Cerchiamo ora le rette unite; partendo da $y'=mx'+q$ e sostituendo si trova $2y=m(x-y)+q$ ovvero $y(m+2)=mx+q$. Si trova che deve essere

$$\frac{m}{m+2}=m \quad \wedge \quad \frac{q}{m+2}=q$$

che ha come soluzioni $m=0 \wedge q=0$ e $m=-1 \wedge q$ qualunque. Dunque sono unite la retta $y=0$ (che è anche puntualmente unita) e le rette del fascio $y=-x+q$. Si verifica anche facilmente che non esistono rette unite verticali. □

Esercizio 16.6 (Interessante, ma impegnativo)**.** *Determinare le trasformazioni lineari (ovvero con* $\vec{\tau}=\vec{0}$*) che trasformano la circonferenza* $x^2+y^2=1$ *nell'ellisse* $4x^2+25y^2=100$*.*

Risoluzione. Sia

$$\begin{cases} x' &= ax+by \\ y' &= cx+dy \end{cases}$$

l'affinità. Potremmo cercare l'inversa dell'affinità e sostituire in $x^2+y^2=1$ al posto di x e y le loro espressioni in funzione di x' e y'. Possiamo però anche operare sull'equazione dell'ellisse, che scriviamo nella forma $4x'^2+25y'^2=100$. Si ottiene

$$4(ax+by)^2+25(cx+dy)^2=100\,, \quad \text{ovvero} \quad (4a^2+25c^2)x^2+(4b^2+25d^2)y^2+(8ab+50cd)xy=100\,.$$

Si può dunque concludere che deve essere

$$\frac{4a^2+25c^2}{100}=1 \wedge \frac{4b^2+25d^2}{100} \wedge 8ab+50cd=0,$$

ovvero

$$\begin{cases} \dfrac{a^2}{25}+\dfrac{c^2}{4} & = & 1 \\ \dfrac{b^2}{25}+\dfrac{d^2}{4} & = & 1 \\ \dfrac{ab}{25}+\dfrac{cd}{4} & = & 0 \end{cases}.$$

La prima e la seconda equazione implicano che devono esistere β e γ tali che

$$\cos\beta=\frac{a}{5} \wedge \sin\beta=\frac{c}{2} \quad \Rightarrow \quad a=5\cos\beta \wedge c=2\sin\beta,$$

e

$$\cos\gamma=\frac{b}{5} \wedge \sin\gamma=\frac{d}{2} \quad \Rightarrow \quad b=5\cos\gamma \wedge d=2\sin\gamma.$$

Sostituendo nella terza si trova

$$\cos\beta\cos\gamma+\sin\beta\sin\gamma=0 \quad \Rightarrow \quad \cos(\beta-\gamma)=0 \quad \Rightarrow \quad \beta=\gamma+\frac{\pi}{2}+k\pi.$$

Per $k=0$ si ottiene:

$$\begin{cases} x' & = & -5x\sin\gamma+5y\cos\gamma \\ y' & = & 2x\cos\gamma+2y\sin\gamma \end{cases}.$$

Per $k=1$ si ottiene:

$$\begin{cases} x' & = & -5x\sin\gamma+5y\cos\gamma \\ y' & = & -2x\cos\gamma+2y\sin\gamma \end{cases}.$$

Si trovano dunque infinite trasformazioni che soddisfano le condizioni richieste. Non sarebbe difficile provare che queste trasformazioni sono la composizione di una rotazione di un angolo qualunque γ o di una riflessione attorno a una retta qualunque per l'origine, con una trasformazione che dilati le x di 5 unità e le y di due unità. La cosa è geometricamente evidente. □

Esercizio 16.7 (Interessante ma impegnativo)**.** *Trovare le affinità che mutano l'ellisse*

$$\frac{(x-1)^2}{9}+\frac{(y+2)^2}{4}=1 \tag{16.66}$$

nella circonferenza

$$x^2+y^2=1. \tag{16.67}$$

Risoluzione. Una soluzione elementare (ma parziale!) del problema consiste nell'osservare che, sicuramente, le affinità dei tipi

$$\begin{cases} x' = \pm\dfrac{x-1}{3} \\ y' = \pm\dfrac{y+2}{2} \end{cases} \tag{16.68}$$

soddisfano le condizioni richieste: si tratta di traslare il centro dell'ellisse nel punto $(0,0)$ e in seguito di contrarre le x e le y rispettivamente di 3 e 2 unità. Purtroppo questo ragionamento non ci garantisce di aver trovato tutte le soluzioni. Consideriamo dunque le equazioni di una generica affinità

$$\begin{cases} x' = ax+by+p \\ y' = cx+dy+q \end{cases} \tag{16.69}$$

e sostituiamo le (16.69) nell'equazione (16.67); otteniamo, dopo semplificazione,

$$(a^2+c^2)x^2+(b^2+d^2)y^2+2(ab+cd)xy+2(ap+cq)x+2(bp+dq)y+p^2+q^2-1=0. \tag{16.70}$$

Confrontiamo la (16.70) con la (16.66), che per comodità riscriviamo eseguendo i calcoli indicati:

$$\frac{1}{9}x^2+\frac{1}{4}y^2-\frac{2}{9}x+y+\frac{1}{9}=0. \tag{16.71}$$

Deve dunque essere:

$$\begin{cases} a^2+c^2={}^1/_9 \\ b^2+d^2={}^1/_4 \\ ab+cd=0 \\ ap+cq=-{}^1/_9 \\ bp+dq={}^1/_2 \\ p^2+q^2-1={}^1/_9 \end{cases}. \tag{16.72}$$

Si tratta di un sistema di 6 equazioni nelle 6 incognite a, b, c, d, p, q; poiché tutte le equazioni sono di 2° grado, il sistema ha grado 64(!!). Un sistema siffatto è in genere irrisolubile; se però riscriviamo le prime due equazioni del sistema nella forma

$$(3a)^2+(3c)^2=1 \quad \wedge \quad (2b)^2+(2d)^2=1,$$

notiamo che devono esistere α e β tali che

$$3a=\cos\alpha\,,\, 3c=\sin\alpha \quad \wedge \quad 2b=\cos\beta\,,\, 2d=\sin\beta\,,$$

ovvero

$$a=\frac{1}{3}\cos\alpha\,,\, c=\frac{1}{3}\sin\alpha \quad \wedge \quad b=\frac{1}{2}\cos\beta\,,\, d=\frac{1}{2}\sin\beta\,. \tag{16.73}$$

Sostituendo le (16.73) nella terza delle (16.72) si trova

$$\frac{1}{6}\cos\alpha\cos\beta+\frac{1}{6}\sin\alpha\sin\beta=0 \quad\Rightarrow\quad \cos(\alpha-\beta)=0 \quad\Rightarrow\quad \alpha=\beta\pm\frac{\pi}{2}.$$

Trattiamo prima il caso

$$\alpha=\beta+\frac{\pi}{2},$$

che porge

$$a=-\frac{1}{3}\sin\beta\,,\; c=\frac{1}{3}\cos\beta\,. \tag{16.74}$$

Sostituendo i valori di a, c, b, d dati dalle (16.74) e (16.73) nella quarta e quinta delle (16.72) e semplificando si trova

$$\begin{cases} p\sin\beta-q\cos\beta=\dfrac{1}{3} \\ p\cos\beta+q\sin\beta=1 \end{cases}. \tag{16.75}$$

Da qui si possono facilmente ricavare p e q, sempre in funzione di β, per esempio usando Cramer (o per sostituzione). Si ottiene:

$$p=\frac{1}{3}\sin\beta+\cos\beta \quad\wedge\quad q=\sin\beta-\frac{1}{3}\cos\beta\,. \tag{16.76}$$

Sostituendo questi valori nell'ultima delle (16.72) si ottiene un'identità: β rimane dunque arbitrario. In conclusione, tutte le affinità del tipo

$$\begin{cases} x' &= -\dfrac{1}{3}x\sin\beta+\dfrac{1}{2}y\cos\beta+\dfrac{1}{3}\sin\beta+\cos\beta \\ y' &= \dfrac{1}{3}x\cos\beta+\dfrac{1}{2}y\sin\beta+\sin\beta-\dfrac{1}{3}\cos\beta \end{cases} \tag{16.77}$$

soddisfano il problema.

Procedendo esattamente allo stesso modo con il caso

$$\alpha=\beta-\frac{\pi}{2},$$

si ottiene un secondo insieme di affinità:

$$\begin{cases} x' &= \dfrac{1}{3}x\sin\beta+\dfrac{1}{2}y\cos\beta-\dfrac{1}{3}\sin\beta+\cos\beta \\ y' &= -\dfrac{1}{3}x\cos\beta+\dfrac{1}{2}y\sin\beta+\sin\beta-\dfrac{1}{3}\cos\beta \end{cases}. \tag{16.78}$$

È immediato verificare che le (16.77) ed (16.78), per $\beta={}^{\pi}/_{2}$ e per $\beta={}^{3\pi}/_{2}$, forniscono le soluzioni particolari trovate all'inizio. In ogni caso, però, il problema ha infinite soluzioni. □

17. Grafici costruibili per via elementare

In questo capitolo proponiamo una rassegna di tecniche per tracciare grafici di funzioni reali di variabile reale, in un sistema cartesiano ortogonale, senza l'uso del calcolo infinitesimale, anche se, occasionalmente, dovremo usare alcuni termini come "asintoto", limite, ecc., limitandoci però solo a concezioni intuitive.

Queste tecniche rivestono grande importanza nelle applicazioni e consentono anche di poter effettuare rapidamente controlli sulla correttezza dei risultati ottenuti con i tradizionali metodi dell'analisi che saranno oggetto dei corsi di studio universitari.

Alcune delle considerazioni qui proposte sono già state trattate in altre parti di questo testo, e sono comunque ripresentate per questioni di completezza.

In tutti i casi che considereremo supporremo di avere tracciato, con qualche tecnica, il grafico di una o più funzioni e ci proporremo di dedurre da essi il grafico di altre funzioni, con semplici operazioni "grafiche".

17.1. Simmetrie

Data una funzione $f(x)$ di cui sappiamo tracciare il grafico, per ottenere il grafico di $f(-x)$ basta considerare il simmetrico di quello di $f(x)$ rispetto all'asse y, per ottenere il grafico di $-f(x)$ basta considerare il simmetrico di quello di $f(x)$ rispetto all'asse delle ascisse. In sostanza, scritta la funzione nella forma $y = f(x)$, la sostituzione di x con $-x$ fornisce la funzione simmetrica rispetto all'asse y, la sostituzione di y con $-y$ fornisce la funzione simmetrica rispetto all'asse x.

È di particolare interesse il caso in cui i grafici di $f(x)$ e $f(-x)$ coincidono: le funzioni con questa proprietà si chiamano *pari*. Se invece coincidono i grafici di $f(x)$ e $-f(-x)$ la funzione si chiama *dispari*. Giova osservare che una funzione non può essere pari o dispari se il suo dominio non è simmetrico rispetto all'origine; la funzione $f(x)$ dell'esempio 17.1 non può essere né pari né dispari, in quanto il dominio naturale è costituito da $\mathbb{R}^+$. Esempi classici di funzioni pari e dispari sono, rispettivamente, la funzione coseno e la funzione seno.

Esempio 17.1. Le funzioni $f(x) = \sqrt{x}$, $g(x) = f(-x) = \sqrt{-x}$, $h(x) = -f(x) = -\sqrt{x}$ hanno i grafici mostrati nella figura 17.1.

Data di nuovo una funzione $f(x)$ di cui sappiamo tracciare il grafico, è facile tracciarne il simmetrico rispetto a una retta verticale di equazione $x = k$ o a una retta orizzontale di equazione $y = k$. Chiamiamo $g(x)$ e $h(x)$, rispettivamente, le funzioni che hanno questi grafici e osserviamo che, se $\mathsf{P}_1(x_1, y_1)$ è un punto sul grafico di $f(x)$, $\mathsf{P}_2(x_2, y_2)$ è il suo simmetrico rispetto a $x = k$ sul grafico di $g(x)$ e $\mathsf{P}_3(x_3, y_3)$ è il suo simmetrico rispetto a $y = k$ sul grafico di $h(x)$, si deve avere

$$\frac{x_1 + x_2}{2} = k \Rightarrow x_2 = -x_1 + 2k \,,\; y_2 = y_1 \quad \text{e} \quad x_3 = x_1 \,,\; \frac{y_1 + y_3}{2} = k \Rightarrow y_3 = -y_1 + 2k.$$

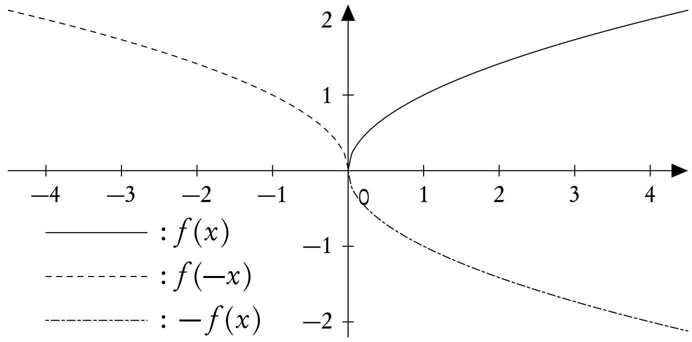


Figura 17.1.: *Grafici di* $f(x) = \sqrt{x}$, $g(x) = f(-x) = \sqrt{-x}$, $h(x) = -f(x) = -\sqrt{x}$

Dunque per ottenere l'espressione analitica di $y = g(x)$ basterà sostituire x con $-x+2k$ in quella di $y = f(x)$, per ottenere l'espressione analitica di $y = h(x)$ basterà sostituire y con $-y+2k$ in quella di $y = f(x)$, ovvero sostituire $f(x)$ con $-f(x)+2k$.

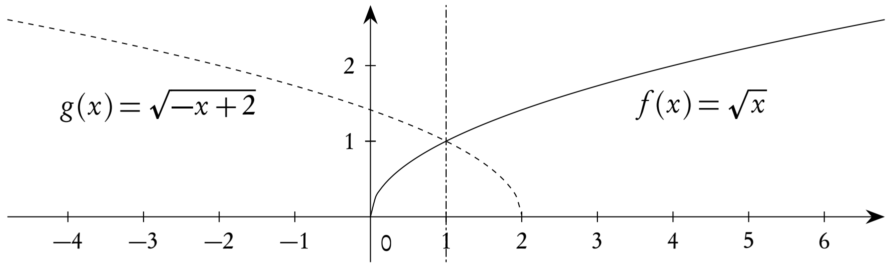


Figura 17.2.: *Grafico di una funzione e della sua simmetrica rispetto alla retta* $x = 1$

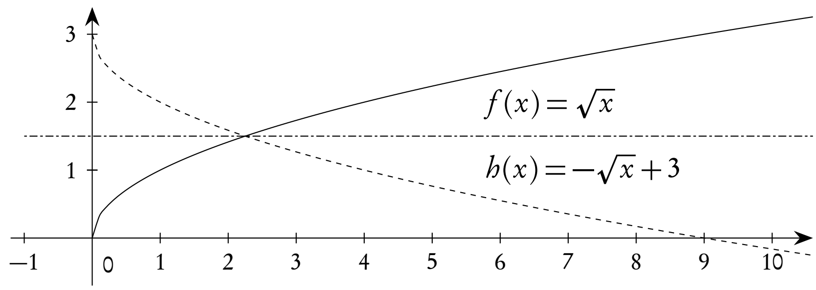


Figura 17.3.: *Grafico di una funzione e della sua simmetrica rispetto alla retta* $y = 3/2$

17.2. Traslazioni

Data una funzione $f(x)$ di cui sappiamo tracciare il grafico, per ottenere il grafico di $g(x) = f(x)+k$ basta traslare rigidamente il grafico di $f(x)$ di un tratto verticale k, traslazione che deve avvenire verso l'alto se $k > 0$, verso il basso se $k < 0$.

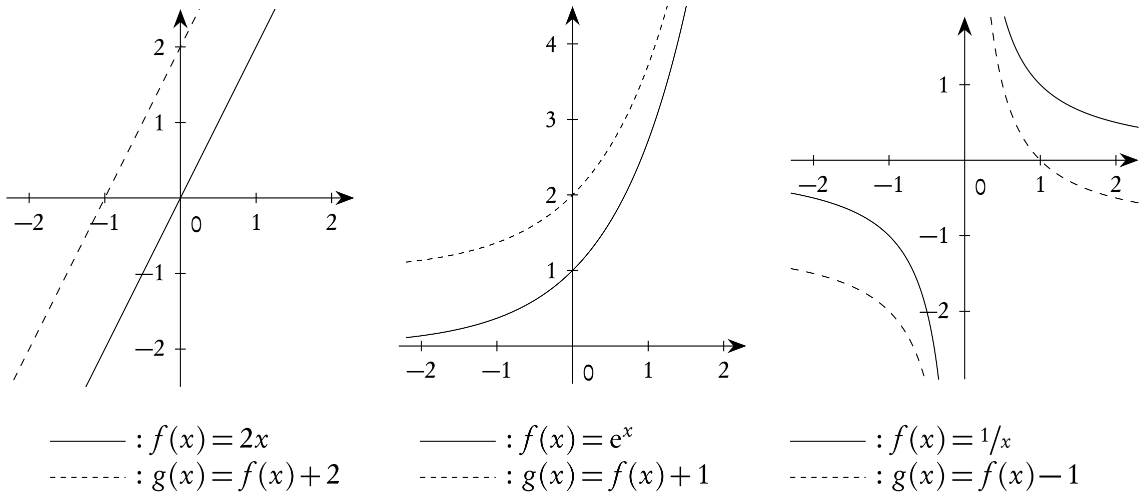


Figura 17.4.: *Grafici di funzioni e di loro traslate verticali*

Data di nuovo una funzione $f(x)$ di cui sappiamo tracciare il grafico, per ottenere il grafico di $g(x)=f(x+k)$ basta traslare rigidamente il grafico di $f(x)$ di un tratto orizzontale k, traslazione che deve avvenire verso sinistra se $k>0$, verso destra se $k<0$.

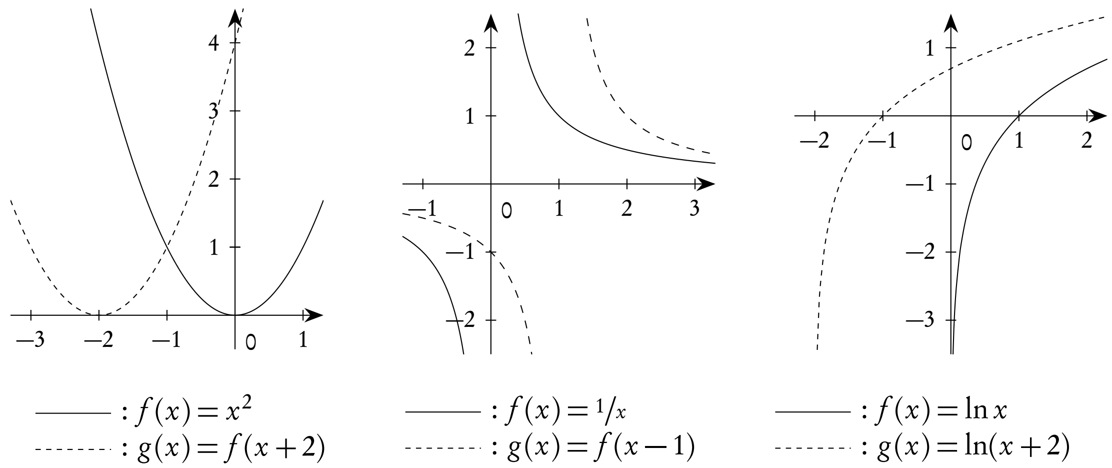


Figura 17.5.: *Grafici di funzioni e di loro traslate orizzontali*

17.3. Cambiamenti di scala

Data una funzione $f(x)$ di cui sappiamo tracciare il grafico, vogliamo tracciare il grafico di $kf(x)$, essendo k un numero reale qualunque. Se $k>0$ si tratta semplicemente di moltiplicare per un fattore k tutte le ordinate della f. Se $k<0$ si ha anche un ribaltamento del grafico rispetto all'asse x. In particolare si ottiene uno "schiacciamento" se $|k|<1$, una "dilatazione" se $|k|>1$. Se $k=1$ naturalmente non si ha alcun effetto, se $k=-1$ si tratta di una semplice simmetria rispetto all'asse x, situazione che abbiamo già considerato. Prestare attenzione al fatto che, ovviamente, i punti sull'asse x non vengono spostati (hanno ordinata zero!).

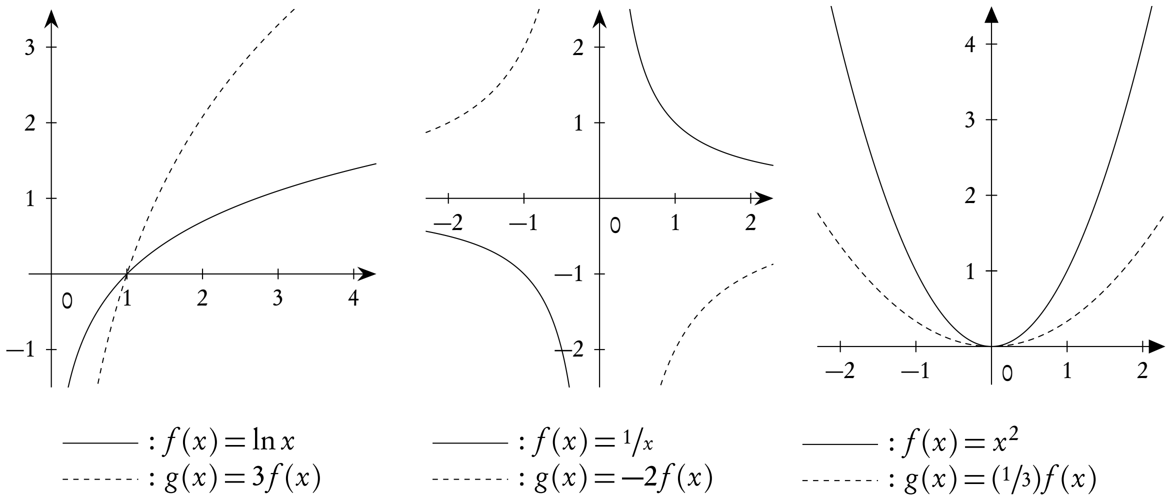


Figura 17.6.: *Grafici di funzioni e cambiamenti di scala sull'asse delle ordinate*

Data di nuovo una funzione $f(x)$ di cui sappiamo tracciare il grafico, vogliamo tracciare il grafico di $f(kx)$, essendo k un numero reale qualunque. Per capire come vanno le cose basta osservare che se per $x = 1$ nella funzione originaria si ottiene un certo valore di y, ora lo stesso valore si ottiene per $x = 1/k$. Se $k > 1$ si tratta dunque di un effetto di schiacciamento orizzontale sul grafico, se $0 < k < 1$ di un effetto di dilatazione orizzontale. Se $k = 1$ naturalmente non si ha alcun effetto. Se $k < 0$ a tutto questo va aggiunta una simmetria rispetto all'asse y. Se $k = -1$ si tratta di una semplice simmetria rispetto all'asse y, situazione che abbiamo già considerato.

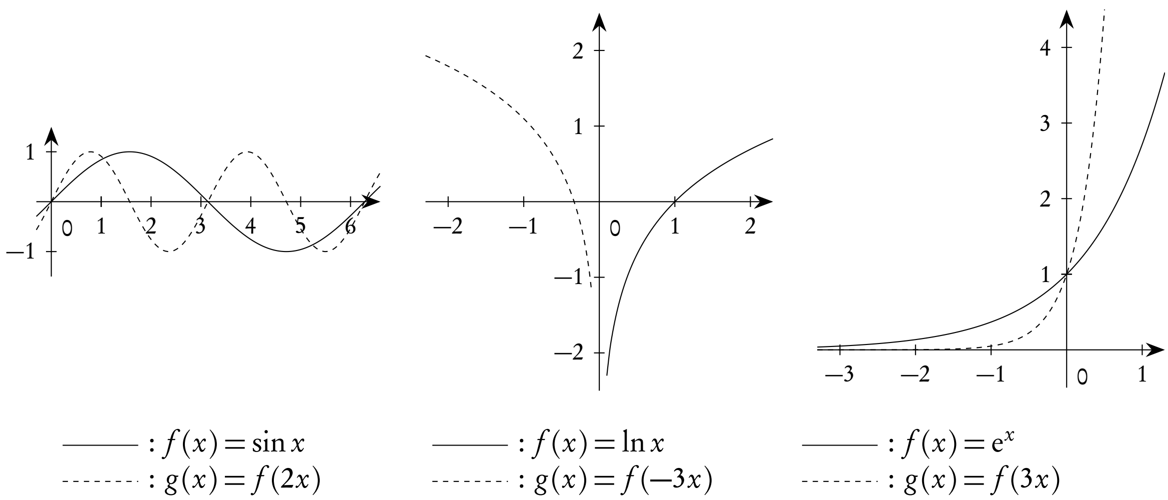


Figura 17.7.: *Grafici di funzioni e cambiamenti di scala sull'asse delle ascisse*

Esempio 17.2. Si vuole dedurre il grafico di $y = 2x^2 + 4x - 1$ da quello di $y = x^2$.

Si comincia con lo scrivere la curva come $y = 2(x^2 + 2x) - 1$. Completando il quadrato tra parentesi si ha poi $y = 2(x+1)^2 - 3$. Si possono allora considerare, nell'ordine, i grafici di $y = x^2$, $y = (x+1)^2$, $y = 2(x+1)^2$, $y = 2(x+1)^2 - 3$, utilizzando le tecniche proposte. Si veda la figura 17.8.

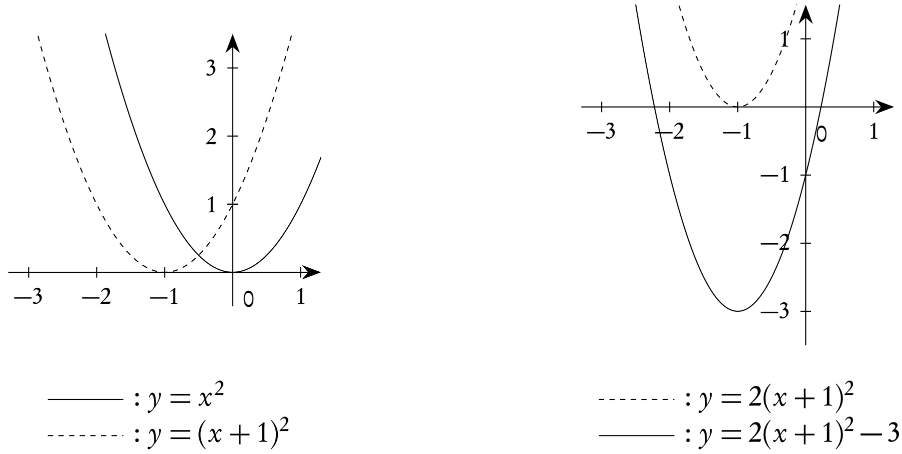


Figura 17.8.: *Grafico di $f(x) = 2x^2 + 4x - 1$ a partire da quello di $f(x) = x^2$*

Il grafico della funzione dell'esempio 17.2 si poteva anche tracciare direttamente, trattandosi di una parabola: abbiamo voluto comunque applicare le tecniche studiate, per poter poi avere un facile controllo.

Occorre prestare la massima attenzione alle combinazioni delle operazioni sopra considerate. Per esempio per tracciare il grafico di $f(2x+1)$ a partire da quello di $f(x)$ si può procedere in uno dei due modi seguenti, come illustrato nella figura 17.9:

— tracciare il grafico di $f(x+1)$ (traslazione verso sinistra di 1 unità) e poi tracciare il grafico di $f(2x+1)$ (schiacciamento orizzontale del grafico di un fattore 2);
— tracciare il grafico di $f(2x)$ (schiacciamento orizzontale del grafico di un fattore 2) e poi tracciare il grafico di $f(2x+1) = f(2(x+1/2))$ (traslazione verso sinistra di $1/2$ unità).

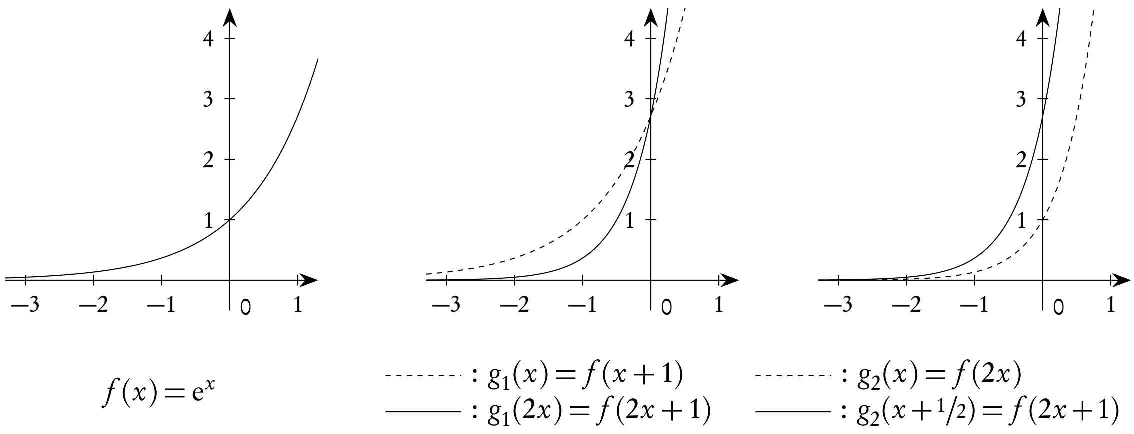


Figura 17.9.: *Diverse combinazioni di traslazioni e cambiamenti di scala*

17.4. Valori assoluti

La presenza di valori assoluti nel problema di tracciare grafici di funzioni per via elementare può prevedere diverse situazioni: cominciamo con il considerare le due situazioni principali.

1. Data una funzione $f(x)$, tracciare il grafico di $|f(x)|$;
2. data una funzione $f(x)$, tracciare il grafico di $f(|x|)$.

Si tratta di due problemi estremamente semplici; basta ricordare la definizione di valore assoluto per concludere come segue:

1. il grafico di $|f(x)|$ si ottiene semplicemente ribaltando, rispetto all'asse x, quella parte del grafico di $f(x)$ che "sta sotto" l'asse stesso;
2. il grafico di $f(|x|)$ si ottiene considerando solo la parte del grafico di $f(x)$ ove $x \geq 0$ e prolungandolo per simmetria rispetto all'asse y, dove $x \leq 0$.

Un esempio è proposto nella figura 17.10.

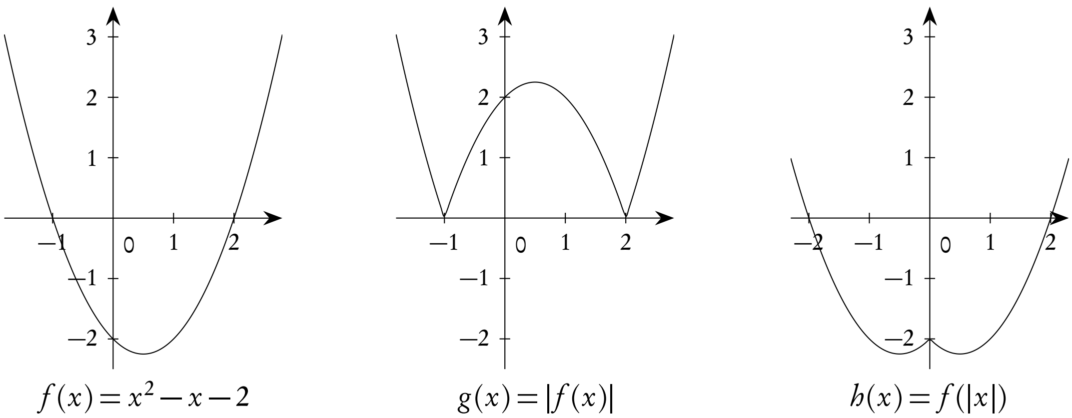


Figura 17.10.: *Grafici di $f(x) = x^2 - x - 2$, $|f(x)|$ e $f(|x|)$*

Altre situazioni in cui compaiono valori assoluti si possono trattare semplicemente tenendo conto della definizione di valore assoluto. Proponiamo alcuni esempi per chiarire il metodo.

Esempio 17.3. Per tracciare il grafico di $f(x) = x + |x-1| - |2x+1|$ si comincia con l'osservare che

$$|x-1| = \begin{cases} -(x-1), & \text{se } x < 1 \\ x-1, & \text{se } x \geq 1 \end{cases}, \quad |2x+1| = \begin{cases} -(2x+1), & \text{se } x < -1/2 \\ 2x+1, & \text{se } x \geq -1/2 \end{cases}.$$

Questo consente di riscrivere la funzione come

$$f(x) = \begin{cases} 2x+2, & \text{se } x < -1/2 \\ -2x, & \text{se } -1/2 \leq x < 1 \\ -2, & \text{se } x \geq 1 \end{cases}.$$

Il grafico è dunque costituito da parti di grafici di funzioni lineari (cioè di rette), tutte facilmente tracciabili. Si veda la figura 17.11.

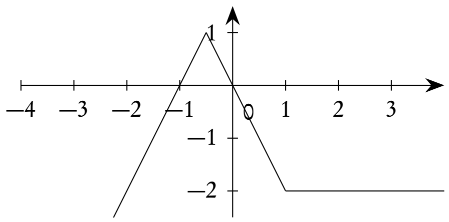


Figura 17.11.: *Grafico di* $f(x) = x + |x-1| - |2x+1|$

Esempio 17.4. Per tracciare il grafico di $f(x) = x^2 + x - |3x+1|$ si può osservare che

$$|3x+1| = \begin{cases} -(3x+1), & \text{se } x < -1/3 \\ 3x+1, & \text{se } x \geq -1/3 \end{cases}.$$

Questo consente di riscrivere la funzione $f(x)$ come

$$f(x) = \begin{cases} x^2 - 2x - 1, & \text{se } x \geq -1/3 \\ x^2 + 4x + 1, & \text{se } x < -1/3 \end{cases}.$$

Il grafico è proposto nella figura 17.12.

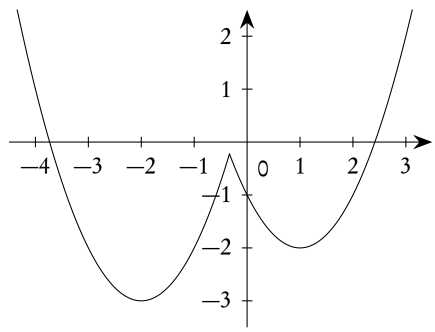


Figura 17.12.: *Grafico di* $f(x) = x^2 + x - |3x+1|$

17.5. Operazioni tra funzioni

Una delle tecniche frequenti per costruire nuovi grafici a partire da grafici noti è quella di considerare somme, prodotti, quozienti, composizioni, potenze, di funzioni note. Si tratta di un problema non semplice e che richiede attenzione e prudenza. Ci limiteremo a considerare solo alcune situazioni molto semplici e relative a funzioni “abbastanza regolari”.

17.5.1. Il passaggio al reciproco

Si tratta, data una funzione $f(x)$, di ottenere il grafico della funzione $1/f(x)$. Si può osservare che in tutti gli intervalli ove $f(x)$ cresce ed è di segno costante, il suo reciproco decresce, e viceversa. Per una trattazione significativa è opportuno anche fare qualche considerazione, seppure a livello intuitivo, su alcuni "comportamenti asintotici". In particolare ci interessano le seguenti osservazioni:

— nei punti in cui $f(x)$ vale zero, il reciproco non è definito e tende all'infinito (quindi con presenza di asintoti verticali);
— nei pressi dei punti dove $f(x)$ tende all'infinito, il reciproco tende a zero.

La figura 17.13 illustra il passaggio dal grafico della funzione $f(x) = x^2 - x - 2$ a quello della sua reciproca.

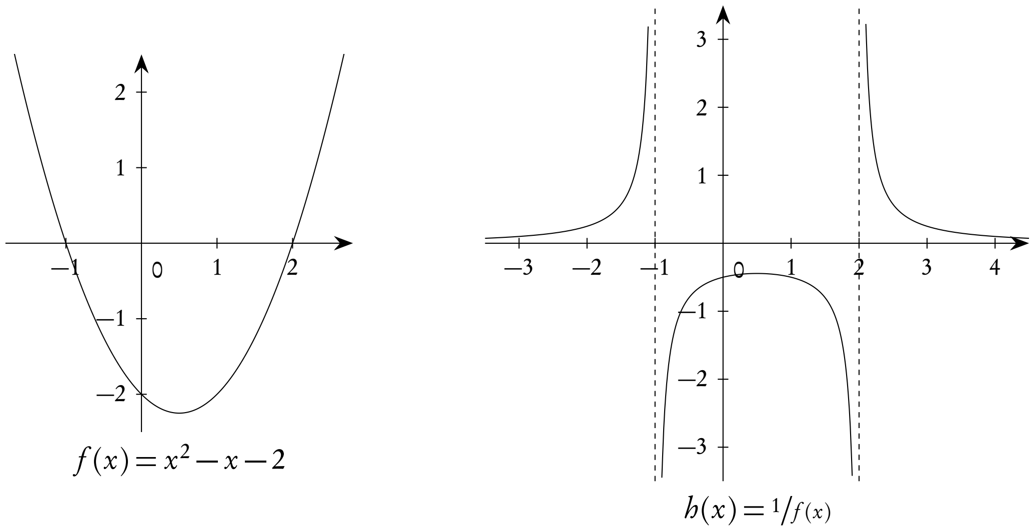


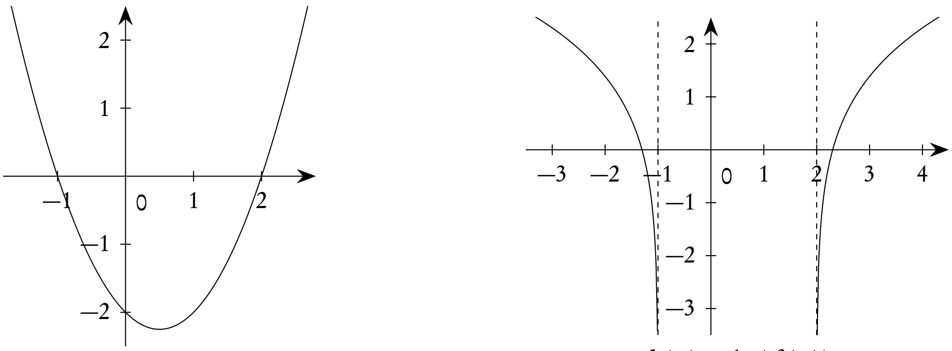
Figura 17.13.: *Grafico di $f(x) = x^2 - x - 2$ e della sua reciproca*

17.5.2. Il logaritmo naturale di una funzione

Poiché la funzione logaritmo naturale è strettamente crescente nel suo dominio, data una funzione $f(x)$, la funzione $\ln(f(x))$ sarà crescente o decrescente esattamente come $f(x)$. Naturalmente occorrerà tenere conto che $\ln(f(x))$ è definita solo dove $f(x)$ è positiva. Per una trattazione significativa è opportuno anche fare qualche considerazione, seppure a livello intuitivo, su alcuni "comportamenti asintotici". In particolare ci interessano le seguenti osservazioni:

— nei punti in cui $f(x)$ vale zero, il $\ln(f(x))$ non è definito e tende a meno infinito (dal lato ove $f(x)$ è positiva), quindi con presenza di asintoto verticale;
— nei pressi dei punti dove $f(x)$ tende a più infinito, anche $\ln(f(x))$ tende a più infinito.

Nel caso di logaritmi in altre basi, basterà solo ricordare che se la base è maggiore di uno, il logaritmo è crescente (e quindi si comporta come il logaritmo naturale), mentre se la base è compresa tra zero ed uno il logaritmo è decrescente e quindi il comportamento di $\ln(f(x))$ sarà opposto a quello di $f(x)$. La figura 17.14 illustra il passaggio dal grafico della funzione $f(x) = x^2 - x - 2$ a quello della funzione $\ln(x^2 - x - 2)$.



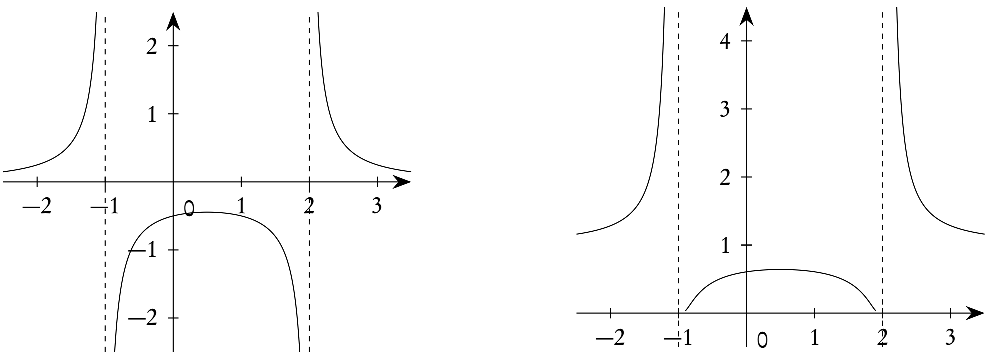
$f(x)=x^2-x-2$ $h(x)=\ln(f(x))$

Figura 17.14.: *Grafico di* $f(x)=x^2-x-2$ *e* $h(x)=\ln(f(x))$

17.5.3. L'esponenziale di una funzione

Poiché la funzione esponenziale (in base e) è strettamente crescente, data una funzione $f(x)$, la funzione $\mathrm{e}^{f(x)}$ sarà crescente o decrescente esattamente come $f(x)$. Naturalmente occorre ricordare che $\mathrm{e}^{f(x)}$ è sempre positiva. Per una trattazione significativa è opportuno anche fare qualche considerazione, seppure a livello intuitivo, su alcuni "comportamenti asintotici". In particolare ci interessano le seguenti osservazioni:

— nei pressi dei punti ove $f(x)$ tende a meno infinito, $\mathrm{e}^{f(x)}$ tende a zero;
— nei pressi dei punti dove $f(x)$ tende a più infinito, anche $\mathrm{e}^{f(x)}$ tende a più infinito.



$f(x)={}^1\!/_{x^2-x-2}$ $h(x)=\mathrm{e}^{f(x)}$

Figura 17.15.: *Grafico di* $f(x)={}^1\!/_{x^2-x-2}$ *e di* $\mathrm{e}^{f(x)}$

17.6. Funzioni lineari in seno e coseno

Sono le funzioni del tipo

$$f(x)=a\sin x+b\cos x+c. \tag{17.1}$$

Come già spiegato nel paragrafo 11.12.2 del capitolo 11, si può operare la seguente trasformazione. Poiché $a^2+b^2\neq 0$, si può raccogliere a^2+b^2, ottenendo:

$$f(x)=\sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{b}{\sqrt{a^2+b^2}}\cos x\right)+c.$$

Poiché

$$\left(\frac{a}{\sqrt{a^2+b^2}}\right)^2+\left(\frac{b}{\sqrt{a^2+b^2}}\right)^2=1,$$

esiste sicuramente α tale che

$$\cos\alpha=\frac{a}{\sqrt{a^2+b^2}}\quad\wedge\quad\sin\alpha=\frac{b}{\sqrt{a^2+b^2}}.$$

La funzione 17.1 può allora essere messa nella forma

$$f(x)=\sqrt{a^2+b^2}\big(\cos\alpha\sin x+\sin\alpha\cos x\big)+c\,,$$

ovvero

$$f(x)=\sqrt{a^2+b^2}\sin(x+\alpha)+c\,, \tag{17.2}$$

che può essere trattata utilizzando le tecniche già viste per dedurre il grafico di queste funzioni da quello della funzione seno.

Esempio 17.5. Per tracciare il grafico di $f(x)=\sin x+\sqrt{3}\cos x+1$, si riscrive la funzione nella forma

$$f(x)=2\sin\left(x+\frac{\pi}{3}\right)+1.$$

Il grafico si ottiene dunque da quello del seno con una traslazione verso sinistra di $\pi/3$, una dilatazione verticale di un fattore 2 e una successiva traslazione verso l'alto di una unità.

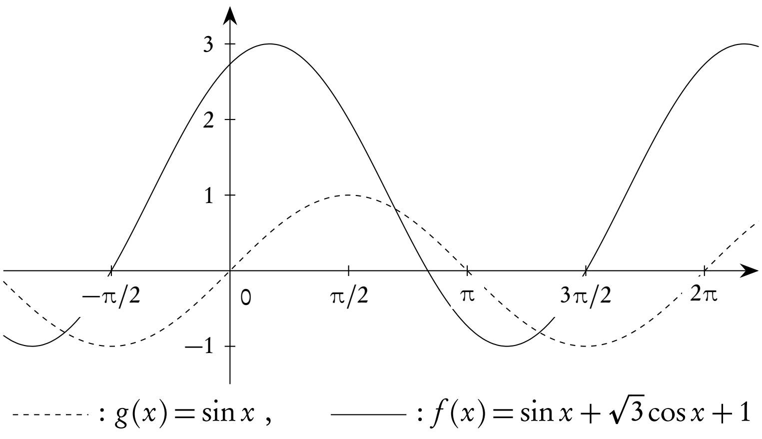


Figura 17.16.: *Grafico di una funzione lineare in seno e coseno*

17.7. Uso delle coniche

17.7.1. Funzioni razionali fratte e iperboli

Vogliamo considerare funzioni del tipo

$$f(x)=\frac{ax+b}{dx+e} \quad \text{e} \quad g(x)=\frac{ax^2+bc+c}{dx+e}, \quad d\neq 0, \tag{17.3}$$

nell'ipotesi che il numeratore non sia multiplo del denominatore[1]. La prima di queste funzioni è già stata trattata nel paragrafo 8.11.1 del capitolo 8: si tratta della cosiddetta funzione omografica.

Riscrivendo le (17.3) nella forma

$$y=\frac{bx+c}{dx+e} \quad \text{e} \quad y=\frac{ax^2+bx+c}{dx+e}, \tag{17.4}$$

riducendo allo stesso denominatore e portando tutto a primo membro, si verifica che si tratta di coniche (equazioni di 2° grado in due incognite). Poiché in entrambi i casi le rette verticali

$$x=-\frac{e}{d} \tag{17.5}$$

sono asintoti verticali (nei pressi di questo valore di x le due funzioni tendono all'infinito) le coniche sono sicuramente iperboli (uniche coniche con asintoti).

Conviene anche riscrivere le funzioni date eseguendo la divisione tra numeratore e denominatore; si ottiene

$$f(x)=\beta+\frac{\gamma}{dx+e} \quad \text{e} \quad g(x)=\alpha x+\beta+\frac{\gamma}{dx+e}, \tag{17.6}$$

con opportuni valori di α, β e γ. Se x tende all'infinito l'ultimo addendo in entrambi i casi tende a 0, dunque $f(x)$ "si avvicina" alla retta orizzontale $y=\beta$, $g(x)$ "si avvicina" alla retta obliqua $y=\alpha x+\beta$: si tratta, in entrambi i casi, del secondo asintoto dell'iperbole. Le bisettrici degli angoli individuati dai due asintoti sono gli assi dell'iperbole e le intersezioni di uno degli assi con l'iperbole forniscono le coordinate dei vertici: a questo punto il tracciamento del grafico è immediato.

Esempio 17.6. Rappresentare graficamente la funzione

$$f(x)=\frac{x-1}{2x+1}.$$

Eseguendo la divisione tra numeratore e denominatore si ottiene

$$f(x)=\frac{1}{2}+\frac{-3/2}{2x+1}.$$

I due asintoti sono dunque

$$x=-\frac{1}{2} \quad \text{e} \quad y=\frac{1}{2}.$$

[1] Nel caso in cui il numeratore sia multiplo del denominatore le (17.3) si riducono a una funzione costante o a una funzione lineare di primo grado, con l'esclusione del punto $x=-e/d$: si tratta di rette private di un punto.

Le bisettrici degli angoli individuati dagli asintoti si possono trovare facilmente e sono

$$y = -x \quad \text{e} \quad y = x + 1\,.$$

Le intersezioni della prima bisettrice con l'iperbole forniscono i due vertici dell'iperbole stessa:

$$\mathsf{V}_1 = \left(\frac{-1-\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2}\right) \quad \text{e} \quad \mathsf{V}_2 = \left(\frac{-1+\sqrt{3}}{2}, \frac{1-\sqrt{3}}{2}\right)$$

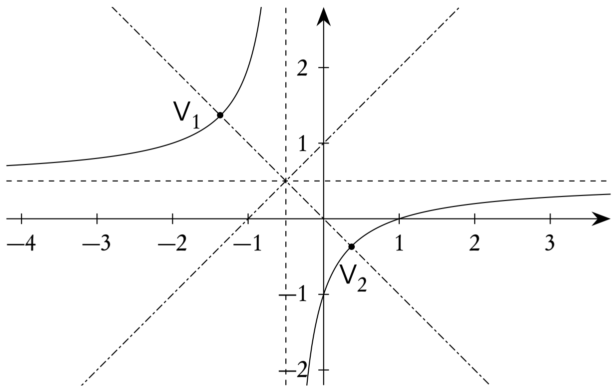


Figura 17.17.: *La funzione* $f(x) = (x-1)/(2x+1)$

Esempio 17.7. Rappresentare graficamente la funzione

$$f(x) = \frac{x^2 - x - 1}{x - 2}\,.$$

Eseguendo la divisione del numeratore per il denominatore si ottiene

$$f(x) = x + 1 + \frac{1}{x-2}\,.$$

I due asintoti sono dunque

$$x = 2 \quad \text{e} \quad y = x + 1, \quad \text{ovvero} \quad \mathsf{r}: x - 2 = 0 \quad \text{e} \quad \mathsf{s}: x - y + 1 = 0\,.$$

Per trovare le bisettrici degli angoli da esse individuati basterà imporre la condizione che un punto $\mathsf{P}(\alpha, \beta)$ sia equidistante da r e da s:

$$|\alpha - 2| = \frac{|\alpha - \beta + 1|}{\sqrt{2}} \quad \Rightarrow \quad \alpha - 2 = \pm\frac{\alpha - \beta + 1}{\sqrt{2}}\,.$$

Distinguendo i due casi e sostituendo α con x e β con y si ottengono le due rette perpendicolari

$$(1-\sqrt{2})x - y + 1 + 2\sqrt{2} = 0 \quad \text{e} \quad (1+\sqrt{2})x - y + 1 - 2\sqrt{2} = 0\,.$$

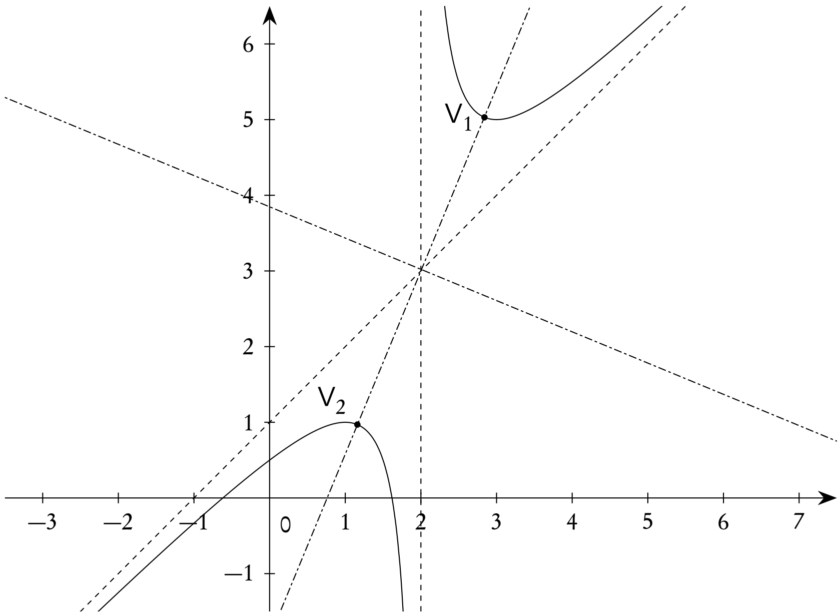


Figura 17.18.: *Grafico della funzione* $f(x)=(x^2-x-1)/(x-2)$

Le intersezioni della seconda bisettrice con l'iperbole forniscono i due vertici dell'iperbole stessa:

$$\mathsf{V}_1=\left(2+\frac{1}{\sqrt[4]{2}},3+\frac{1}{\sqrt[4]{2}}+\sqrt[4]{2}\right)\quad\text{e}\quad\mathsf{V}_2=\left(2+\frac{1}{\sqrt[4]{2}},3-\frac{1}{\sqrt[4]{2}}-\sqrt[4]{2}\right).$$

È interessante osservare che in una funzione come quella dell'esempio 17.7 è possibile determinare, sempre per via elementare, i massimi e minimi relativi. Per questo basta osservare che, detto m il minimo relativo e M il massimo relativo, le rette di equazione $y=m$ e $y=M$ sono tangenti all'iperbole. Per trovarle basterà considerare il sistema tra l'iperbole stessa e una generica retta orizzontale $y=k$, imponendo che l'equazione risolvente abbia una sola soluzione, ovvero abbia discriminante nullo.

$$\begin{cases} y=k \\ y=\frac{x^2-x-1}{x-2} \end{cases} \quad\Rightarrow\quad \begin{cases} y=k \\ x^2-x(k+1)+2k-1=0 \end{cases}.$$

Uguagliando a zero il discriminante si trova $k_1=1$, $k_2=5$. Il primo corrisponde al massimo relativo, il secondo al minimo relativo.

17.7.2. Funzioni irrazionali e coniche

Vogliamo considerare funzioni del tipo

$$f(x)=ax+b+c\sqrt{dx^2+ex+f},\quad \text{con } d \text{ ed } e \text{ non entrambi nulli,} \tag{17.7}$$

e con il trinomio di secondo grado sotto radice non sempre negativo.

Scrivendo la (17.7) nella forma

$$y = ax + b + c\sqrt{dx^2 + ex + f} \quad \text{ovvero} \quad y - ax - b = c\sqrt{dx^2 + ex + f}, \tag{17.8}$$

si verifica subito che deve essere

$$y - ax - b \geq 0 \quad \text{oppure} \quad y - ax - b \leq 0, \tag{17.9}$$

a seconda che c sia positivo o negativo.

Dopo aver scritto la condizione (17.9) si possono elevare al quadrato ambo i membri, ottenendo un'equazione di secondo grado in due incognite, che dunque rappresenta sempre una conica, di cui si dovrà considerare solo la parte che soddisfa la condizione posta. I calcoli possono essere complessi, specie se $a \neq 0$ nella (17.7), in quanto l'elevazione al quadrato fa comparire il termine misto: si tratta dunque di una conica con assi di simmetria non paralleli agli assi coordinati. In ogni caso, al fine di una ricerca delle proprietà geometriche del grafico ottenuto, è opportuno comunque sempre tenere presente che una funzione di questo tipo è parte di una conica.

Esempio 17.8. Vogliamo tracciare il grafico di

$$f(x) = \sqrt{x^2 - 1}.$$

Si ha

$$\begin{cases} y \geq 0 \\ y^2 = x^2 - 1 \end{cases} \Rightarrow \quad x^2 - y_2 = 1 \quad .$$

Si tratta dunque di una semiiperbole equilatera di centro l'origine e semiassi uguali ad 1, situata al di sopra dell'asse x.

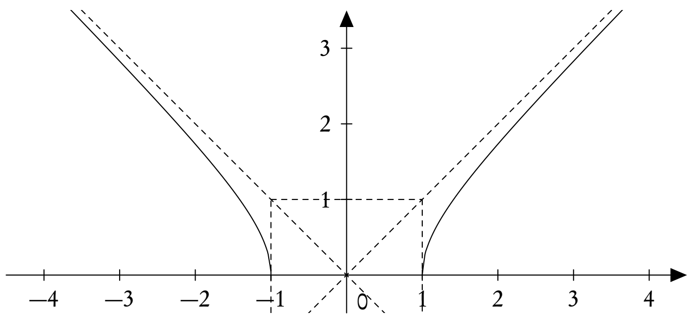


Figura 17.19.: *La funzione* $f(x) = \sqrt{x^2 - 1}$

Esempio 17.9. Vogliamo tracciare il grafico di

$$f(x) = \sqrt{|x^2 - 8x + 7|}$$

Procedendo come indicato si ha

$$\begin{cases} y \geq 0 \\ y^2 = |x^2 - 8x + 7| \end{cases} .$$

Considerando due casi, a seconda che l'argomento del valore assoluto sia positivo o negativo, si ottiene

$$\begin{cases} y \geq 0 \\ x \leq 1 \vee x \geq 7 \\ (x-4)^2 - y^2 = 9 \end{cases}, \quad \begin{cases} y \geq 0 \\ 1 < x < 7 \\ (x-4)^2 + y^2 = 9 \end{cases}.$$

Si tratta dunque dell'unione tra una semiiperbole equilatera di centro $(4,0)$ e semiassi uguali a 3 e una semicirconferenza con lo stesso centro e raggio uguale a 3, entrambe situate sopra l'asse delle ascisse.

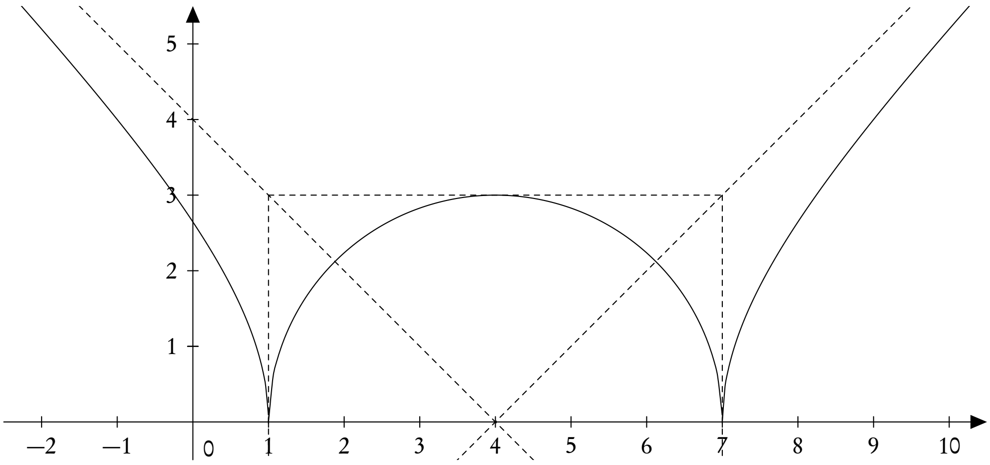


Figura 17.20.: *Grafico della funzione* $f(x) = \sqrt{|x^2 - 8x + 7|}$

A. Il syllabus dell'UMI

L'Unione Matematica Italiana ha pubblicato, nel 1999, un *Syllabus di Matematica* contenente un elenco delle conoscenze e capacità richieste per l'accesso all'università, in particolare alle facoltà scientifiche. Questo syllabus contiene anche una interessante raccolta di quesiti e problemi che invitiamo a risolvere.

Riportiamo un estratto dall'introduzione del syllabus, perché ci pare molto interessante e significativo.

> In genere, nell'insegnamento della matematica a livello universitario, molte nozioni vengono riprese dall'inizio e quindi potrebbero -in linea di principio- essere ignorate dagli studenti che si accingono ad entrare nell'Università. Tuttavia, la velocità e l'ampiezza con cui si sviluppano i corsi universitari di contenuto matematico sono tali che risulta difficile seguirli se le conoscenze elementari non sono già in parte bene assestate e se la mente non è già stata affinata ed allenata in modo assiduo ed intelligente. Per queste ragioni si è ritenuto opportuno elencare in questo Syllabus anche alcuni argomenti di base che risultano essere solitamente trattati (di nuovo e più approfonditamente) nei corsi universitari e la cui conoscenza non è quindi da considerare un prerequisito irrinunciabile.

Riportiamo di seguito gli argomenti elencati nel syllabus, limitatamente alla colonna dei saperi, raggruppati nei cinque temi previsti.

1. Strutture numeriche, aritmetica.
   — I numeri naturali: operazioni aritmetiche e loro proprietà.
   — La divisione con resto.
   — Numeri primi.
   — Massimo comun divisore e minimo comune multiplo.
   — Le frazioni numeriche: operazioni e ordinamento.
   — I numeri interi relativi. I numeri razionali relativi.
   — Rappresentazione dei numeri come allineamenti; allineamenti con virgola, finiti o periodici.
   — Idea intuitiva dei numeri reali.
   — Disuguaglianze e relative regole di calcolo.
   — Valore assoluto.
   — Potenze e radici.
   — Media aritmetica e media geometrica di due numeri positivi.
   — Logaritmi e loro proprietà.
   — Logaritmo decimale e sua relazione con la rappresentazione decimale dei numeri.
   — Basi numeriche.
2. Algebra elementare, equazioni, disequazioni.
   — Elementi di calcolo letterale, uso delle parentesi.
   — Polinomi.

- Prodotti notevoli. Potenza n-esima di un binomio.
- Divisione con resto nei polinomi. Regola di Ruffini. I polinomi come funzioni. Il principio di identità dei polinomi.
- Espressioni razionali fratte.
- Identità ed equazioni: nozione di soluzione.
- Equazioni algebriche di primo e secondo grado.
- Sistemi lineari di due equazioni in due incognite.
- Disequazioni.
- Disequazioni algebriche di primo e secondo grado.
- Disequazioni con espressioni fratte. Radicali, disequazioni con radicali.

3. Insiemi, elementi di logica, calcolo combinatorio, relazioni, funzioni.

- Linguaggio elementare degli insiemi, appartenenza, inclusione, intersezione, unione, complementazione, insieme vuoto.
- Coppie ordinate (prodotto cartesiano).
- Relazioni, funzioni o applicazioni. Relazioni di equivalenza e di ordine.
- Funzioni iniettive, suriettive, biiettive (corrispondenze biunivoche).
- Composizione di funzioni, funzione identica, funzione inversa di una funzione biiettiva.
- Permutazioni, disposizioni semplici e con ripetizione, combinazioni semplici.
- Connettivi logici: negazione, congiunzione, disgiunzione.
- Implicazione. Condizioni sufficienti, condizioni necessarie.
- Conoscere il significato dei termini: assioma, definizione, teorema, lemma, corollario, ipotesi, tesi. Dimostrazioni per assurdo.
- Quantificatori: $\forall$ (per ogni) e $\exists$ (esiste).

4. Geometria.

- Geometria euclidea piana: incidenza, ordinamento, parallelismo, congruenza (in alcuni testi: uguaglianza). Esistenza e unicità della parallela e della perpendicolare per un punto ad una retta assegnata.
- Lunghezza di un segmento (distanza tra due punti); corrispondenza biunivoca tra i punti di una retta e i numeri reali.
- Ampiezza degli angoli: misura in gradi. Lunghezza della circonferenza e misura degli angoli in radianti. Somma degli angoli interni di un triangolo. Relazioni tra gli angoli formati da due rette parallele tagliate da una trasversale.
- Criteri di equiscomponibilità dei poligoni e nozione elementare di area. Area del cerchio. Relazioni tra aree di figure simili.
- Misure e proporzionalità tra grandezze.
- Luoghi geometrici notevoli (asse di un segmento, bisettrice di un angolo, circonferenza, ecc.). Figure convesse.
- Proprietà delle figure piane: criteri di congruenza dei triangoli. Punti notevoli dei triangoli (baricentro, incentro, circocentro, ortocentro). Parallelogrammi. Teoremi di Talete, di Euclide, di Pitagora. Proprietà segmentarie e angolari del cerchio (corde, secanti, tangenti, arco sotteso da un angolo). Angoli al centro e alla circonferenza.

- Trasformazioni geometriche del piano: isometrie e similitudini. Simmetrie rispetto ad una retta e rispetto ad un punto, traslazioni, rotazioni, omotetie e loro composizioni.
- Coordinate cartesiane: equazioni di rette e circonferenze. Equazioni di semplici luoghi geometrici (parabole, ellissi, iperboli) in sistemi di riferimento opportuni.
- Trigonometria: seno, coseno, tangente di un angolo. Identità trigonometrica fondamentale $\sin^2 \alpha + \cos^2 \alpha = 1$. Formule di addizione.
- Geometria euclidea dello spazio: mutue posizioni di due rette, di due piani, di una retta e di un piano (angoli, parallelismo, perpendicolarità). Diedri e triedri.
- Sfera, cono, cilindro.
- Poliedri convessi, parallelepipedi, piramidi, prismi, poliedri regolari. Formula di Eulero.
- Idea intuitiva di volume dei solidi. Formule per il calcolo del volume e dell'area della superficie di parallelepipedo, piramide, prisma, cilindro, cono e sfera. Relazioni tra aree e tra volumi di solidi simili.
- Consapevolezza dell'esistenza di geometrie in cui sono negati alcuni assiomi della geometria euclidea classica (geometrie non-euclidee).

5. Successioni e funzioni numeriche.
   - Nozione di successione. Successioni definite assegnando il termine generale e successioni definite per ricorrenza.
   - Progressioni aritmetiche e geometriche.
   - Le funzioni numeriche e i loro grafici. Dominio di una funzione. Proprietà qualitative: crescenza, decrescenza, zeri, limitatezza, massimi e minimi relativi e assoluti.
   - Proprietà di alcune funzioni elementari: polinomi di primo e secondo grado, funzione potenza $x \mapsto x^{m/n}$; funzioni logaritmo ed esponenziale; funzioni trigonometriche e loro grafici. La funzione logaritmo come inversa dell'esponenziale. Periodicità delle funzioni trigonometriche.

Questo testo di Matematica di base tratta tutti questi argomenti e contiene qualche ulteriore approfondimento: l'idea è quella di fornire un manuale che possa essere utilizzato anche *dopo* aver seguito i corsi universitari, qualora si avvertisse la necessità di richiamare concetti e tecniche normalmente non presenti nei testi specialistici.

A proposito dell'assenza, nella Matematica di base, dei fondamenti del calcolo differenziale ed integrale, è opportuno segnalare, come fa il syllabus, che esso viene svolto in maniera approfondita e dettagliata sia negli insegnamenti di Analisi Matematica per i corsi di laurea in Matematica, Fisica, Ingegneria e Informatica, sia negli insegnamenti di Matematica, Istituzioni di Matematica e di Matematica Generale di altri corsi di laurea. Anche gli studenti che non hanno avuto occasione di acquisire nozioni di calcolo differenziale e integrale negli studi secondari possono dunque ragionevolmente pensare di iscriversi a uno dei corsi di laurea citati. Tuttavia conviene avvertire che gli strumenti del calcolo differenziale e integrale sono spesso usati fin dall'inizio in altri insegnamenti (Fisica, per esempio). Inoltre, soprattutto nei corsi di laurea dove l'insegnamento è impartito sulla base di semestri intensivi non è facile apprenderne in poche settimane sia i fondamenti teorici sia la manualità.

Notazioni utilizzate

Le notazioni utilizzate in questo testo sono quelle di default nel sistema tipografico $\LaTeX\,2_\varepsilon$, notazioni che, nella maggior parte dei casi, concordano con quelle previste dalla normativa UNI CEI ISO 80000-2 : 2010. Nella tabella che segue abbiamo riportato, per ragioni di completezza, anche alcuni simboli che non sono stati esplicitamente utilizzati in questo testo.

Segnaliamo inoltre che, nella numerazione dei teoremi, definizioni, osservazioni, ecc., abbiamo scelto di usare una numerazione progressiva per capitolo. Altri testi usano invece numerazioni progressive separatamente per i teoremi, le definizioni, ecc. Si tratta naturalmente solo di una questione di gusto personale.

La scrittura di un testo contenente molta matematica è sempre un'impresa ardua e che richiede molto tempo e fatica. Un aiuto indispensabile è fornito da un sistema di composizione come quello che abbiamo adottato (e che costituisce ormai lo standard de facto per i testi scientifici). Per chi fosse interessato a conoscere $\LaTeX\,2_\varepsilon$ segnaliamo che si tratta di un sistema di composizione tipografica di livello professionale e assolutamente gratuito. Tutte le informazioni utili si possono trovare sul sito ufficiale della comunità degli sviluppatori, `http://www.ctan.org/` e, in lingua italiana, sul sito degli Utilizzatori italiani di $\TeX$ e $\LaTeX$, `http://www.guitex.org`. Alcuni manuali introduttivi e consigli per iniziare si trovano anche sul sito personale dell'autore, `http://www.batmath.it`.

Osservazioni

Come già detto, le notazioni usate in questo testo sono quelle previste dalla normativa ufficiale UNI CEI ISO 80000-2 : 2010. Questa normativa in realtà è obbligatoria per chi usa la matematica per motivi tecnici: fisici, ingegneri, ecc. Non esiste obbligo alcuno per i "matematici puri". Riteniamo tuttavia opportuno adeguarsi, indipendentemente dagli obblighi ufficiali, per ragioni di uniformità: in tempi in cui i documenti possono circolare con estrema facilità e in cui i software di calcolo diventano sempre più diffusi, è opportuno evitare, per quanto possibile, ogni possibile ambiguità. Nella maggior parte dei casi le differenze tra le notazioni usate dai vari autori e quelle "ufficiali" sono solo di forma (caratteri corsivi oppure tondi, caratteri in grassetto o no, ecc.), e il lettore attento si adeguerà facilmente. Ci sono però alcune situazioni che meritano una speciale attenzione e le segnaliamo qui di seguito, anche se le abbiamo richiamate nella tabella che segue.

— Le notazioni sui logaritmi sono tutt'altro che omogenee. Oltre alle notazioni da noi proposte (al solito quelle "ufficiali") si usa anche $\log x$ per indicare il logaritmo in base "e", ancora $\log x$ per indicare il logaritmo in base 10 e $\operatorname{Log} x$ sempre per indicare il logaritmo in base 10. Si presti la massima attenzione e si consulti sempre l'elenco delle notazioni (quando disponibile!).

— La funzione $\operatorname{floor} x = \lfloor x \rfloor$ è spesso identificata con la funzione parte intera ed indicata con la notazione $[x]$. Anche i moderni software di calcolo simbolico (come *Mathematica*) definiscono invece la funzione parte intera secondo la normativa ISO, riportata nella tabella che segue.

— Discorso analogo per la funzione parte frazionaria, definita spesso come $\operatorname{frac}(x) = x - \operatorname{floor} x$: si noti che con questa definizione la funzione in oggetto diventa sempre non negativa, per ogni x.

— Alcuni autori (pochi per la verità) definiscono la funzione $\operatorname{sgn} x$ come $x/|x|$: la differenza con la definizione da noi data è che in questo modo la funzione non è definita in 0.
— Il prodotto vettoriale di due vettori è spesso indicato con $\vec{a} \wedge \vec{b}$, anziché con $\vec{a} \times \vec{b}$.
— Il prodotto scalare di due vettori è spesso indicato con $\vec{a} \times \vec{b}$ anziché con $\vec{a} \cdot \vec{b}$.

Elenco delle notazioni

"," - "."	Separatore decimale. Le regole prescrivono l'uso della virgola come separatore decimale nelle lingue diverse dall'inglese, dove invece si deve usare il punto. Tuttavia in questo testo abbiamo preferito fare uno strappo e usare il punto.
$p \wedge q$	p "et" q, congiunzione logica.
$p \vee q$	p "or" q, disgiunzione logica.
$\neg p$	"not" p, negazione.
$p \Rightarrow q$	p implica q.
$p \Leftrightarrow q$	p è equivalente a q.
$\forall$	Per ogni, quantificatore universale.
$\exists$	Esiste, quantificatore esistenziale.
$\exists!$, $\exists^1$	Esiste un solo.
$x \in A$, $A \ni x$	x appartiene ad A.
$x \notin A$, $A \not\ni x$	x non appartiene ad A.
$\{\, x \in A \mid \mathscr{P}(x) \,\}$	Insieme degli x di A per cui vale la proprietà $\mathscr{P}(x)$.
$\operatorname{card} A$, $\lvert A\rvert$	Cardinalità dell'insieme A.
$B \subseteq A$, $A \supseteq B$	B è un sottoinsieme di A oppure A è un soprainsieme di B; sono tollerate anche le scritture $B \subset A$ e $A \supset B$, ma in questo caso per i sottoinsiemi propri si deve usare $B \subsetneq A$ oppure $B \supsetneq A$.
$B \subset A$, $A \supset B$	B è un sottoinsieme proprio di A.
$A \cup B$	Unione di insiemi.
$A \cap B$	Intersezione di insiemi.
$A \setminus B$	Differenza di insiemi.
$A \triangle B$	Differenza simmetrica di due insiemi.
(a, b)	Coppia ordinata; se si usa la virgola come separatore decimale, e se a o b sono numeri con la virgola, va usato il ";" al posto della virgola come separatore della coppia.
$A \times B$	Prodotto cartesiano di insiemi.
$\complement_U A$	Complementare dell'insieme A rispetto all'insieme U.
$a \stackrel{\text{def}}{=} b$, $a := b$, $a =_{\text{def}} b$	a è uguale a b per definizione.
$a \propto b$	a è proporzionale a b.
$a \approx b$	a è circa uguale a b.
$a \ll b$	a è molto minore di b.
$a \gg b$	a è molto maggiore di b.

Continua nella pagina successiva

Segue dalla pagina precedente

$a \mid b$	a divide b (negli interi).
$\mathrm{MCD}(a,b)$	Massimo comun divisore di a e b.
$\mathrm{mcm}(a,b)$	Minimo comune multiplo di a e b.
$a \equiv b \mod k$	a è congruo a b modulo k.
$\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, $\mathbf{P}$	Insieme dei naturali (compreso lo zero), degli interi, dei razionali, dei reali, dei complessi, dei primi; si possono usare anche i simboli $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{P}$, e noi in questo testo abbiamo sempre usato questi; $\mathbf{N}^*$ oppure $\mathbb{N}^*$ indica i naturali senza lo zero (scritture analoghe per gli altri insiemi); $\mathbb{Z}^+$ indica gli interi maggiori o uguali a zero (scritture analoghe per gli altri insiemi); per indicare altre restrizioni si possono usare scritture del tipo $\mathbf{N}_{\geq 3}$, con ovvio significato.
$[a,b]$, $]a,b]$, $[a,b[$, $]a,b[$	Intervallo di reali chiuso, aperto a sinistra, aperto a destra, aperto; la normativa prevede anche i simboli $(a,b]$, $[a,b)$, (a,b) per gli intervalli aperti a sinistra, aperti a destra, aperti: ritengo che questi simboli vadano evitati, soprattutto l'ultimo per la confusione che può sorgere con il simbolo di coppia di reali.
$]-\infty,b]$, $]-\infty,b[$	Intervallo inferiormente illimitato chiuso, intervallo inferiormente illimitato aperto.
$[a,+\infty[$, $]a,+\infty[$	Intervallo superiormente illimitato chiuso, intervallo superiormente illimitato aperto.
$a \cdot b$, ab	Simboli usati per la moltiplicazione con operandi letterali.
3×5	Simbolo per la moltiplicazione con operandi numerici. Tuttavia a volte si usa anche il punto centrato come per gli operandi letterali.
$\mathsf{AB} \parallel \mathsf{CD}$, $\mathsf{r} \parallel \mathsf{s}$	La retta AB è parallela alla retta CD, la retta r è parallela alla retta s.
$\mathsf{AB} \perp \mathsf{CD}$, $\mathsf{r} \perp \mathsf{s}$	La retta AB è perpendicolare alla retta CD, la retta r è perpendicolare alla retta s.
$\overline{\mathsf{AB}}$	Segmento di estremi A e B.
$\overrightarrow{\mathsf{AB}}$	Vettore da A a B.
$d(\mathsf{A},\mathsf{B})$	Distanza tra A e B, lunghezza del segmento $\overline{\mathsf{AB}}$, modulo del vettore $\overrightarrow{\mathsf{AB}}$. Per la lunghezza del segmento $\overline{\mathsf{AB}}$ si usa anche la notazione $\lvert\overline{\mathsf{AB}}\rvert$, anche se non prevista nella normativa ISO.
$\lvert\overrightarrow{\mathsf{AB}}\rvert$, $\lVert\overrightarrow{\mathsf{AB}}\rVert$	Modulo o norma del vettore $\overrightarrow{\mathsf{AB}}$, anche in alternativa al simbolo $d(\mathsf{A},\mathsf{B})$.
$\overset{\frown}{\mathsf{AB}}$	Arco di estremi A e B.
$\mathsf{A}\hat{\mathsf{O}}\mathsf{B}$	Angolo di vertice O, individuato dalle semirette OA ed OB.
$\mathcal{A}$, $\mathcal{A}(\mathsf{ABC}\dots)$	Area, area della figura di vertici A, B, C, ...
$2p$	Perimetro di un poligono.
$n!$	Fattoriale di n.

Continua nella pagina successiva

Segue dalla pagina precedente

$\mathsf{C}_n^k = \binom{n}{k}$	Combinazioni (semplici) di n oggetti di classe k, ovvero coefficienti binomiali.
${}^{\mathsf{R}}\mathsf{C}_n^k$	Combinazioni con ripetizione di n oggetti di classe k.
V_n^k	Disposizioni ("variazioni") semplici di n oggetti di classe k.
${}^{\mathsf{R}}\mathsf{V}_n^k$	Disposizioni ("variazioni") con ripetizione di n oggetti di classe k.
P_n	Permutazioni di n oggetti distinti.
$\mathsf{P}_n^{n_1,n_2,\dots,n_k}$	Permutazioni di n oggetti, di cui n_1 uguali tra di loro, n_2 uguali tra di loro, $\dots n_k$ uguali tra di loro.
$\lvert a\rvert$, $\operatorname{abs} a$	Valore assoluto di a.
$\operatorname{sgn} a$	Segno del numero reale a, definito come segue: $\operatorname{sgn} a = -1$ per $a < 0$, $\operatorname{sgn} a = 0$ per $a = 0$, $\operatorname{sgn} a = 1$ per $a > 0$.
$\operatorname{ent} a$, $\lfloor a \rfloor$, $\operatorname{floor} a$	Il più grande intero minore o uguale al numero reale a, detta anche *funzione floor*; questa funzione è normalmente chiamata *funzione parte intera* nei testi di matematica ed è indicata con il simbolo $[a]$; secondo lo standard ISO (e secondo i software più diffusi), invece, la funzione parte intera è definita come nella linea seguente; in ogni caso riteniamo assolutamente da evitare il simbolo $[a]$.
$\operatorname{int} a$	Parte intera del numero reale a, definita come $\operatorname{int} a = \operatorname{sgn} a \cdot \lfloor \operatorname{abs} a \rfloor$: si veda la nota alla riga precedente.
$\operatorname{frac} a$	Parte frazionaria del numero reale a, definita come $\operatorname{frac} a = a - \operatorname{int} a$; questa definizione costituisce lo standard ISO (ed è implementata con questo nome dai software più diffusi), mentre nei testi di matematica è di solito definita come $a - \operatorname{floor} a$.
$\lceil a \rceil$, $\operatorname{ceil} a$	Il più piccolo intero maggiore o uguale al numero reale a, detta anche *funzione ceil* o *funzione ceiling*.
$\min(a, b)$	Minimo di a e b.
$\max(a, b)$	Massimo di a e b.
$\sin x$, $\cos x$	Le funzioni seno e coseno.
$\tan x$	La funzione tangente; evitare la scrittura $\operatorname{tg} x$.
$\cot x$	La funzione cotangente; evitare la scrittura $\operatorname{ctg} x$.
$\sec x$	La funzione secante.
$\csc x$, $\operatorname{cosec} x$	La funzione cosecante.
$\arcsin x$	La funzione arcoseno.
$\arccos x$	La funzione arcocoseno.
$\arctan x$	La funzione arcotangente; evitare la scrittura $\operatorname{arctg} x$.
$\operatorname{arccot} x$	La funzione arcocotangente; evitare la scrittura $\operatorname{arcctg} x$.
$\operatorname{arcsec} x$	La funzione arcsecante.
$\operatorname{arccsc} x$	La funzione arccosecante; evitare la scrittura $\operatorname{arccosec} x$.

Continua nella pagina successiva

Segue dalla pagina precedente

$\sinh x$	La funzione seno iperbolico.
$\cosh x$	La funzione coseno iperbolico.
$\tanh x$	La funzione tangente iperbolica.
$\operatorname{arsinh} x$	La funzione inversa del seno iperbolico.
$\operatorname{arcosh} x$	La funzione inversa del coseno iperbolico.
$\operatorname{artanh} x$	La funzione inversa della tangente iperbolica.
$f : A \to B$	Funzione di dominio A e codominio B (B non è l'insieme delle immagini).
$f : x \mapsto f(x)$	La funzione f manda $x \in A$ su $f(x) \in B$; $f(x)$ è un'espressione (di natura qualsiasi) che fornisce il valore della funzione f su x.
$g \circ f$	Composizione della funzione f con la funzione g.
Δf	Incremento finito della funzione f.
$\dfrac{\mathrm{d}f}{\mathrm{d}x}$, $\mathrm{d}f/\mathrm{d}x$, f'	Derivata della funzione f (per funzioni di una variabile); se la variabile è il tempo, si può usare $\dot{f}$ al posto di f'.
$\left(\dfrac{\mathrm{d}f}{\mathrm{d}x}\right)_{x=a}$, $f'(a)$	Derivata della funzione f calcolata nel punto a.
$\dfrac{\mathrm{d}^n f}{\mathrm{d}x^n}$, $f^{(n)}$	Derivata n-esima della funzione f.
$\int f(x)\,\mathrm{d}x$	Integrale indefinito della funzione f.
$\int_a^b f(x)\,\mathrm{d}x$	Integrale definito della funzione f da a a b.
$[f(x)]_a^b$, $f(x)\big\vert_a^b$	$f(b) - f(a)$.
e^x, $\exp x$	Esponenziale di x in base e.
a^x, $\exp_a x$	Esponenziale di x in base a.
$\log x$	Logaritmo di x, da usare quando non è necessario precisare la base; da notare che in molti testi (e spesso anche nelle calcolatrici e nei software) questa scrittura è usata per il logaritmo in base 10; purtroppo la stessa scrittura è usata anche in alcuni testi per il logaritmo naturale: è meglio attenersi alla norma ufficiale.
$\ln x$, $\log_{\mathrm{e}} x$	Logaritmo di x in base e.
$\lg x$, $\log_{10} x$	Logaritmo di x in base 10.
$\log_a x$	Logaritmo di x in base a.
$\operatorname{lb} x$, $\log_2 x$	Logaritmo binario (in base 2).
$\lvert z\rvert$, $\operatorname{abs} z$	Valore assoluto del numero complesso $z = x + \mathrm{i}y$, definito come $\sqrt{x^2 + y^2}$.
$\operatorname{Re} z$, $\operatorname{Im} z$	Parti reale e immaginaria di un numero complesso z: $z = \operatorname{Re} z + \mathrm{i} \operatorname{Im} z$.

Continua nella pagina successiva

Segue dalla pagina precedente

$\arg z$	Argomento del numero complesso z.
$\overline{z}$, z^*	Complesso coniugato di z: il primo è usato in matematica, il secondo in fisica e ingegneria.
$\operatorname{sgn} z$	Funzione segno del numero complesso z: $\operatorname{sgn} z = z/\lvert z\rvert$, $\operatorname{sgn} 0 = 0$.
$\boldsymbol{v}$, $\vec{v}$	Simboli per i vettori.
$\vec{a} \times \vec{b}$	Prodotto vettoriale di $\vec{a}$ e $\vec{b}$.
$\lvert\vec{v}\rvert$, $\lVert\vec{v}\rVert$	Modulo o norma del vettore $\vec{v}$.
$\vec{a} \cdot \vec{b}$	Prodotto scalare di $\vec{a}$ e $\vec{b}$.
$\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$	Base canonica di V_3.
$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$, $\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$	Scritture di una matrice .
$\det \boldsymbol{A}$, $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$	Determinante di una matrice quadrata.
$\boldsymbol{A}^{\mathsf{T}}$	Trasposta di una matrice.
$\operatorname{rank} \boldsymbol{A}$	Rango di una matrice.
$\boldsymbol{E}$, $\boldsymbol{I}$	Matrice unità.
$\operatorname{tr} \boldsymbol{A}$	Traccia di una matrice quadrata.

Alfabeto greco

Riportiamo, per l'importanza che ha nelle notazioni matematiche, l'alfabeto greco con la relativa pronuncia.

alfa	α	A	nu (ni)	ν	N
beta	β	B	csi	ξ	Ξ
gamma	γ	Γ	omicron	o	O
delta	δ	Δ	pi	π	Π
epsilon	ε	E	ro	ϱ	R
zeta	ζ	Z	sigma	σ	Σ
eta	η	H	tau	τ	T
theta	ϑ	Θ	upsilon	υ	Υ
iota	ι	I	fi	φ	Φ
cappa	$\varkappa$	K	chi	χ	X
lambda	λ	Λ	psi	ψ	Ψ
mu (mi)	μ	M	omega	ω	Ω

Come si può notare, in molti casi le lettere greche maiuscole non differiscono nella grafia dalle lettere dell'alfabeto latino.

Segnaliamo anche una lettera dell'alfabeto ebraico di uso comune in matematica.

aleph $\aleph$

Bibliografia

[1] G. Anichini, A. Carbone, P. Chiarelli e G. Conti. *Precorso di matematica*. Milano-Torino: Pearson Italia, 2010.

[2] L. Brusotti. «Poligoni e Poliedri». In: *Enciclopedia delle matematiche elementari*. Vol. II-I. Milano: Hoepli, 1937 (cit. a p. 308).

[3] H.S.M. Coxeter. *Projective Geometry*. New York: Springer, 2003 (cit. a p. 11).

[4] G. De Marco. *Analisi zero*. Padova: decibel editrice, 1981 (cit. a p. 62).

[5] K. Devlin. *Il gene della matematica*. Milano: Longanesi, 2002 (cit. a p. 37).

[6] Istituto di Matematica Applicata dell' Universitá di Padova, cur. *Richiami di matematica propedeutici agli studi di Ingegneria*. Non in vendita. Padova: La ciclografica Borghero, 1974.

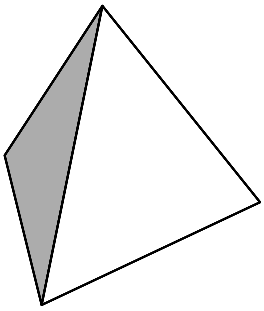
[7] A. Sabbatini. «Sui metodi elementari per la risoluzione dei problemi geometrici». In: *Questioni riguardanti le matematiche elementari*. Vol. II. Bologna: Zanichelli, 1924 (cit. a p. 327).

[8] M.E. Shanks, C.R. Fleenor e C.F. Brumfield. *Pre-Calculus Mathematics*. Reading, Massachussets: Addison-Wesley Publishing Company, 1981.

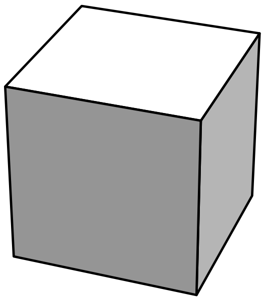
[9] Unione Matematica Italiana, cur. *Syllabus di matematica*. 1999. URL: `http://umi.dm.unibo.it/downloads/syllabus.pdf`.

[10] M. Verri e M. Bramanti. *Politest: il test di ingegneria al Politecnico di Milano*. Milano: Polipress, 2006.

Indice analitico



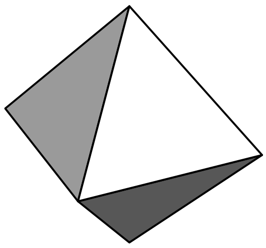
Matematica di Base
1 - Richiami teorici



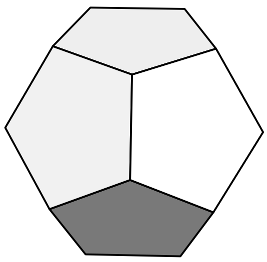
Luciano Battaia

http://www.batmath.it

Versione 1.1 del 21 giugno 2021



Questo testo è rivolto agli studenti che si accingono ad affrontare un corso universitario nel quale la Matematica sia un requisito essenziale del processo formativo. È ormai consuetudine che i candidati all'iscrizione ad uno di questi corsi di tipo tecnico-scientifico si cimentino in un test di matematica di base, contenente quesiti relativi agli argomenti matematici principali che sono stati oggetto del loro studio alle scuole secondarie superiori (calcolo differenziale ed integrale esclusi). In questo primo volume si sono voluti richiamare i concetti essenziali di questa *matematica elementare*, corredando i richiami teorici con una serie di esempi ed esercizi esplicativi.



Luciano Battaia
Già docente di matematica e fisica presso il Liceo Scientifico Grigoletti di Pordenone. Già titolare di corsi di insegnamento e di corsi di esercitazioni di Matematica di Base, Analisi Matematica, Istituzioni di Analisi, Matematica Generale, Matematica e Statistica, Matematica e Biomatematica, Meccanica Razionale, Fisica Matematica, presso le Università di Padova, Trieste e Udine.