DIDATTICA DELLA MATEMATICA I

Appunti del corso per la
scuola di specializzazione
per gli insegnanti della scuola secondaria
Padova 2005/2006

Enrico Gregorio

Dipartimento di Informatica — Sezione di Matematica
Università di Verona
Strada le Grazie 15 – Ca' Vignal, 37134 Verona (Italy)

Enrico.Gregorio@univr.it

Introduzione

Lo scopo di questo corso è di illustrare alcuni argomenti che non sempre fanno parte del programma di studio delle scuole superiori, o, se vengono trattati, lo sono in modo diverso.

Trovo che sia utile discutere questioni matematiche da punti di vista diversi e nuovi, con un occhio sempre alla possibilità di astrazione. Non sempre questa è desiderabile, soprattutto con studenti molto giovani; tuttavia è bene che essa sia introdotta un po' alla volta, per mostrare quanto potente sia il metodo che dal particolare va al generale, trovando sempre nuovi campi di applicazione.

La prima sezione tratta del metodo di induzione matematica: non occorre dire che si tratta di un metodo fondamentale nelle dimostrazioni di molti fatti e che le argomentazioni del tipo "e così via" nascondono una dimostrazione per induzione. È poi evidente che il metodo non va confuso con l'induzione di leggi fisiche o di altra natura dai dati sperimentali: questo è ciò che distingue la matematica dalle altre scienze. Certo, il matematico fa congetture, si lascia guidare dalle simmetrie e dalle analogie, ma non enuncia mai una legge basandosi su casi particolari.

La seconda sezione tratta della divisione con resto nei numeri naturali: l'argomento si presta, nello stesso modo, ad essere esteso alla divisione tra polinomi; l'esistenza del massimo comun divisore fra due polinomi (in una indeterminata) può essere dimostrata esattamente come si fa lì per i numeri naturali, tramite l'algoritmo euclideo. Occorre far notare come i polinomi *dovrebbero* essere introdotti con una sola indeterminata, solo per questi ha senso il teorema di Ruffini, ad esempio. Il fatto che anche per i polinomi in più di una indeterminata valga la decomposizione unica come prodotto di irriducibili è un risultato di difficile dimostrazione.

La sezione seguente è un esempio di applicazione dell'algebra astratta (in particolare della teoria degli anelli) ad argomenti di teoria dei numeri come le congruenze e la funzione di Eulero. Non sono argomenti difficili e possono essere trattati anche con metodi classici; l'approccio con l'algebra astratta, oltre che a impratichire lo studente nell'applicazione di regole di calcolo, pratica che può tornare utile nello studio di algoritmi per l'informatica, usa il minimo di strumenti e di calcoli, che rendono spesso oscuri i ragionamenti.

Un'altra possibile applicazione è la costruzione dei numeri reali mediante successioni di Cauchy. Questa costruzione renderebbe possibile definire in modo molto più veloce il concetto di funzione esponenziale. Tuttavia il metodo esposto nella sezione seguente è, a mio avviso, nettamente superiore.

Si introduce infatti il logaritmo (naturale) come primitiva della funzione $x \mapsto 1/x$, definita per $x > 0$. Le proprietà di questa nuova funzione sono dedotte dai teoremi del calcolo differenziale e integrale. Ciò fornisce, tra l'altro, applicazioni significative di questi risultati.

La sezione termina con una breve storia dei logaritmi che può essere usata per mostrare come i concetti vengono via via raffinati e le tecniche perfezionate: presentare la matematica come una scienza statica la rende noiosa e poco affascinante per gli studenti. Una delle domande più frequenti che vengono rivolte ai matematici è proprio questa: "Ma che cosa c'è ancora da scoprire?" E invece si tratta di una scienza viva, con problemi aperti e problemi ancora da inventare.

Le due sezioni seguenti introducono i concetti di serie e di serie di potenze, che non compaiono mai nei programmi liceali, anche se non sono argomenti di così difficile comprensione: l'unico teorema di difficile dimostrazione è quello di derivazione per serie. Lo sviluppo di una funzione in serie di potenze è, in realtà, una delle tecniche fondamentali in molte applicazioni e può servire di introduzione ad altri sviluppi, come quello in serie di Fourier, oltre che nel calcolo di limiti.

L'ultima sezione tratta delle funzioni trigonometriche in modo astratto; anche questa può essere usata per capire come l'astrazione ha molti vantaggi, al prezzo di una difficoltà iniziale: arrivare in cima a un monte è difficile, ma la vista che si può godere di lassù merita la fatica fatta.

1. Induzione

È indubbio che uno degli scogli su cui si incaglia il ragionamento è il concetto di *infinito*. In forma ancora primitiva, questo concetto è incontrato fin dal momento in cui ci si accorge che si può sempre *nominare* un numero più grande di qualunque numero dato, basta *aggiungere uno*.

L'osservazione che i numeri naturali sono infiniti non turbò i matematici per molti secoli; la matematica si accontentava di pensare sempre alle cose infinite come a enti *indefinitamente estendibili*; forse il primo teorema sulla infinità è quello dell'esistenza di infiniti numeri primi, che incontreremo più avanti. Il suo enunciato, in forma simile a quella datane da Euclide, è: *data una lista di numeri primi, possiamo esibire un altro numero primo che non è nella lista*.

In tempi più recenti, verso il diciassettesimo secolo, ci si rese conto che in molti casi è necessario *dimostrare* un enunciato che comprende infiniti casi. Come fare?

Un tipico esempio è: *dimostrare che ogni numero naturale maggiore di uno o è primo o può essere decomposto come prodotto di numeri primi*. L'intuizione ci dice che così è: ogni volta che abbiamo un numero naturale, abbiamo almeno una procedura che ci permette di decomporlo come prodotto di primi, il *crivello di Eratostene*.

Se il numero da decomporre è n, si scrivono tutti i numeri da 2 a n, si cancellano i multipli di 2, poi quelli di 3 che è il primo numero non cancellato, poi quelli di 5 e così via, cancellando sempre i multipli del successivo primo numero non già cancellato. Se durante questo procedimento si cancella n, abbiamo trovato un fattore primo di n, di fatto il più piccolo numero primo che divide n; altrimenti concludiamo che n è primo.

Una volta che sappiamo che n non è primo, possiamo eseguire la divisione, semplicemente contando i passi necessari per arrivare dal numero primo dato a n. Ne segue che possiamo ripetere la procedura con il quoziente.

Ovviamente questo non è il modo più economico per fattorizzare un numero naturale, ma è certamente un metodo che non richiede concetti superiori: per cancellare i multipli di m basta solo saltare m numeri nella lista, cioè basta solo saper contare.

Non è tanto il metodo che si usa, la cosa importante, ma l'idea fondamentale che sta sotto la dimostrazione: abbiamo ridotto il problema di fattorizzare n a quello di fattorizzare *un numero più piccolo*.

Questa è una forma dell'induzione matematica; fu usata da Fermat, che la chiamò *metodo di discesa*. Vediamo un'altra forma del metodo: indichiamo con s_n la somma dei numeri da 0 a n; vogliamo *dimostrare* che

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

Che $s_0 = 0 = 0(0+1)/2$ è ovvio. Ora compiamo il passo decisivo: supponiamo di sapere che la formula è valida per n e la dimostriamo per $n+1$. Infatti

$$s_{n+1} = s_n + n + 1 \stackrel{*}{=} \frac{n(n+1)}{2} + n + 1 = \frac{(n+1)(n+1+1)}{2}.$$

Abbiamo indicato con un asterisco il punto in cui si usa l'*ipotesi induttiva*, cioè l'ipotesi che la formula sia valida per il numero n.

Ora, il conto fatto è valido indipendentemente da n: siccome la formula è valida per $n = 0$, lo è per $n = 1$, quindi per $n = 2$ e così via. Dunque la formula è valida *per ogni n*.

Vedremo molti altri esempi di dimostrazioni per induzione, ma tutti si riducono a *ridurre la dimostrazione a un passo precedente*. Con un po' più di rigore, vediamo come si può enunciare il *principio di induzione*.

Prima forma: *Sia data una successione $A_0, A_1, A_2, \ldots$ di proposizioni che costituiscono, insieme, la proposizione A. Se sappiamo che A_0 è vera e che dalla verità di A_n segue la verità di A_{n+1}, allora la proposizione A è vera.*

Seconda forma: *Sia data una successione $A_0, A_1, A_2, \ldots$ di proposizioni che costituiscono, insieme, la proposizione A. Se sappiamo che A_0 è vera e che dalla verità di $A_1, A_2, \ldots, A_n$ segue la verità di A_{n+1}, allora la proposizione A è vera.*

La dimostrazione della validità di A_0 si chiama *base dell'induzione*; il resto della dimostrazione si chiama *passo induttivo*.

Il secondo esempio è un'applicazione della prima forma. Il primo esempio che abbiamo dato sfrutta invece l'induzione nella seconda forma; la proposizione A è: *Se n è un numero naturale, allora $n = 0$, oppure $n = 1$, oppure n è primo, oppure n è prodotto di primi.*

Non è difficile (esercizio) dimostrare l'equivalenza delle due forme del principio. Spesso è comodo usare la prima, talvolta è necessario servirsi della seconda.

Osserviamo inoltre che molto spesso occorre partire da una base di induzione diversa dal caso 0, ma questo è irrilevante: se vogliamo dimostrare la validità di tutte le proposizioni del tipo $A_k, A_{k+1}, \ldots$ (con $k > 0$), basta porre $B_r = A_{r+k}$ e fare induzione su r.

Esempio 1.1. Ci servirà in futuro questo conto: fissiamo un numero a e definiamo $s_0 = 0$ e, per $n \in \mathbf{N}$, $s_{n+1} = s_n + a^n$. Perciò, per $n \geq 1$,

$$s_n = 1 + a + a^2 + \cdots + a^{n-1}.$$

Verifichiamo, che per $a \neq 1$, abbiamo

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

La cosa è banale per $n = 0$. Supponiamola vera per n: allora

$$s_{n+1} = s_n + a^n = \frac{1 - a^n}{1 - a} + a^n = \frac{1 - a^n + a^n - a^{n+1}}{1 - a} = \frac{1 - a^{n+1}}{1 - a},$$

come si voleva. Un modo costruttivo di dimostrare la stessa formula è di accorgersi che

$$as_n = s_{n+1} - 1 = s_n + a^n - 1$$

e ricavare s_n.

L'esempio di prima mostra che è spesso facile dimostrare per induzione una formula data; molto meno facile è spesso *trovare* la formula giusta.

Un bambino di terza elementare mi domandò una volta: "Infinito è pari o dispari?"

Evidentemente gli avevano detto che oltre tutti i numeri esiste questo "infinito". È facile rispondere: "Infinito non è un numero", ma credo sia più istruttivo spiegare perché non lo è. Gli chiesi di dire come si fa a sapere che una certa quantità di oggetti è pari o dispari: si toglie una coppia di oggetti alla volta dal mucchio finché si può. Se non ne avanzano gli oggetti erano un numero pari, se ne avanza uno, gli oggetti erano un numero dispari.

Se eseguiamo la stessa operazione su un mucchio infinito di oggetti, non arriviamo mai alla fine, quindi l'attributo "pari o dispari" non si applica all'infinito.

2. La divisione con resto e l'algoritmo di Euclide

Sappiamo tutti che 44 gatti in fila per sei danno un resto di due. In generale, data una quantità di oggetti da dividere fra più persone, sappiamo come fare per eseguire l'operazione e sappiamo che potrà avanzare qualcosa.

Ovviamente il fatto che questa divisione è sempre possibile e dà un risultato univoco va dimostrato. Nel seguito faremo uso delle proprietà elementari delle operazioni fra numeri naturali senza menzionarle; può essere un utile esercizio quello di esplicitare le proprietà usate.

Teorema 2.1. *Dati due numeri naturali a e b, con $b \neq 0$, esistono e sono unici due numeri naturali q e r tali che*

(1) $a = bq + r$,
(2) $r < b$.

Dimostrazione. (Unicità) Dobbiamo verificare che, se $a = bq + r$ e $a = bq' + r'$, con $r < b$ e $r' < b$, allora $q = q'$ e $r = r'$.

Se $r = r'$, allora, da $bq + r = bq' + r'$ segue $bq = bq'$ e, essendo $b \neq 0$, otteniamo $q = q'$. Dunque basta dimostrare che non può essere $r \neq r'$.

Se $r' < r$, possiamo scrivere $r - r' = b(q' - q)$. Ma il numero a sinistra è certamente minore di b e maggiore di zero. Perciò anche quello a destra è diverso da zero; ma, in tal caso $q' - q \geq 1$ e quindi $b(q' - q) \geq b$: questo è assurdo.

Se $r' > r$, la dimostrazione è analoga.

(Esistenza) Useremo l'induzione su a nella seconda forma: l'idea è quella di "togliere b oggetti dal mucchio" e quindi ridursi a un caso precedente.

Se $a = 0$ l'enunciato è certamente vero: $0 = b0 + 0$.

Supponiamo allora $a > 0$ e che l'enunciato sia vero per ogni numero minore di a.

Se $a < b$, non c'è niente da fare: $a = b0 + a$ e $a < b$ per ipotesi.

Se invece $a \geq b$, sappiamo che $a - b < a$. Quindi, per ipotesi induttiva, esistono q e r tali che

(1) $a - b = bq + r$,
(2) $r < b$.

Ma allora $a = b + bq + r = b(q + 1) + r$ e abbiamo concluso. □

Definizione 2.2. Se $a = bq + r$, con $r < b$, diciamo che q è il *quoziente* e r è il *resto* della divisione di a per b.

Osservazione 2.3. Nella dimostrazione precedente non è possibile usare direttamente l'induzione nella prima forma: il procedimento delineato è proprio quello di togliere b oggetti dal mucchio, poi altri b, e così via. È chiaro a tutti che non si può ridurre il problema della divisione di $n + 1$ per b a quello della divisione di n per b.

La seconda condizione, quella che il resto sia minore del divisore, è ovviamente importantissima, per garantire il risultato unico della divisione.

La prima applicazione che possiamo dare di questo è proprio il teorema sull'infinità dei numeri primi; la dimostrazione è costruttiva, ed assomiglia molto a quella data da Euclide, che è la prima comparsa in un testo, i famosi "Elementi della geometria".

Definizione 2.4. Diciamo che a è un *multiplo* di b se esiste un naturale q tale che $a = bq$; tale q è unico per il teorema precedente. Diremo anche che b *divide* a o che b è un divisore di a, e useremo la notazione $b \mid a$. La negazione di questo si scrive $b \nmid a$.

Un numero naturale p è *primo* se $p > 1$ e, da $b \mid p$, segue $b = 1$ oppure $b = p$. Perciò un numero maggiore di 1 è primo se ha, come unici divisori, 1 e p stesso.

Notiamo che 0 e 1 non sono numeri primi; questa scelta è dovuta al desiderio di enunciare precisamente il teorema sulla fattorizzazione unica in prodotto di primi.

Lemma 2.5. *Sia n un numero naturale maggiore di 1. Allora esiste un numero primo p che divide n.*

Dimostrazione. Induzione su n, con base $n = 2$, nella seconda forma.

Certamente 2 è un numero primo e $2 \mid 2$.

Supponiamo ora che il risultato sia vero per ogni numero minore di n, con $n > 2$. Se n è primo, abbiamo finito, perché $n \mid n$. Altrimenti $n = km$, con $1 < m < n$. Ma allora, per ipotesi induttiva, esiste un numero primo p tale che $p \mid m$. Quindi $p \mid n$. □

Teorema 2.6. *Esistono infiniti numeri primi.*

Dimostrazione. Supponiamo di avere una lista di numeri primi distinti, $p_1, p_2, \ldots, p_r$. Poniamo $n = p_1 p_2 \ldots p_r + 1$. Per il lemma, esiste un primo p che divide n. Ma questo primo p non può comparire nella lista, perché il resto della divisione di n per p_i è 1 ($i = 1, 2, \ldots, n$).

Perciò nessuna lista finita di numeri primi può elencarli tutti. □

Dicevamo che la dimostrazione è costruttiva: infatti basta partire da $p_1 = 2$. Allora:

- $p_1 + 1 = 3$ è primo, sia p_2;
- $p_1 p_2 + 1 = 7$ è primo, sia p_3;
- $p_1 p_2 p_3 + 1 = 43$ è primo, sia p_4;
- $p_1 p_2 p_3 p_4 + 1 = 1807$ è primo, sia p_5;
- $p_1 p_2 p_3 p_4 p_5 + 1 = 3\,263\,443$ è primo, sia p_6;
- $p_1 p_2 p_3 p_4 p_5 p_6 + 1 = 10\,650\,056\,950\,807$ non è primo, ma si fattorizza come

$$10\,650\,056\,950\,807 = 547 \cdot 607 \cdot 1033 \cdot 31051.$$

Si può proseguire ponendo uno dei primi trovati all'ultimo passo, ad esempio $p_7 = 547$.

Notiamo che questa procedura trova numeri primi, ma certamente non tutti: non è noto alcun algoritmo capace di generare tutti i numeri primi, e credo non ne possa esistere alcuno.

Esistono numerosissime congetture sui numeri primi. La più celebre è quella di Goldbach: *ogni numero pari è somma di due numeri primi.* Esistono risultati parziali, ma non è ancora stata né provata né dimostrata falsa. Nel 1931 fu data la dimostrazione, da Shnirelman, che ogni numero naturale è somma di al più 300 000 primi. Vinogradov, da parte sua, dimostrò successivamente che ogni numero naturale "sufficientemente grande" è somma di al più quattro numeri primi. L'enunciato dice, più precisamente: *esiste un numero naturale N tale che ogni numero naturale $n > N$ si può scrivere come somma di al più quattro numeri primi.*

La differenza tra i due risultati è evidente: il primo dà una proprietà vera per ogni numero naturale; il secondo dice essenzialmente che è assurdo supporre che esistano infiniti numeri naturali che non si possono esprimere come somma di al più quattro numeri primi.

Famosissimo è il cosiddetto "Ultimo teorema di Fermat": *se $n > 2$, non esistono numeri naturali x, y e z diversi da zero tali che $x^n + y^n = z^n$.*

L'enunciato originale fu appuntato da Fermat a margine della sua copia delle "Coniche" di Apollonio, con il commento: "Ho trovato una dimostrazione mirabile di questo teorema, ma il margine è troppo piccolo per contenerla". La dimostrazione è stata data più di trecento anni dopo, da Andrew Wiles, nel 1997.

Citiamo questo parlando di primi, perché è facile ridurre la questione a esponenti primi (esercizio: la riduzione, non il teorema!).

Un'altra congettura, dimostrata falsa da Eulero è ancora di Fermat: *i numeri della forma* $F_n = 2^{2^n}+1$ *sono primi.* Di fatto si ha:

$$
\begin{aligned}
F_0 &= 2^{2^0}+1 = 3\\
F_1 &= 2^{2^1}+1 = 5\\
F_2 &= 2^{2^2}+1 = 17\\
F_3 &= 2^{2^3}+1 = 257\\
F_4 &= 2^{2^4}+1 = 65\,537\\
F_5 &= 2^{2^5}+1 = 4\,294\,967\,297 = 641\cdot 6\,700\,417\\
F_6 &= 2^{2^6}+1 = 18\,446\,744\,073\,709\,551\,617 = 274\,177\cdot 67\,280\,421\,310\,721\\
F_7 &= 2^{2^7}+1 = 340\,282\,366\,920\,938\,463\,463\,374\,607\,431\,768\,211\,457\\
&\qquad = 59\,649\,589\,127\,497\,217\cdot 5\,704\,689\,200\,685\,129\,054\,721
\end{aligned}
$$

e non si conoscono altri numeri della forma F_n che siano effettivamente primi. Si noti come i numeri di Fermat F_n crescano rapidamente, ciò che rende molto arduo fattorizzarli. Il calcolo della fattorizzazione di F_7 ha richiesto un paio di minuti di lavoro ad un calcolatore piuttosto veloce, utilizzando il programma *Mathematica*™.

Possiamo estendere il concetto di divisione con resto anche agli interi.

Teorema 2.7. *Siano a e b numeri interi, con $b \neq 0$. Allora esistono e sono unici due numeri interi q e r tali che:*

(1) $a = bq + r$;
(2) $0 \leq r < b$.

Dimostrazione. L'unicità si sistema esattamente come nel teorema sui naturali. Rimane l'esistenza, che divideremo in quattro casi.

Caso $a \geq 0$ e $b > 0$: non c'è nulla di nuovo da dimostrare.

Caso $a \geq 0$ e $b < 0$: sappiamo che $a = (-b)q + r$, con $0 \leq r < b$. Allora $a = b(-q) + r$.

Caso $a < 0$ e $b > 0$: abbiamo $-a = bq + r$, con $0 \leq r < b$. Se $r = 0$, possiamo scrivere $a = b(-q) + 0$ ed abbiamo finito. Se $r > 0$, invece, scriviamo $a = b(-q) - r = b(-q) - b + b - r = b(-q-1) + (b-r)$ e $0 \leq b - r < b$.

Caso $a < 0$ e $b < 0$: abbiamo $a = (-b)q + r = b(-q) + r$, con $0 \leq r < b$, per il caso precedente. □

Non molti sanno che il manuale scientifico più famoso della storia, gli "Elementi" di Euclide, non contiene solo lo sviluppo della geometria piana e solida, ma anche molte parti di aritmetica, fra cui tutta la teoria delle proporzioni. Vi sono trattate anche questioni che ora andrebbero sotto il titolo di teoria dei numeri: ne abbiamo già visto un esempio a proposito dei numeri primi.

Di fondamentale importanza nella teoria elementare della misura è il concetto di massimo comun divisore e la trattazione di Euclide comprende un metodo per il calcolo.

Definizione 2.8. Dati due numeri naturali a e b, diremo che d è il *massimo comun divisore* di a e b quando:
(1) $d \mid a$ e $d \mid b$;
(2) se $c \mid a$ e $c \mid b$, allora $c \mid d$.

Osservazione 2.9. Notiamo che la definizione usa non tanto l'ordinamento usuale dei numeri naturali, quanto l'ordinamento per divisibilità. Questo spiega il motivo dell'importanza delle coppie di naturali il cui massimo comun divisore è 1: infatti 1 è il minimo dei naturali ordinati per divisibilità.

Questa definizione, inoltre, ammette una facile estensione all'ambito dei polinomi, dove non è possibile darne una che usi altri tipi di ordinamento.

È facile dimostrare che il massimo comun divisore, *se esiste*, è unico (esercizio). Meno facile è provarne l'esistenza.

Lemma 2.10. *Siano a e b numeri naturali, con $a \geq b$. Il massimo comun divisore fra a e b esiste se e solo se esiste il massimo comun divisore fra $a-b$ e b. In tal caso i due massimi comuni divisori sono uguali.*

Dimostrazione. ($\Rightarrow$) Sia d il massimo comun divisore fra a e b; allora $a = da'$ e $b = db'$. Quindi $a - b = d(a' - b')$ e quindi d soddisfa la prima proprietà per la coppia $a-b$, b.

Supponiamo ora che $c \mid (a-b)$ e $c \mid b$. Allora $a - b = cx$ e $b = cy$; perciò $a = cx + b = c(x+y)$, da cui $c \mid a$. Poiché d soddisfa la seconda condizione per la coppia a, b, abbiamo che $c \mid d$, cioè che d soddisfa la seconda condizione per la coppia $a-b$, b, come richiesto.

($\Leftarrow$) Analoga. □

Questo lemma è il passo fondamentale per l'algoritmo di Euclide per il calcolo del massimo comun divisore, oltre che per la dimostrazione dell'esistenza. Il fatto che si possa "togliere il minore dei due numeri dal maggiore", insieme alla procedura già usata per la divisione con resto, indica la validità dell'enunciato seguente.

Corollario 2.11. *Siano a e b numeri naturali non nulli e sia $a = bq + r$, con $r < b$. Il massimo comun divisore fra a e b esiste se e solo se esiste il massimo comun divisore fra b e r. In tal caso i due massimi comuni divisori sono uguali.*

Dimostrazione. Identica alla precedente (esercizio). □

Ora, se abbiamo i due numeri a e b (non nulli, altrimenti il calcolo del massimo comun divisore è banale) possiamo procedere così: ogni passo è la divisione con resto. Non è restrittivo supporre $a \geq b$; poniamo $r_1 = b$.
(1) Eseguiamo $a = r_1 q_1 + r_2$; se $r_2 = 0$ interrompiamo.
(2) Eseguiamo $r_1 = r_2 q_2 + r_3$; se $r_3 = 0$ interrompiamo.
(3) Eseguiamo $r_2 = r_3 q_3 + r_4$; se $r_4 = 0$ interrompiamo.
(4) ...
(n) Eseguiamo $r_{n-1} = r_n q_n + r_{n+1}$; se $r_{n+1} = 0$ interrompiamo.

Sappiamo che a un certo punto dovremo interrompere il procedimento: infatti, per definizione di resto della divisione, abbiamo

$$r_0 > r_1 > r_2 > \cdots > r_{n-1} > r_n$$

e quindi, al più in $b+1$ passi, dovremo trovare uno dei resti successivi che è zero. Ma questo significa che possiamo supporre che $r_{n+1} = 0$, quindi che il massimo comun divisore fra r_n e r_{n+1} esiste ed è r_n.

Se applichiamo il corollario n volte, abbiamo che r_n è il massimo comun divisore fra a e b. (Esercizio: rendere preciso il ragionamento con una dimostrazione per induzione.)

Teorema 2.12. *Dati due numeri naturali a e b, il massimo comun divisore fra a e b esiste e si denota con* $\mathrm{mcd}(a,b)$.

Ovviamente questo non è l'unico metodo per il calcolo del massimo comun divisore, ma è certamente il più economico per numeri grandi, di cui è complicato calcolare la fattorizzazione in primi. Inoltre l'algoritmo fornisce anche un importantissimo risultato, noto come teorema di Bézout.

Esempio 2.13. Calcoliamo il massimo comun divisore fra 18847957 e 565051.

(1) $18847957 = 565051 \cdot 33 + 201274$;
(2) $565051 = 201274 \cdot 2 + 162503$;
(3) $201274 = 162503 \cdot 1 + 38771$;
(4) $162503 = 38771 \cdot 4 + 7419$;
(5) $38771 = 7419 \cdot 5 + 1676$;
(6) $7419 = 1676 \cdot 4 + 715$;
(7) $1676 = 715 \cdot 2 + 246$;
(8) $715 = 246 \cdot 2 + 223$;
(9) $246 = 223 \cdot 1 + 23$;
(10) $223 = 23 \cdot 9 + 16$;
(11) $23 = 16 \cdot 1 + 7$;
(12) $16 = 7 \cdot 2 + 2$;
(13) $7 = 2 \cdot 3 + 1$;
(14) $2 = 1 \cdot 2 + 0$.

Dunque $\mathrm{mcd}(18847957, 565051) = 1$. Sfido chiunque ad eseguire il calcolo con una semplice calcolatrice a quattro operazioni, con il metodo della fattorizzazione in primi, in meno di venti minuti, il tempo da me impiegato anche per scrivere tutti i passaggi.

Teorema 2.14 (Bézout)**.** *Siano a e b numeri naturali e sia $d = \mathrm{mcd}(a,b)$. Allora esistono due numeri interi x e y tali che*

$$d = ax + by.$$

Dimostrazione. Consideriamo i passi usati per il calcolo di d con l'algoritmo di Euclide, supponendo che $d = r_n$ sia l'ultimo resto non nullo.

Per definizione abbiamo $r_1 = a \cdot 0 + b \cdot 1 = ax_1 + by_1$. Dal primo passo otteniamo

$$r_2 = a - r_1 q_1 = a - (ax_1 + by_1)q_1 = a(1 + x_1 q_1) + b(-y_1 q_1) = ax_2 + by_2.$$

Dal secondo passo otteniamo

$$r_3 = r_1 - r_2 q_2 = ax_1 + by_1 - (ax_2 + by_2)q_2 = a(x_1 - x_2 q_2) + b(y_1 - y_2 q_2) = ax_3 + by_3.$$

Ovviamente questo si può ripetere, fino all'ultimo passo, trovando quindi $r_n = ax_n + by_n$. □

Questo teorema fornisce, ad esempio, una facile dimostrazione di una proprietà caratteristica dei numeri primi. Attenzione: questa dimostrazione usa i numeri interi (anche negativi)!

Teorema 2.15. *Sia p un numero naturale, $p > 1$; p è primo se e solo, per ogni coppia di numeri naturali a e b, da $p \mid ab$ segue $p \mid a$ oppure $p \mid b$.*

Dimostrazione. ($\Rightarrow$) Supponiamo che p sia primo, che $p \mid ab$ e che $p \nmid a$. Allora $\mathrm{mcd}(a,p) = 1$, perché questo massimo comun divisore può essere solo 1 oppure p, dal momento che p è primo. Non può però essere p, perché $p \nmid a$.

Per il teorema di Bézout, possiamo scrivere $1 = ax + py$, per opportuni numeri *interi* x e y. Allora

$$b = b \cdot 1 = abx + pby.$$

Essendo $p \mid ab$, otteniamo che $p \mid b$.

($\Leftarrow$) Supponiamo che p abbia la proprietà suddetta e scriviamo $p = ab$, con $0 < a < p$. Siccome $p \nmid a$, deve essere $p \mid b$, quindi $b = pc$. Ne segue $p = ab = apc$, cioè $0 = p(1 - ac)$. Ma allora $ac = 1$ e quindi $a = 1$. □

Corollario 2.16. *Sia p un numero primo e supponiamo che $p \mid a_1 a_2 \dots a_r$. Allora $p \mid a_i$, per un opportuno i, $1 \leq i \leq r$.*

Dimostrazione. Esercizio (induzione su r). □

Usando questa proprietà possiamo dimostrare l'unicità della fattorizzazione in primi dei numeri naturali maggiori di uno.

Definizione 2.17. Diciamo che una lista di primi $\langle p_1, p_2, \dots, p_r \rangle$ (non necessariamente distinti) è una *fattorizzazione in primi* di $n > 1$ se

$$p_1 \leq p_2 \leq \cdots \leq p_r \quad \text{e} \quad p_1 p_2 \dots p_r = n.$$

Comprendiamo quindi anche il caso $r = 1$, dove il "prodotto" è p_1. Diremo che r è la *lunghezza* della fattorizzazione.

Due fattorizzazioni $\langle p_1, p_2, \dots, p_r \rangle$ e $\langle q_1, q_2, \dots, q_s \rangle$ sono uguali se $r = s$ e $p_1 = q_1$, $p_2 = q_2$, …, $p_r = q_r$.

Possiamo allora enunciare il teorema fondamentale dell'aritmetica.

Teorema 2.18. *Sia $n > 1$ un numero naturale; allora esiste una fattorizzazione in primi di n e due fattorizzazioni in primi di n sono uguali.*

Dimostrazione. L'esistenza è già stata dimostrata (completare i dettagli).

Supponiamo ora che $\langle p_1, p_2, \dots, p_r \rangle$ e $\langle q_1, q_2, \dots, q_s \rangle$ siano due fattorizzazioni in primi di n.

Dimostriamo per induzione su r.

Se $r = 1$, $n = p_1$ è primo. Se $s = 1$ abbiamo finito; supponiamo $s > 1$, da cui $p_1 \mid q_1 q_2 \dots q_s$. Il corollario 2.15 dice allora che p_1 divide q_i per un opportuno i, $1 \leq i \leq s$. Ma q_i è primo, quindi $p_1 = q_i$: assurdo.

Supponiamo di avere l'asserto per fattorizzazioni di lunghezza $r - 1$ e dimostriamolo per fattorizzazioni di lunghezza r ($r > 0$). Abbiamo

$$n = p_1 p_2 \dots p_{r-1} p_r = q_1 q_2 \dots q_s,$$

quindi $p_r \mid q_1 q_2 \dots q_s$. Di nuovo, $p_r \mid q_i$, per un certo i, e $p_r = q_i$. Ora abbiamo $p_r = q_i \leq q_s$. In modo del tutto analogo abbiamo $q_s \leq p_r$, quindi $p_r = q_s$.

Allora

$$n' = p_1 p_2 \dots p_{r-1} = q_1 \dots q_{s-1}$$

è un numero che ammette una fattorizzazione di lunghezza $r - 1$. Per ipotesi induttiva, le due fattorizzazioni sono uguali, cioè $r - 1 = s - 1$ e $p_1 = q_1$, $p_2 = q_2$, …, $p_{r-1} = q_{r-1}$. □

Avendo il teorema fondamentale dell'aritmetica, è facile ricavare l'altro metodo per il calcolo del massimo comun divisore, avendo a disposizione la fattorizzazione dei numeri: *si prendono i divisori primi comuni, una sola volta e con il minimo esponente.*

3. Anelli e campi

In questa sezione daremo qualche applicazione dell'algebra astratta alla teoria dei numeri. In particolare useremo alcuni elementi della teoria degli anelli per dimostrare il "piccolo teorema di Fermat" e per discutere la funzione di Eulero.

Definizione 3.1. Un *anello* è un insieme A dotato di due operazioni, addizione e moltiplicazione, denotate rispettivamente con + e $\cdot$ (ma di solito il puntino viene sottinteso) tali che:

(A1) + è associativa, cioè $a+(b+c)=(a+b)+c$, per ogni $a,b,c\in A$;
(A2) + ammette un elemento neutro 0, tale che $a+0=a=0+a$, per ogni $a\in A$;
(A3) ogni elemento $a\in A$ ammette opposto rispetto a +, cioè un elemento b tale che $a+b=0=b+a$;
(A4) $\cdot$ è associativa, cioè $a(bc)=(ab)c$, per ogni $a,b,c\in A$;
(A5) $\cdot$ ammette un elemento neutro 1, tale che $a1=a=1a$, per ogni $a\in A$;
(A6) $\cdot$ è distributiva rispetto a +, cioè

$$a(b+c)=ab+ac, \qquad (a+b)c=ac+bc,$$

per ogni $a,b,c\in A$.

L'anello A si dice:

- *commutativo* se $ab=ba$, per ogni $a,b\in A$;
- un *dominio* se è commutativo, $0\neq 1$ e, per ogni $a,b\in A$, da $ab=0$ segue $a=0$ oppure $b=0$;
- un *campo* se è commutativo, $0\neq 1$ e, per ogni $a\in A$, $a\neq 0$, esiste $b\in A$ con $ab=1$ (cioè ogni elemento diverso da 0 ammette inverso rispetto alla moltiplicazione).

Un elemento a dell'anello A si dice *invertibile* se esiste un elemento b tale che $ab=1=ba$.

La notazione è quella usuale dei numeri, con le stesse regole di precedenza per le operazioni. Ad esempio, $ab+ac$ significa fare ab, poi ac e sommare il risultato. Viceversa, $a(b+c)$ significa eseguire il prodotto fra a e la somma $b+c$. Non deve far temere il fatto che si usi la terminologia usuale, perché le regole di calcolo sono esattamente le stesse dell'algebra letterale con i numeri.

Proposizione 3.2. *Sia A un anello. Allora:*

(1) *l'elemento neutro dell'addizione è unico;*
(2) *ogni elemento $a\in A$ ha un unico opposto rispetto all'addizione, denotato con $-a$;*
(3) *l'elemento neutro della moltiplicazione è unico;*
(4) *se un elemento $a\in A$ ha inverso rispetto alla moltiplicazione, questo è unico e si denota con a^{-1};*
(5) *se $a\in A$, allora $a0=0a=0$;*
(6) *se $a,b\in A$, allora $-(ab)=(-a)b=a(-b)$;*
(7) *l'addizione è commutativa, cioè $a+b=b+a$, per ogni $a,b\in A$.*

Dimostrazione. Le asserzioni (1), (2), (3) e (4) sono ben note in algebra astratta. Supponiamo che X sia un insieme dotato di un'operazione associativa $*$. Se e ed e' sono entrambi elementi neutri per $*$, abbiamo

$$e = ee' = e'.$$

Se y e y' sono inversi di x rispetto a $*$, allora

$$y = e * y = (y' * x) * y = y' * (x * y) = y' * e = y'.$$

Vediamo la (5). Certamente $0 = 0 + 0$, quindi, ponendo $x = a0$, abbiamo

$$x = a0 = a(0 + 0) = a0 + a0 = x + x,$$

Ne segue che $0 = x + (-x) = (x + x) + (-x) = x + (x + (-x)) = x + 0 = x$. Analogamente si verifica che $0a = 0$.

Vediamo la (6). Eseguiamo la somma fra ab e $(-a)b$:

$$ab + (-a)b = (a + (-a))b = 0b = 0,$$

quindi $(-a)b$ è un opposto di ab, quindi $(-a)b = -(ab)$, per l'unicità dell'opposto. Analogamente $-(ab) = a(-b)$.

Vediamo la (7). Sviluppiamo $(1 + 1)(a + b)$ in due modi:

$$(1 + 1)(a + b) = 1(a + b) + 1(a + b) = (a + b) + (a + b) = a + b + a + b,$$
$$(1 + 1)(a + b) = (1 + 1)a + (1 + 1)b = (a + a) + (b + b) = a + a + b + b.$$

Abbiamo omesso le parentesi per l'associatività dell'addizione; possiamo sommare $-a$ a sinistra dei due termini, ottenendo

$$b + a + b = a + b + b$$

e sommare $-b$ a destra dei due termini, ottenendo

$$b + a = a + b,$$

come richiesto. □

In virtù della (6) scriveremo $-ab = -(ab)$, non c'è ambiguità. Abbrevieremo anche $a + (-b)$ scrivendo $a - b$.

Vediamo ora la ragione vera per cui non si può "dividere per 0".

Proposizione 3.3. *Sia A un anello. Se 0 ammette un inverso rispetto alla moltiplicazione, allora $A = \{0\}$.*

Dimostrazione. Sia b un elemento tale che $0b = 1$; allora, per quanto visto prima, $0 = 1$. Se $a \in A$, abbiamo $a = a1 = a0 = 0$. □

Questo spiega anche le condizioni $0 \neq 1$ nelle definizioni di dominio e di campo, un anello in cui $0 = 1$ consiste di un solo elemento e non è molto utile.

Esempi di anelli sono quello **Z** degli interi, **Q** dei razionali, **R** dei reali. Gli ultimi due sono campi, **Z** è un dominio che non è un campo.

Proposizione 3.4. *Ogni campo è un dominio.*

Dimostrazione. Sia A un campo e supponiamo che $a, b \in A$, con $ab = 0$. Se $b = 0$, non c'è nulla da dimostrare; se $b \neq 0$, abbiamo

$$0 = 0b^{-1} = abb^{-1} = a1 = a,$$

quindi $a = 0$. □

Esistono domini che non sono campi, **Z** è un esempio. Il seguente risultato mostra una tecnica fondamentale nelle questioni di finitezza.

Proposizione 3.5. *Ogni dominio finito è un campo.*

Dimostrazione. Sia A un dominio e sia $a \in A$, $a \neq 0$. Consideriamo l'applicazione $f_a \colon A \to A$ definita da $x \mapsto ax$; è evidente che f_a è iniettiva: se $f_a(x) = f_a(y)$, abbiamo $ax = ay$, cioè $0 = ax - ay = a(x-y)$; poiché siamo in un dominio e $a \neq 0$, ne segue $x - y = 0$, cioè $x = y$.

L'applicazione f_a è allora suriettiva, perché A è un insieme finito, dunque esiste $x \in A$ tale che $f_a(x) = ax = 1$. Dunque $x = a^{-1}$. □

Nel seguito indicheremo con $\mathbf{0}$ e con $\mathbf{1}$ lo zero e l'uno dei numeri interi; parimenti indicheremo i numeri interi con carattere nero. Questa convenzione verrà abbandonata, non appena sarà chiaro che non esistono possibili ambiguità.

Definizione 3.6. Se A è un anello, $a \in A$ e $\mathbf{n} \in \mathbf{N}$, definiamo

$$\mathbf{0}a = 0, \quad (\mathbf{n}+\mathbf{1})a = \mathbf{n}a + a,$$
$$a^{\mathbf{0}} = 1 \quad a^{\mathbf{n}+\mathbf{1}} = a^{\mathbf{n}}a.$$

Se $\mathbf{n} \in \mathbf{Z}$ e $\mathbf{n} < 0$, poniamo anche $\mathbf{n}a = (-\mathbf{n})(-a)$.

Un ben noto esercizio di algebra astratta è di verificare che valgono le usuali proprietà.

Proposizione 3.7. *Sia A un anello e siano $a, b \in A$, $\mathbf{m}, \mathbf{n} \in \mathbf{Z}$. Allora*

$$(\mathbf{m}+\mathbf{n})a = (\mathbf{m}a) + (\mathbf{n}a),$$
$$\mathbf{n}(a+b) = (\mathbf{n}a) + (\mathbf{n}b).$$

Se $\mathbf{m}, \mathbf{n} \geq 0$, allora

$$a^{\mathbf{m}+\mathbf{n}} = a^{\mathbf{m}}a^{\mathbf{n}}.$$

Se, inoltre, A è commutativo, allora

$$(ab)^{\mathbf{n}} = a^{\mathbf{n}}b^{\mathbf{n}}.$$

L'elemento $\mathbf{n}a$ si dice il multiplo secondo $\mathbf{n}$ di a; $a^{\mathbf{n}}$ si dice la potenza di esponente $\mathbf{n}$ di a.

Esiste un altro concetto fondamentale, quello di omomorfismo.

Definizione 3.8. Siano A e B anelli e sia $\alpha \colon A \to B$ un'applicazione. Diremo che α è un omomorfismo se

(1) $\alpha(x+y) = \alpha(x) + \alpha(y)$, per ogni $x, y \in A$;
(2) $\alpha(xy) = \alpha(x)\alpha(y)$, per ogni $x, y \in A$;
(3) $\alpha(1) = 1$.

Notiamo che nei primi membri si usano le operazioni di A e nel secondo membro le operazioni di B.

Un omomorfismo α si dice *isomorfismo* se è un'applicazione biiettiva. In tal caso α^{-1} è un omomorfismo (esercizio). Due anelli A e B tali che esista un isomorfismo $\alpha \colon A \to B$ si dicono *isomorfi*. Due anelli isomorfi sono indistinguibili dal punto di vista dell'algebra: tutto quello che si può dire di uno si può dire anche dell'altro.

Proposizione 3.9. *Siano A e B anelli e sia $\alpha \colon A \to B$ un omomorfismo. Allora $\alpha(0) = 0$ e, per ogni $x \in A$, $\alpha(-x) = -\alpha(x)$.*

Dimostrazione. Poniamo $b = \alpha(0)$: abbiamo

$$b = \alpha(0) = \alpha(0+0) = \alpha(0) + \alpha(0) = b + b$$

e, come in precedenza, $b = 0$.

Abbiamo poi

$$\alpha(x) + \alpha(-x) = \alpha(x-x) = \alpha(0) = 0,$$

quindi $\alpha(-x)$ è l'opposto di $\alpha(x)$. □

L'anello $\mathbf{Z}$ degli interi ha una proprietà molto importante. Nella dimostrazione seguente denoteremo con 1_A l'uno dell'anello A.

Proposizione 3.10. *Se A è un anello, esiste un unico omomorfismo $\chi_A \colon \mathbf{Z} \to A$.*

Dimostrazione. Supponiamo che χ_A esista. Se $\mathbf{n} \in \mathbf{Z}$, $\mathbf{n} > 0$, abbiamo

$$\chi_A(\mathbf{n}) = \chi_A(\underbrace{1+1+\cdots+1}_{\mathbf{n}\text{ volte}}) = \underbrace{\chi_A(1)+\cdots+\chi_A(1)}_{\mathbf{n}\text{ volte}} = \underbrace{1_A+\cdots+1_A}_{\mathbf{n}\text{ volte}} = \mathbf{n}1_A$$

(si faccia induzione su $\mathbf{n}$).

Ne segue che $\chi_A(\mathbf{n}) = \mathbf{n}1_A$, anche per $\mathbf{n} < 0$. Dunque χ_A, se esiste, è unico.

Definiamo allora $\chi_A(\mathbf{n}) = \mathbf{n}1_A$ e verifichiamo che si tratta di un omomorfismo.

Se $\mathbf{m}, \mathbf{n} \in \mathbf{Z}$, abbiamo

$$\chi_A(\mathbf{m}+\mathbf{n}) = (\mathbf{m}+\mathbf{n})1_A = \mathbf{m}1_A + \mathbf{n}1_A = \chi_A(\mathbf{m}) + \chi_A(\mathbf{n}),$$

per le proprietà dei multipli. È chiaro dalla definizione che $\chi_A(\mathbf{1}) = \mathbf{1}1_A = 1_A$.

Si tratta ora di verificare che $\chi_A(\mathbf{mn}) = \chi_A(\mathbf{m})\chi_A(\mathbf{n})$, cioè che

$$(\mathbf{mn})1_A = (\mathbf{m}1_A)(\mathbf{n}1_A).$$

Si lascia per esercizio: dividere in quattro casi e, nel caso $\mathbf{m}, \mathbf{n} \geq 0$, fare induzione su $\mathbf{n}$. □

A questo punto è chiaro che si può abbandonare la convenzione di scrivere gli interi in carattere diverso; la scrittura $0a$ può ad esempio indicare il prodotto fra due elementi di A oppure il multipli secondo 0 di a: non fa differenza, in entrambi i casi l'elemento indicato è lo zero di A. Allo stesso modo non è necessario distinguere fra l'uno degli interi e quello dell'anello A. L'unico caso in cui ci potrebbe essere ambiguità è quello dell'anello $\mathbf{Z}$, ma qui multipli e prodotti coincidono!

Se $\alpha \colon A \to B$ è un omomorfismo di anelli, definiamo

$$\ker \alpha = \{a \in A : \alpha(a) = 0\}.$$

Questo sottoinsieme si chiama il *nucleo* di α (in inglese *kernel*). Notiamo che $0 \in \ker \alpha$.

Proposizione 3.11. *Sia $\alpha \colon A \to B$ un omomorfismo di anelli. Allora α è un'applicazione iniettiva se e solo se* $\ker \alpha = \{0\}$.

Dimostrazione. ($\Rightarrow$) Se $x \in \ker \alpha$, allora $\alpha(x) = 0 = \alpha(0)$. Per l'iniettività, $x = 0$.

($\Leftarrow$) Supponiamo che $\alpha(x) = \alpha(y)$; allora

$$0 = \alpha(x) - \alpha(y) = \alpha(x) + \alpha(-y) = \alpha(x-y),$$

quindi $x - y \in \ker \alpha$ e $x - y = 0$, per ipotesi. □

Definizione 3.12. Un sottoinsieme I dell'anello A si dice un *ideale* se:

(1) $0 \in I$;
(2) se $x, y \in I$, allora $x + y \in I$;
(3) se $x \in I$, allora $-x \in I$;
(4) se $x \in I$ e $a \in A$, allora $ax \in I$ e $xa \in I$.

Esempi banali di ideali dell'anello A sono $\{0\}$ e A.

Proposizione 3.13. *Sia $\alpha \colon A \to B$ un omomorfismo di anelli. Allora* $\ker \alpha$ *è un ideale di A.*

Dimostrazione. Che $0 \in \ker \alpha$ è noto. Siano $x, y \in \ker \alpha$; allora

$$\alpha(x+y) = \alpha(x) + \alpha(y) = 0 + 0 = 0,$$
$$\alpha(-x) = -\alpha(x) = -0 = 0.$$

Se poi $a \in A$, abbiamo

$$\alpha(ax) = \alpha(a)\alpha(x) = \alpha(a)0 = 0,$$
$$\alpha(xa) = \alpha(x)\alpha(a) = 0\alpha(a) = 0,$$

quindi la tesi. □

L'aspetto importante del concetto di ideale è che gli ideali di A sono esattamente i nuclei degli omomorfismi di dominio A.

Sia I un ideale di A; definiamo una relazione $\sim_I$ in A ponendo

$$x \sim_I y \text{ quando } x - y \in I.$$

È immediato verificare che $\sim_I$ è una relazione di equivalenza su A (esercizio). Di fatto $\sim_I$ ha altre due proprietà importanti. Supponiamo che $x \sim_I y$ e che $x' \sim_I y'$; allora

$$x + x' \sim_I y + y',$$
$$xx' \sim_I yy'.$$

La prima è dimostrata facilmente:

$$(x + x') - (y + y') = x + x' - y - y' = (x - y) + (x' - y') \in I.$$

La seconda ha bisogno di un trucco:

$$xx' - yy' = xx' - xy' + xy' - yy' = x(x' - y') + (x - y)y' \in I$$

per le proprietà di I (quali?).

Indicheremo con $[x]_I$ la classe di equivalenza di x rispetto a $\sim_I$ e con A/I l'insieme quoziente.

Su A/I definiamo due operazioni:

$$(*) \qquad [x]_I + [x']_I = [x + x']_I, \qquad [x]_I[x']_I = [xx']_I.$$

Se dimostriamo che queste sono *buone definizioni*, cioè non dipendono dai rappresentanti delle classi di equivalenza, abbiamo introdotto su A/I la struttura di anello, perché queste operazioni soddisfano certamente la definizione.

Ora, se $x \sim_I y$ e $x' \sim_I y'$, abbiamo $x + x' \sim_I y + y'$ e $xx' \sim_I yy'$, cioè quanto richiesto.

L'anello A/I con queste operazioni si dice *anello quoziente* di A modulo I. l'elemento neutro per l'addizione è $[0]_I$ e quello per la moltiplicazione è $[1]_I$.

Proposizione 3.14. *Se I è un ideale dell'anello A, allora $\pi_I \colon A \to A/I$ definita da $\pi_I(x) = [x]_I$ è un omomorfismo e $\ker \pi_I = I$.*

Dimostrazione. La prima parte è banale. Inoltre $\pi_I(x) = [0]_I$ se e solo se $x \sim_I 0$, cioè se e solo se $x \in I$. □

Di particolare importanza sono gli ideali di $\mathbf{Z}$. Se $n \in \mathbf{Z}$, indichiamo con $n\mathbf{Z}$ l'insieme di tutti i multipli interi di n. È facile vedere che $n\mathbf{Z}$ è un ideale di $\mathbf{Z}$.

Proposizione 3.15. *Sia I un ideale di $\mathbf{Z}$; allora esiste uno ed un solo $n \geq 0$ tale che $I = n\mathbf{Z}$.*

Dimostrazione. Se $I = \{0\}$, allora $I = 0\mathbf{Z}$. Supponiamo $I \neq \{0\}$: allora esiste $x \in I$, $x \neq 0$. Perciò, per definizione, $-x \in I$; dunque non è restrittivo fissare $n \in I$, $n > 0$. Possiamo anche fare di meglio: possiamo supporre che n sia il *minimo* intero positivo tale che $n \in I$. Difatti, se $a \in I$, $a > 0$,

$$\{m \in I : 0 < m < a\}$$

è finito e perciò ha un più piccolo elemento.

È evidente che $n\mathbf{Z} \subseteq I$, per le proprietà degli ideali. Sia ora $x \in I$: possiamo eseguire la divisione con resto di x per n:

$$x = nq + r, \qquad 0 \leq r < n.$$

Ma $x \in I$ e $nq \in I$, quindi $r = x - nq \in I$, ancora per le proprietà degli ideali. Se fosse $r > 0$ si contraddirebbe la minimalità di n, quindi $r = 0$ e $x \in n\mathbf{Z}$. Dunque $I \subseteq n\mathbf{Z}$ e l'uguaglianza $I = n\mathbf{Z}$ è provata.

L'unicità di n è ovvia. □

Qual è la relazione di equivalenza associata a $n\mathbf{Z}$? Per $n = 0$ la risposta è facile: $x \sim_{0\mathbf{Z}} y$ se e solo se $x = y$. In tal caso le classi di equivalenza consistono tutte di un solo elemento.

Supponiamo allora $n > 0$: abbiamo

$$x \sim_{n\mathbf{Z}} x' \quad \text{se e solo se} \quad x - x' \in n\mathbf{Z}.$$

Se allora $x - x' \in n\mathbf{Z}$, eseguiamo le divisioni di x e x' per n:

$$x = nq + r \quad (0 \leq r < n), \qquad x' = nq' + r' \quad (0 \leq r' < n).$$

Ammettiamo, per assurdo, che $r \neq r'$; non è allora restrittivo assumere $r > r'$. Allora $x - x' = n(q - q') + (r - r')$ e $0 \leq r - r' < n$. Ma questo contraddice il fatto che $x - x'$ sia un multiplo di n: contraddizione. Dunque $r = r'$.

Viceversa, prendiamo $x, x' \in \mathbf{Z}$ tali che la divisione di x e di x' per n dia lo stesso resto r; allora è evidente che $x - x' \in n\mathbf{Z}$.

Proposizione 3.16. *Sia $n > 0$. Allora, dati $x, x' \in \mathbf{Z}$, abbiamo $x \sim_{n\mathbf{Z}} x'$ se e solo se la divisione di x e x' per n dà lo stesso resto. In particolare il seguente è l'elenco completo e senza ripetizioni delle classi di equivalenza in $\mathbf{Z}/n\mathbf{Z}$:*

$$[0]_{n\mathbf{Z}}, [1]_{n\mathbf{Z}}, [2]_{n\mathbf{Z}}, \ldots, [n-1]_{n\mathbf{Z}}.$$

Dimostrazione. Se $x \in \mathbf{Z}$ e $x = nq + r$, con $0 \leq r < n$, è chiaro che $x \sim_{n\mathbf{Z}} r$. Se poi $0 \leq r < s < n$, le classi di equivalenza $[r]_{n\mathbf{Z}}$ e $[s]_{n\mathbf{Z}}$ sono distinte. □

L'anello $\mathbf{Z}/n\mathbf{Z}$ si chiama *anello delle classi resto modulo n*, per $n > 0$.

Ritorniamo all'omomorfismo χ_A: è evidente ora che esiste un unico $n \geq 0$ tale che $\ker \chi_A = n\mathbf{Z}$. Questo n si chiama *caratteristica* dell'anello A.

Un anello di caratteristica 1 ha un solo elemento $0 = 1$. Un anello di caratteristica 0 è infinito, in quanto χ_A è un'applicazione iniettiva.

Proposizione 3.17. *Sia A un anello di caratteristica n. Se A è un dominio, allora $n = 0$ oppure n è un numero primo.*

Dimostrazione. Se $n > 0$ non è primo, è certamente maggiore di uno e possiamo fattorizzarlo come $n = hk$, con $h, k > 1$. Allora $h, k \notin \ker \chi_A = n\mathbf{Z}$ e quindi $x = \chi_A(h) \neq 0$ e $y = \chi_A(k) \neq 0$. D'altra parte

$$xy = \chi_A(h)\chi_A(k) = \chi_A(hk) = \chi_A(n) = 0,$$

e quindi A non sarebbe un dominio: assurdo. □

Non è vero il viceversa, in generale; lo è nel caso in cui l'anello di cui si parla è $\mathbf{Z}/n\mathbf{Z}$.

Proposizione 3.18. *Se p è un numero primo, allora $\mathbf{Z}/p\mathbf{Z}$ è un dominio e quindi un campo.*

Dimostrazione. Siano $x, y \in \mathbf{Z}$ tali che $[x]_{p\mathbf{Z}}[y]_{p\mathbf{Z}} = [0]_{p\mathbf{Z}}$. Allora $[xy]_{p\mathbf{Z}} = [0]_{p\mathbf{Z}}$, cioè xy è un multiplo di p. Per il teorema 2.15, uno fra x e y è multiplo di p, quindi $[x]_{p\mathbf{Z}} = [0]_{p\mathbf{Z}}$ oppure $[y]_{p\mathbf{Z}} = [0]_{p\mathbf{Z}}$. Siccome poi $\mathbf{Z}/p\mathbf{Z}$ è finito, esso è un campo. □

È istruttivo trovare l'inverso di un elemento $[a]_{p\mathbf{Z}} \in \mathbf{Z}/p\mathbf{Z}$. Naturalmente occorre supporre che a non sia un multiplo di p: ne segue che $\mathrm{mcd}(a,p) = 1$ e quindi che esistono $x, y \in \mathbf{Z}$ tali che $1 = ax + py$. Passando alle classi di equivalenza, si ha

$$[1]_{p\mathbf{Z}} = [ax + py]_{p\mathbf{Z}} = [a]_{p\mathbf{Z}}[x]_{p\mathbf{Z}} + [p]_{p\mathbf{Z}}[y]_{p\mathbf{Z}} = [a]_{p\mathbf{Z}}[x]_{p\mathbf{Z}}.$$

Dunque $[x]_{p\mathbf{Z}}$ è l'inverso cercato. Tuttavia lo stesso ragionamento funziona più in generale.

Proposizione 3.19. *Sia $n > 0$ e sia $0 < r < n$. Allora $[r]_{n\mathbf{Z}}$ è invertibile in $\mathbf{Z}/n\mathbf{Z}$ se e solo se* $\mathrm{mcd}(r,n) = 1$.

Dimostrazione. ($\Rightarrow$) Se $\mathrm{mcd}(r,n) = 1$, esistono $x, y \in \mathbf{Z}$ tali che $rx + ny = 1$. Con lo stesso conto precedente, vediamo che $[r]_{n\mathbf{Z}}[x]_{n\mathbf{Z}} = [1]_{n\mathbf{Z}}$.

($\Leftarrow$) Supponiamo che esista $x \in \mathbf{Z}$ tale che $[r]_{n\mathbf{Z}}[x]_{n\mathbf{Z}} = [1]_{n\mathbf{Z}}$. Allora $rx - 1$ è un multiplo di n, poniamo $rx - 1 = ny$. Vogliamo provare che da ciò segue che $\mathrm{mcd}(r,n) = 1$.

Infatti, se $d = \mathrm{mcd}(r,n)$, possiamo scrivere $r = dr'$ e $n = dn'$. Ma allora $1 = rx - ny = dr'x - dn'y = d(r'x - n'y)$, da cui la tesi. □

Definizione 3.20. Se $n > 0$, $\varphi(n)$ denota il numero degli interi r tali che $0 < r < n$ e $\mathrm{mcd}(r,n) = 1$. Si pone poi $\varphi(0) = 0$ e $\varphi(1) = 0$. L'applicazione $\varphi\colon \mathbf{N} \to \mathbf{N}$ si chiama *funzione φ di Eulero*.

Per dimostrare il teorema seguente in modo rapido si può fare riferimento a un risultato di teoria dei gruppi.

Teorema 3.21. *Se G è un gruppo finito con n elementi, allora, per ogni $g \in G$, $g^n = 1$.*

Possiamo applicare questo teorema in quanto l'insieme degli elementi invertibili di un anello è un gruppo rispetto alla moltiplicazione (esercizio). Ne viene come corollario un celebre teorema di Eulero, il quale generalizza il piccolo teorema di Fermat.

Teorema 3.22 (Eulero)**.** *Se a e n sono numeri naturali e* $\mathrm{mcd}(a,n) = 1$, *allora*

$$a^{\varphi(n)} \sim_{n\mathbf{Z}} 1.$$

Dimostrazione. La cosa è banale per $n = 0$ oppure $n = 1$. Supponiamo allora $n > 1$.

Dire che $\mathrm{mcd}(a,n) = 1$ significa esattamente dire che $[a]_{n\mathbf{Z}}$ appartiene al gruppo degli elementi invertibili di $\mathbf{Z}/n\mathbf{Z}$, che ha $\varphi(n)$ elementi. Per il teorema 3.21,

$$[a]_{n\mathbf{Z}}^{\varphi(n)} = [1]_{n\mathbf{Z}}$$

che è precisamente la tesi. □

Teorema 3.23 (Fermat)**.** *Se p è un numero primo e a è un numero naturale tale che $p \nmid a$, allora*

$$a^{p-1} \sim_{p\mathbf{Z}} 1.$$

Dimostrazione. Si tratta di applicare il teorema di Eulero al caso di $n = p$; in questo caso $\mathbf{Z}/p\mathbf{Z}$ è un campo, quindi gli elementi invertibili sono esattamente $p - 1$ (l'unico non invertibile è lo zero). In altre parole, $\varphi(p) = p - 1$. □

Ecco un'altra applicazione; vogliamo trovare una proprietà del numero $(p-1)!$, dove p è un numero primo.

Teorema 3.24 (Wilson)**.** *Se p è un numero primo, allora*

$$(p-1)! \equiv -1 \pmod{p}.$$

Dimostrazione. Prendiamo un elemento $a \in \mathbf{Z}/p\mathbf{Z}$ tale che $a^2 \neq [1]$. Allora anche $(a^{-1})^2 \neq [1]$, quindi se eseguiamo il prodotto di tutti gli elementi non nulli di $\mathbf{Z}/p\mathbf{Z}$ possiamo accoppiare ciascun elemento con il suo inverso, esclusi $[1]$ e $-[1] = [p-1]$. Dunque questo prodotto è proprio $-[1]$. Ma allora abbiamo esattamente la tesi:

$$[1] \cdot [2] \cdot [3] \cdot \cdots \cdot [p-2] \cdot [p-1] = -[1]$$

è equivalente a

$$(p-1)! \equiv -1 \pmod{p}. \qquad \square$$

Notiamo anche che questo è un *test di primalità*: il numero naturale n è primo se e solo se $(n-1)! \equiv -1 \pmod{n}$. Una direzione è il teorema di Wilson. Supponiamo che n non sia primo, quindi che $p \mid n$, con p primo, $p < n$. Allora è chiaro che $p \mid (n-1)!$, quindi $p \mid \mathrm{mcd}((n-1)!, n)$ e dunque $[(n-1)!]$ non è invertibile in $\mathbf{Z}/n\mathbf{Z}$. In particolare $[(n-1)!] \neq -[1]$.

Esempio 3.25. Tutte le congruenze siano modulo 29:

$$\begin{array}{rrrr}
1 \equiv 1 & 1 \cdot 2 \equiv 2 & 2 \cdot 3 \equiv 6 & 6 \cdot 4 \equiv 24 \\
24 \cdot 5 \equiv 4 & 4 \cdot 6 \equiv 24 & 24 \cdot 7 \equiv 23 & 23 \cdot 8 \equiv 10 \\
10 \cdot 9 \equiv 3 & 3 \cdot 10 \equiv 1 & 1 \cdot 11 \equiv 11 & 11 \cdot 12 \equiv 16 \\
16 \cdot 13 \equiv 5 & 5 \cdot 14 \equiv 12 & 12 \cdot 15 \equiv 6 & 6 \cdot 16 \equiv 9 \\
9 \cdot 17 \equiv 8 & 8 \cdot 18 \equiv 28 & 28 \cdot 19 \equiv 10 & 10 \cdot 20 \equiv 26 \\
26 \cdot 21 \equiv 24 & 24 \cdot 22 \equiv 6 & 6 \cdot 23 \equiv 22 & 22 \cdot 24 \equiv 6 \\
6 \cdot 25 \equiv 5 & 5 \cdot 26 \equiv 14 & 14 \cdot 27 \equiv 1 & 1 \cdot 28 \equiv 28
\end{array}$$

e effettivamente $(29-1)! = 28! \equiv 28 \equiv -1 \pmod{29}$.

Vediamo invece con le congruenze modulo 15:

$$1 \equiv 1 \qquad 1 \cdot 2 \equiv 2 \qquad 2 \cdot 3 \equiv 6 \qquad 6 \cdot 4 \equiv 9 \qquad 9 \cdot 5 \equiv 0$$

e da questo momento in poi tutti i prodotti sono $[0]$ in $\mathbf{Z}/15\mathbf{Z}$.

Come altra applicazione dell'algebra astratta si può dare la costruzione dei numeri reali con le successioni di Cauchy di numeri razionali.

Definizione 3.26. Se X è un insieme, una *successione* in X è un'applicazione $s \colon \mathbf{N} \to X$. Per motivi espositivi, invece dell'usuale notazione $s(n)$ si usa quella s_n, per indicare l'immagine di n tramite questa applicazione.

Sia ora A un anello e indichiamo con $\mathscr{S}(A)$ l'insieme di tutte le successioni in A. Su questo insieme definiamo due operazioni: per $s, t \in \mathscr{S}(A)$, poniamo

$$\begin{array}{rcl}
s + t \colon \mathbf{N} \to A & \qquad & st \colon \mathbf{N} \to A \\
n \mapsto s_n + t_n & & n \mapsto s_n t_n
\end{array}$$

Evidentemente $s + t$ e st sono ancora due successioni in A. Le due operazioni si chiamano addizione e moltiplicazione puntuali.

Proposizione 3.27. *L'insieme $\mathscr{S}(A)$ è un anello, rispetto alle operazioni di addizione e moltiplicazione puntuali.*

Dimostrazione. Noiosa, ma banale: esercizio. Occorre comunque notare quali sono gli elementi importanti: lo zero è la successione "costante 0", che indicheremo con $\mathbf{0}$; l'uno è la successione "costante 1", che indicheremo con $\mathbf{1}$. Abbiamo allora, per ogni $n \in \mathbf{N}$,

$$\mathbf{0}_n = 0, \qquad \mathbf{1}_n = 1.$$

L'opposto dell'elemento s si ottiene prendendo l'opposto di ogni termine della successione: $(-s)_n = -s_n$. □

L'anello $\mathscr{S}(A)$ non è certamente un campo, anche se lo è A: ad esempio esistono elementi s e t non nulli tali che $st = 0$; si prendano le successioni s con $s_0 = 1$ e $s_n = 0$ per $n > 0$ e t con $t_0 = 0$ e $t_1 = 1$ per $n > 0$. Esiste anche un omomorfismo iniettivo $\iota_A \colon A \to \mathscr{S}(A)$ definito con

$$\iota_A(a) \colon n \mapsto a.$$

In altre parole, $\iota_A(a)$ è la successione "costante a". Il fatto che ι_A sia un omomorfismo è evidente.

Definizione 3.28. Sia A un anello e sia $B \subseteq A$; diremo che B è un sottoanello di A se:

(1) $0 \in B$, $1 \in B$;
(2) se $x, y \in B$, allora $x + y \in B$ e $xy \in B$;
(3) se $x \in B$, allora $-x \in B$.

Se B è un sottoanello di A, allora B è esso stesso un anello, in quanto in esso sono definite le operazioni di addizione e di moltiplicazione e le proprietà richieste sono evidentemente soddisfatte.

Proposizione 3.29. *Se $\alpha \colon A \to B$ è un omomorfismo di anelli, allora* $\mathrm{Im}(\alpha)$ *è un sottoanello di B.*

Dimostrazione. Ricordiamo che un elemento $x \in B$ appartiene a $\mathrm{Im}(\alpha)$ se e solo se esiste $a \in A$ con $\alpha(a) = x$. Le proprietà richieste si verificano immediatamente. □

Ad esempio, $\mathrm{Im}(\iota_A)$ è un sottoanello di $\mathscr{S}(A)$, ed è isomorfo ad A (qual è l'isomorfismo?).

Esempio 3.30. L'anello $\mathscr{S}(A)$ ha anche molti ideali: consideriamo infatti l'insieme Z delle successioni *definitivamente nulle*, dove $s \in \mathscr{S}(A)$ è definitivamente nulla se esiste $n_0 \in \mathbf{N}$ tale che $s_n = 0$, per ogni $n \geq n_0$ (diremo anche, con linguaggio improprio, che la successione è nulla da un certo punto in poi). L'insieme Z è un ideale di $\mathscr{S}(A)$.

Se A è il campo $\mathbf{Q}$ dei numeri razionali, possiamo considerare il sottoanello $\mathscr{C}(\mathbf{Q})$ delle successioni di Cauchy: una successione $s \in \mathscr{S}(\mathbf{Q})$ si dice una *successione di Cauchy* se, per ogni $\varepsilon \in \mathbf{Q}$, $\varepsilon > 0$, esiste $\bar{n} \in \mathbf{N}$ tale che, per ogni $m, n \geq \bar{n}$,

$$|s_m - s_n| < \varepsilon.$$

In altre parole le differenze dei termini in s sono "arbitrariamente piccole". Che $\mathscr{C}(\mathbf{Q})$ sia un sottoanello è ovvio; è anche facile vedere che l'insieme $\mathscr{Z}(\mathbf{Q})$ delle successioni convergenti a zero è un ideale di $\mathscr{C}(\mathbf{Q})$. Possiamo perciò formare l'anello quoziente $\mathscr{C}(\mathbf{Q})/\mathscr{Z}(\mathbf{Q})$: questo è un campo ed è isomorfo al campo dei numeri reali. Infatti ogni numero reale è limite di una successione (di Cauchy) di numeri razionali ed ogni successione di Cauchy nei reali è convergente. Il passaggio all'anello quoziente serve proprio a "identificare" successioni di Cauchy che convergono allo stesso numero reali. Di fatto non occorre conoscere in anticipo i reali: basta *definire* i numeri reali come l'anello $\mathscr{C}(\mathbf{Q})/\mathscr{Z}(\mathbf{Q})$ e vedere immersi in essi i numeri razionali tramite l'omomorfismo iniettivo che si ottiene componendo $\iota_{\mathbf{Q}}$ con la proiezione sul quoziente.

Esempio 3.31. Esibiremo ora una successione di Cauchy in $\mathbf{Q}$ che non è convergente. Questo dimostra, se ce ne fosse bisogno, che l'omomorfismo canonico $\mathbf{Q} \to \mathscr{C}(\mathbf{Q})/\mathscr{Z}(\mathbf{Q})$, ottenuto componendo $\iota_{\mathbf{Q}}$ con la proiezione sul quoziente, non è suriettivo (e quindi non è un isomorfismo).

Poniamo $s_n = 1/n!$, dove il simbolo $n!$ denota il prodotto

$$n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-2) \cdot (n-1) \cdot n,$$

con la convenzione che $0! = 1$ e $1! = 1$. Allora $2! = 2$, $3! = 6$, $4! = 24$ e così via; $n!$ si legge *n fattoriale*. Chiaramente $n! \geq 2^{n-1}$, per $n \geq 1$ (induzione).

Definiamo una nuova successione $t \in \mathscr{S}(\mathbf{Q})$ con

$$t_n = s_0 + s_1 + \cdots + s_n.$$

Più precisamente, $t_0 = s_0$ e $t_{n+1} = t_n + s_{n+1}$. Usando la formula di 1.1, abbiamo che

$$\begin{aligned} t_n &= \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{(n-1)!} + \frac{1}{n!} \\ &\leq 1 + 1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-2}} + \frac{1}{2^{n-1}} \\ &= 1 + \frac{1-(1/2)^n}{1-(1/2)} = 1 + 2\left(1 - \frac{1}{2^n}\right) < 3. \end{aligned}$$

Dunque t è crescente e limitata e perciò è una successione di Cauchy. Ovviamente abbiamo anche $t_n > 2$, per ogni n.

Supponiamo, per assurdo, che t converga ad un numero razionale p/q (con p e q interi positivi), che sappiamo allora essere l'estremo superiore di $\{t_n : n \in \mathbf{N}\}$. In particolare $p/q > t_n$, per ogni n (perché?).

Poiché $2 < t_n < 3$, per ogni n, abbiamo anche $2 < p/q \leq 3$. Supponiamo, per assurdo, che $p/q = 3$ e consideriamo la successione t definita da $t_n = t_{n+k} - t_k$, dove $k \geq 0$ è fissato. È evidente (esercizio) che t converge a $3 - t_k$. Inoltre $t_0 = 0$ e, per $n > 0$,

$$\begin{aligned} t_n &= \frac{1}{(k+1)!} + \frac{1}{(k+2)!} + \cdots + \frac{1}{(k+n)!} \\ &\leq \frac{1}{(k+1)!}\left(1 + \frac{1}{k+1} + \frac{1}{(k+1)^2} + \cdots + \frac{1}{(k+1)^{n-1}}\right) \\ &\leq \frac{1}{(k+1)!}\left(1 + \frac{1}{1 - \dfrac{1}{k+1}}\right) \\ &= \frac{1}{(k+1)!}\frac{k+1}{k} = \frac{1}{k \cdot k!} \end{aligned}$$

Questo è indipendente da n, quindi $3 - t_k \leq 1/(k \cdot k!)$. Ora, già per $k = 2$, abbiamo

$$3 - t_2 = \frac{1}{2} \quad \text{e} \quad \frac{1}{2 \cdot 2!} = \frac{1}{4},$$

che produce una contraddizione.

Ne segue che $q \geq 2$, perché p/q non può essere intero. Fissiamo $\varepsilon > 0$ tale che

$$\frac{1}{q} + \varepsilon \cdot q! < 1.$$

Allora esiste n_0 tale che, per ogni $n \geq n_0$, si abbia

$$0 < \frac{p}{q} - t_n < \varepsilon.$$

Prendiamo allora $n \geq n_0$ e $n > q$. Abbiamo allora

$$\frac{p}{q} < t_q + \frac{1}{(q+1)!} + \frac{1}{(q+2)!} + \cdots + \frac{1}{n!} + \varepsilon,$$

e, moltiplicando ambo i membri per $q!$, otteniamo

$$0 < p \cdot (q-1)! - t_q \cdot q! < \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots + \frac{1}{(q+1)(q+2)\ldots(n-1)n} + \varepsilon \cdot q!$$

e $p \cdot (q-1)! - t_q \cdot q!$ è un numero intero, come si vede sviluppando

$$t_q \cdot q! = q! + q! + \frac{q!}{2!} + \cdots + \frac{q!}{(q-1)!} + \frac{q!}{q!},$$

e osservando che da $h \leq k$ segue $h! \mid k!$, per $h, k \in \mathbf{N}$. Ora, detto a il termine a destra dell'ultima disuaglianza, abbiamo

$$\begin{aligned} a &\leq \frac{1}{q+1} + \frac{1}{(q+1)^2} + \cdots + \frac{1}{(q+1)^{n-q}} + \varepsilon \cdot q! = \frac{1 - \dfrac{1}{(q+1)^{n-q+1}}}{1 - \dfrac{1}{q+1}} + \varepsilon \cdot q! - 1 \\ &= \frac{q+1}{q}\left(1 - \frac{1}{(q+1)^{n-q+1}}\right) + \varepsilon \cdot q! - 1 < \frac{q+1}{q} + \varepsilon \cdot q! - 1 \\ &= \frac{1}{q} + \varepsilon \cdot q! < 1 \end{aligned}$$

e ciò è assurdo.

Osservazione 3.32. Chi conosce l'Analisi ricorderà che il limite della successione t dell'esempio precedente è il *numero di Napier*, di solito indicato con e, che si prende come base dei logaritmi naturali. La dimostrazione precedente è proprio del fatto che e è irrazionale. È noto, ma più difficile da dimostrare, che e è un numero *trascendente*, cioè non è radice di alcun polinomio a coefficienti interi. Lo stesso vale per π, e fu dimostrato da Lindemann nel 1882; poco dopo Weierstrass dimostrò un risultato più generale, da cui segue anche la trascendenza di e.

La parte in cui dimostriamo che la successione t non converge a 3 può essere usata per calcolare il valore di e con quante cifre decimali si voglia: infatti quello che abbiamo effettivamente dimostrato è che

$$e - t_k \leq \frac{1}{k \cdot k!}.$$

Per esempio,

$$\frac{1}{12 \cdot 12!} < 1.8 \cdot 10^{-10}$$

e quindi $t_{12} = 2.71828182\ldots$ è il valore di e esatto fino all'ottava cifra decimale.

4. Aree di figure curvilinee

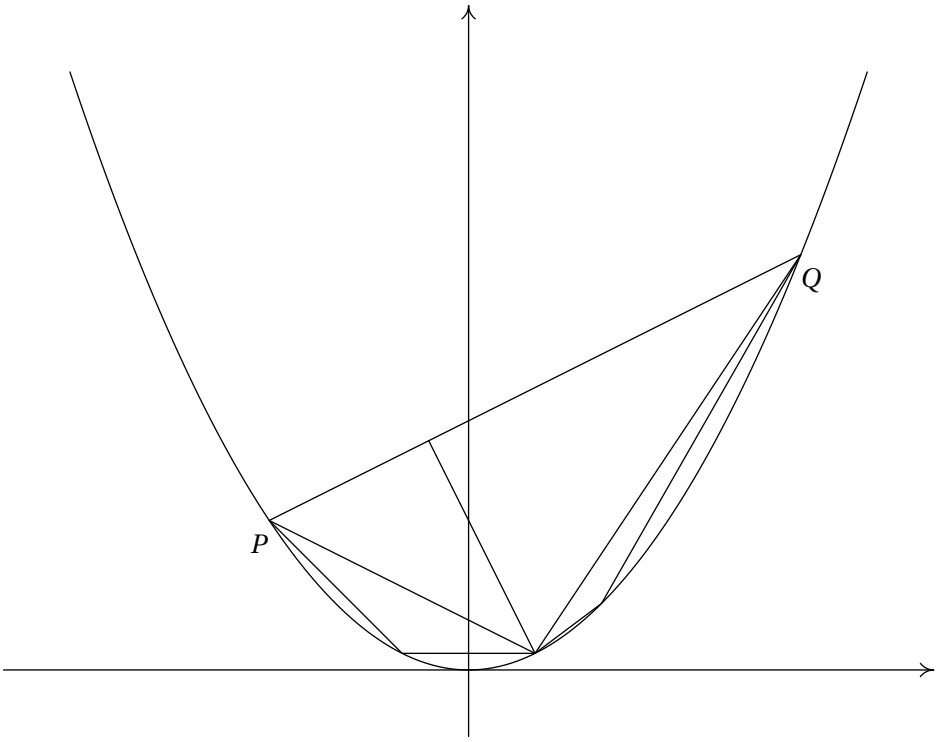

FIGURA 1. Calcolo dell'area del segmento di parabola

Consideriamo la parabola di equazione $y = ax^2$, con $a > 0$. Prendiamone due punti, $P(p, ap^2)$ e $Q(q, aq^2)$, dove $p < q$. Il *segmento parabolico* è la parte di piano compresa fra la retta PQ e la parabola.

Di questo segmento vogliamo calcolare l'area. Il primo a eseguire questo calcolo fu Archimede, con il suo famoso metodo di esaustione. Seguiremo un procedimento analogo al suo, e calcoleremo le aree di opportuni triangoli interni al segmento parabolico.

Il triangolo di area massima con vertici P e Q e il terzo vertice sulla parabola (naturalmente interno all'arco di estremi P e Q), si ottiene considerando il punto $R(r, ar^2)$ con $p < r < q$ tale che la distanza di R dalla retta PQ sia massima.

L'equazione della retta PQ si scrive facilmente:

$$\frac{y-ap^2}{aq^2-ap^2} = \frac{x-p}{q-p}$$

cioè, con facili calcoli,

$$y = a(p+q)x - apq.$$

La distanza del punto R da tale retta è allora

$$\begin{aligned} f(r) &= \frac{|ar^2 - ar(p+q) + apq|}{\sqrt{1+a^2(p+q)^2}} \\ &= \frac{a}{\sqrt{1+a^2(p+q)^2}} |r^2 - r(p+q) + pq| \\ &= \frac{a}{\sqrt{1+a^2(p+q)^2}} |(r-p)(r-q)| \end{aligned}$$

e ci basta considerare il massimo del fattore sotto valore assoluto. Questo, come è evidente, si ottiene per $r = (p+q)/2$.

Archimede sapeva dimostrare questo per via puramente geometrica, usando la definizione di parabola come luogo dei punti equidistanti da un punto fisso e da una retta fissa. Noi, avendo a disposizione la geometria analitica, facciamo molta meno fatica.

Chiameremo *base* del segmento parabolico la misura del segmento PQ e *altezza* la distanza massima appena determinata.

Qual è l'area del massimo triangolo inscritto? È facile calcolarla, base per altezza diviso due; la base è

$$b = \sqrt{(q-p)^2 + (aq^2 - ap^2)^2} = (q-p)\sqrt{1+a^2(p+q)^2},$$

mentre l'altezza è

$$h = f\left(\frac{p+q}{2}\right) = \frac{a}{\sqrt{1+a^2(p+q)^2}} \frac{(q-p)^2}{4}.$$

Perciò l'area è

$$A = \frac{1}{2}bh = \frac{a}{8}(q-p)^3.$$

Dunque quest'area dipende solo dal parametro a della parabola e dalla *proiezione* del segmento PQ sull'asse delle ascisse.

Vogliamo ora ripetere questo procedimento; a questo scopo, chiameremo A_0 l'area appena trovata e porremo $d = q - p$. Se osserviamo con attenzione, il triangolo massimo appena costruito definisce altri due segmenti parabolici più piccoli. La proiezione dei loro estremi avrà misura esattamente $d/2$, per come abbiamo determinato il punto di distanza massima. Quindi i due triangoli di area massima inscritti in questi segmenti, per lo stesso motivo di prima, hanno area

$$A_1 = \frac{a}{8}\left(\frac{d}{2}\right)^3.$$

Ciascuno di questi due triangoli definisce a sua volta due segmenti parabolici e in questi i triangoli di area massima (che sono in tutto quattro) hanno area

$$A_2 = \frac{a}{8}\left(\frac{d}{4}\right)^3.$$

Possiamo ripetere il procedimento quante volte vogliamo. Al passo n avremo 2^n segmenti parabolici e i triangoli di area massima contenuti in essi hanno ciascuno area

$$A_n = \frac{a}{8}\left(\frac{d}{2^n}\right)^3.$$

Se sommiamo le aree così ottenute, avremo un'approssimazione sempre migliore dell'area del segmento parabolico; la somma fino al passo n è

$$\begin{aligned} S_n &= A_0 + 2A_1 + 4A_2 + \cdots + 2^{n-1}A_{n-1} + 2^n A_n \\ &= \frac{ad^3}{8}\left(1 + \frac{1}{2^2} + \frac{1}{2^4} + \cdots + \frac{1}{2^{2n-2}} + \frac{1}{2^{2n}}\right) \\ &= \frac{ad^3}{8}\frac{1 - \dfrac{1}{2^{2n+2}}}{1 - \dfrac{1}{4}} \end{aligned}$$

Quando n diventa grande, la frazione $1/2^{2n+2}$ al numeratore diventa piccolissima; inoltre ogni punto del segmento parabolico appartiene a uno dei triangoli costruiti, per n sufficientemente grande. Perciò è chiaro che l'area del segmento parabolico è

$$S = \frac{ad^3}{8}\frac{1}{1 - \dfrac{1}{4}} = \frac{ad^3}{8}\frac{4}{3}.$$

Non ci interessa scriverla così, ma piuttosto in termini di base e altezza del segmento: le loro espressioni sono state calcolate prima e quindi

$$S = \frac{2}{3}bh.$$

Questa è la formula dovuta ad Archimede. È forse interessante conoscere come lo scienziato siracusano abbia raggiunto questo risultato. Prima ha formulato un'ipotesi, dimostrandola poi con metodi geometrici e l'esaustione; ma come è arrivato all'ipotesi? Semplice: costruendosi modelli di segmenti parabolici di legno e pesandoli!

Con lo stesso metodo di esaustione e le prove pratiche, riuscì a dimostrare numerose altre formule per aree e volumi; applicò il metodo delle triangolazioni anche per calcolare il rapporto fra la circonferenza e il diametro, ottenendo fra gli altri il valore approssimato $22/7$ per π che, per fini pratici, è ottimo. È però da notare che era in grado di calcolarlo molto più esattamente.

Il fatto interessante è che il problema corrispondente dell'area del segmento circolare (o, più in generale, ellittico) non porta a una formula così semplice. Se indichiamo con α l'angolo al centro di un settore circolare, l'area del segmento è

$$\frac{\alpha r^2}{2} - \frac{r^2}{2}\operatorname{sen}\alpha.$$

Questa non ammette semplificazioni significative in termini di base e altezza del segmento circolare; con termini classici queste si chiamano *corda* e *sagitta*. Più interessante è invece l'area del *settore circolare*, che è proporzionale all'ampiezza dell'angolo al centro.

C'è un'altra curva molto interessante, l'iperbole. Come sappiamo, la sua equazione normale è

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

e l'iperbole si dice *equilatera* se $a = b$. In questo caso i suoi asintoti sono perpendicolari e la possiamo allora vedere nel sistema di riferimento che ha come assi gli asintoti con equazione

$$xy = k$$

dove $k > 0$. Studieremo, per semplicità, il caso $k = 1$; gli altri si possono ottenere facilmente con una trasformazione affine.

In un certo senso l'iperbole equilatera è parente della circonferenza, quindi sembra interessante studiare l'area del *settore iperbolico*.

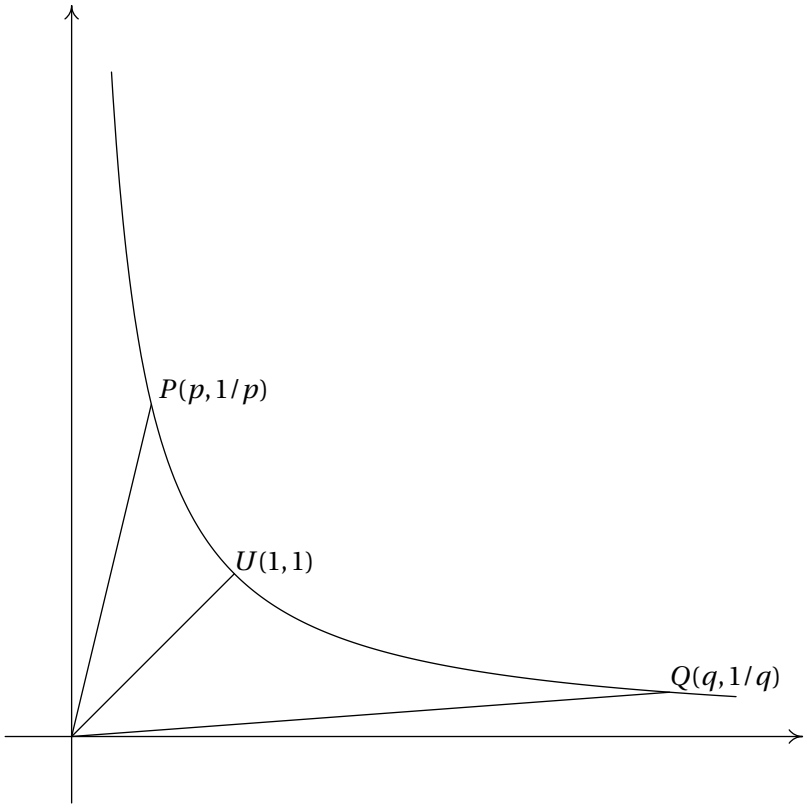

FIGURA 2. Settore iperbolico

Prendiamo una decisione che serve a semplificare i calcoli ed è analoga a quella che ci permette di misurare gli angoli dando valori sia positivi che negativi.

L'area del settore iperbolico OUQ sarà considerata positiva, mentre quella del settore iperbolico OUP sarà considerata negativa. Chiameremo $L(Q)$ l'area (con segno) del settore OUQ. Perché questa convenzione? La risposta è facile: l'area (questa volta senza considerare il segno) del settore OPQ è

$$|L(Q) - L(P)|$$

e questo vale anche se il punto P ha coordinate $(p, 1/p)$ con $p > 1$. Ci limiteremo allora a studiare settori iperbolici OUQ con $q > 1$.

Facciamo una divagazione. Vogliamo una formula semplice per calcolare l'area di un triangolo che abbia un vertice nell'origine. Siano $A(a, b)$ e $C(c, d)$ gli altri due vertici. La retta OC ha equazione $dx - cy = 0$, quindi la distanza di A da questa retta è

$$\frac{|ad - bc|}{\sqrt{c^2 + d^2}},$$

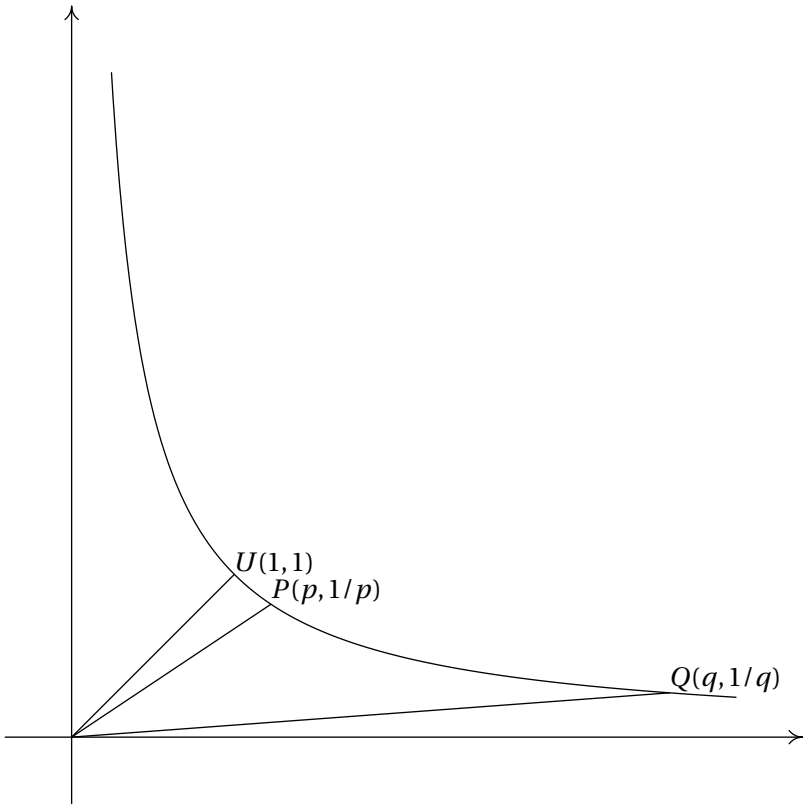

FIGURA 3. Area del settore iperbolico

quindi l'area del triangolo OAB, prendendo OC come base, è

$$\frac{1}{2}\sqrt{c^2+d^2}\frac{|ad-bc|}{\sqrt{c^2+d^2}} = \frac{1}{2}|ad-bc|$$

e la quantità in modulo è proprio la stessa che si otterrebbe dalla formula della regola di Cramer per la tabella

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Se i punti A e B stanno sull'iperbole, diciamo $A(a,1/a)$ e $B(b,1/b)$ (con $a,b>0$), dovremo allora calcolare il modulo di

$$a\frac{1}{b} - b\frac{1}{a} = \frac{a^2-b^2}{ab}$$

che è positivo se $a > b$.

Per approssimare l'area del settore iperbolico OUQ suddividiamo l'arco UQ in sottoarchi definiti da punti $Q_0 = U$, Q_1, Q_2, …, $Q_n = Q$, e calcoliamo l'area dei triangoli OUQ_1, OUQ_2, …, OUQ_n. Se infittiamo la suddivisione, otteniamo un'approssimazione sempre migliore dell'area del settore.

Quali punti di suddivisione scegliamo? Vogliamo n punti, non necessariamente equidistanti e una scelta possibile è di prendere

$$Q_k\left(\sqrt[n]{q^k}, \frac{1}{\sqrt[n]{q^k}}\right) \qquad k = 0,1,2,\ldots,n.$$

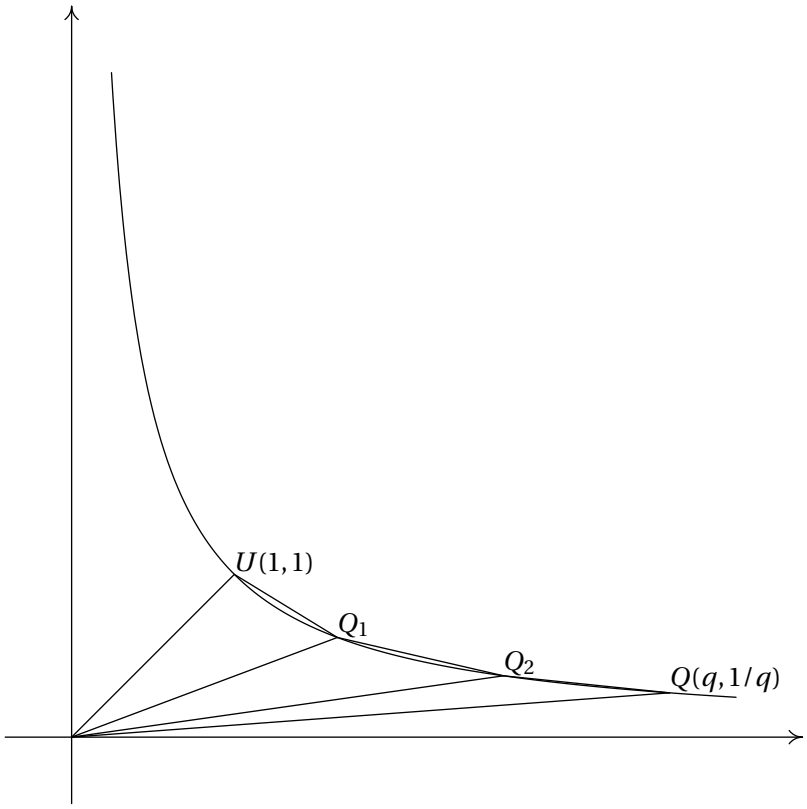

FIGURA 4. Approssimazione dell'area del settore iperbolico

L'area del triangolo $OQ_{k-1}Q_k$ è allora, tenendo conto della formula precedente,

$$\frac{1}{2}\left(\sqrt[n]{q^k}\frac{1}{\sqrt[n]{q^{k-1}}}-\sqrt[n]{q^{k-1}}\frac{1}{\sqrt[n]{q^k}}\right)=\frac{1}{2}\left(\sqrt[n]{q}-\frac{1}{\sqrt[n]{q}}\right)$$

dunque l'approssimazione con n punti è

$$l_{[n]}(q)=\frac{n}{2}\left(\sqrt[n]{q}-\frac{1}{\sqrt[n]{q}}\right).$$

Il risultato è valido, con le stesse convenzioni, quando $0<q<1$ e, in tal caso, $l_{[n]}(q)<0$ come ci attendevamo.

Se indichiamo con $l(q)$ l'area effettiva del settore iperbolico OUQ, i valori $l_{[n]}(q)$ ne sono un'approssimazione sempre migliore, quando n diventa grande. A differenza che nel caso del segmento parabolico e del settore circolare, abbiamo una funzione che ha proprietà inusuali.

Vediamo, per esempio, l'area del settore iperbolico OUQ, dove $Q(q,1/q)$, con $q = ab$. L'approssimazione al passo n è

$$\begin{aligned} l_{[n]}(ab) &= \frac{n}{2}\left(\sqrt[n]{ab} - \frac{1}{\sqrt[n]{ab}}\right) \\ &= \frac{n}{2}\left(\sqrt[n]{ab} - \frac{\sqrt[n]{b}}{\sqrt[n]{a}} + \frac{\sqrt[n]{b}}{\sqrt[n]{a}} - \frac{1}{\sqrt[n]{ab}}\right) \\ &= \frac{n}{2}\left(\sqrt[n]{b}\left(\sqrt[n]{a} - \frac{1}{\sqrt[n]{a}}\right) + \frac{1}{\sqrt[n]{a}}\left(\sqrt[n]{b} - \frac{1}{\sqrt[n]{b}}\right)\right) \\ &= l_{[n]}(a)\sqrt[n]{b} + l_{[n]}(b)\frac{1}{\sqrt[n]{a}}. \end{aligned}$$

Quando n diventa grande, $\sqrt[n]{a}$ si discosta di pochissimo da 1; perciò otteniamo la formula fondamentale

$$l(ab) = l(a) + l(b).$$

La possiamo usare allora per calcolare l'area di un settore iperbolico qualunque: ricordiamo che, se $P(p,1/p)$ e $Q(q,1/q)$, l'area è $|l(q) - l(p)|$; ponendo $ab = q$ e $a = p$, avremo $b = q/p$. Dunque

$$l(q) - l(p) = l(ab) - l(a) = l(b) = l(q/p).$$

La funzione l *trasforma prodotti in somme*. Per chi già conosce i logaritmi, questa è la loro proprietà!

Esiste un altro modo di introdurre i logaritmi, usando l'area di un'altra figura legata all'iperbole.

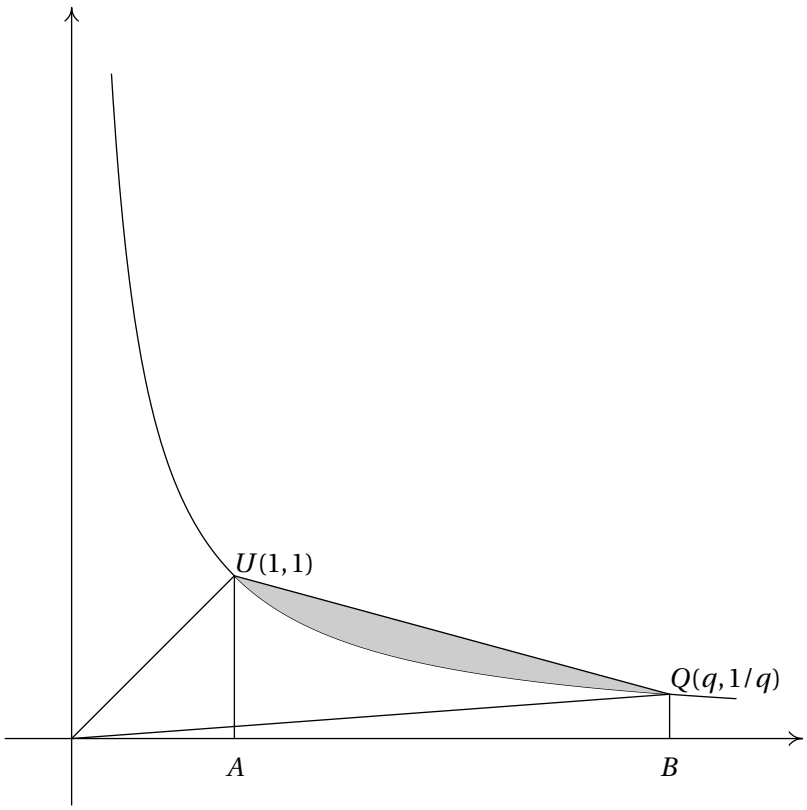

FIGURA 5. Altra definizione di logaritmo

I punti A e B hanno coordinate, rispettivamente, $(1,0)$ e $(q,0)$. Vogliamo calcolare l'area S della figura delimitata dal segmento AU, dall'arco di iperbole UQ, dal segmento QB e dal segmento AB. La parte ombreggiata è l'intersezione fra il triangolo OUQ e il trapezio $AUQB$; quindi la sua area è la differenza fra l'area del triangolo OUQ e l'area del settore iperbolico OUQ, quindi

$$\frac{1}{2}\left(q\cdot 1-1\cdot\frac{1}{q}\right)-l(q)=\frac{q^2-1}{2q}-l(q).$$

L'area del trapezio $AUQB$ vale, invece,

$$\frac{1}{2}\left(1+\frac{1}{q}\right)(q-1)=\frac{1}{2}\frac{q+1}{q}(q-1)=\frac{q^2-1}{2q}.$$

Perciò abbiamo, evidentemente, $S=l(q)$ e quindi un altro modo di rappresentare $l(q)$ come area.

5. Logaritmi

La costruzione dei numeri reali come classi di equivalenza di successioni di Cauchy di numeri razionali permetterebbe di definire facilmente la funzione esponenziale: sia $a \in \mathbf{R}$, $a > 0$; per ogni numero razionale r, è ben definita la potenza a^r. Se allora $s \in \mathscr{S}(\mathbf{Q})$ è una successione convergente al numero reale x, possiamo considerare la successione $t \in \mathscr{S}(\mathbf{R})$ definita da

$$t_n = a^{s_n}.$$

Se dimostriamo che t è una successione di Cauchy, abbiamo allora che t converge a un numero reale che denoteremo con a^x. Come sempre, è necessario anche dimostrare che questo numero reale non dipende dalla particolare successione s che abbiamo scelto.

Tuttavia esiste un altro modo di definire l'esponenziale che evita molte delle verifiche necessarie. L'esposizione che faremo usa solo il concetto di integrale, che può certamente essere sviluppato senza alcun riferimento alla funzione esponenziale. Il passo cruciale è di rovesciare l'ordine delle cose e definire il logaritmo.

È ben noto che, per $n \neq -1$ e $a > 0$,

$$\int_1^a t^n \, dt = \frac{a^{n+1} - 1}{n+1}.$$

Ovviamente la formula non può essere estesa per $n = -1$. Tuttavia, siccome la funzione $x \mapsto 1/x$ è continua in $(0, +\infty)$, l'integrale

$$\int_1^a \frac{1}{t} \, dt$$

esiste, per ogni $a > 0$.

Definizione 5.1. Sia $a \in \mathbf{R}$, $a > 0$; definiamo

$$\log a = \int_1^a \frac{1}{t} \, dt$$

e lo chiamiamo il *logaritmo naturale* di a. Si veda la figura 6 on the facing page, per l'interpretazione geometrica come area.

Dalla definizione di integrale sappiamo che $\log 1 = 0$ e che $\log a < 0$, per $0 < a < 1$. Osserviamo poi che il dominio di definizione della funzione logaritmo è $(0, +\infty)$. Dal teorema fondamentale del calcolo integrale segue anche che

$$\log' x = \frac{1}{x},$$

quindi che log è una funzione crescente, avendo derivata positiva in ogni punto; si tratta poi di una funzione *concava*, avendo derivata seconda $\log'' x = -1/x^2$ che è ovunque negativa. Il grafico è tracciato nella figura 7 on page 34.

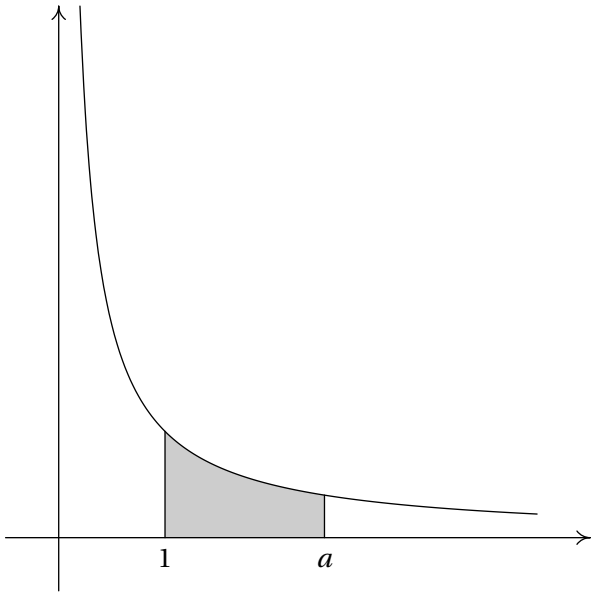

FIGURA 6. Definizione di $\log a$

Proposizione 5.2. *Se $a, b > 0$, allora*

$$\log(ab) = \log a + \log b.$$

Dimostrazione. Consideriamo la funzione f definita, per $x > 0$, da

$$f(x) = \log(ax)$$

e calcoliamo la derivata:

$$f'(x) = a \cdot \log'(ax) = a\frac{1}{ax} = \frac{1}{x}.$$

Perciò f e log hanno la stessa derivata e quindi, per il teorema di Lagrange, esiste una costante c tale che, per ogni $x > 0$, si abbia

$$f(x) = \log x + c.$$

Calcolando per $x = 1$ abbiamo

$$c = 0 + c = \log 1 + c = f(1) = \log(a \cdot 1) = \log a.$$

Ne segue che $\log(ax) = \log a + \log x$, per ogni $x > 0$. □

Corollario 5.3. *Se $a, b > 0$, allora*

$$\log\left(\frac{a}{b}\right) = \log a - \log b.$$

Dimostrazione. Abbiamo $\log a = \log(b(a/b)) = \log b + \log(a/b)$. □

In particolare, $\log a^{-1} = -\log a$.

Corollario 5.4. *Se m è un numero intero positivo e $a > 0$, allora $\log a^n = n \log a$. Di conseguenza*

$$\lim_{x \to +\infty} \log x = +\infty, \quad \lim_{x \to 0^+} \log x = -\infty$$

Dimostrazione. Per induzione su n, abbiamo che $\log a^n = n \log a$. Per $a = 2$, ad esempio,

$$\lim_{n \to +\infty} \log 2^n = \lim_{n \to +\infty} n \log 2 = +\infty.$$

□

Corollario 5.5. *Se r è un numero razionale, allora* $\log a^r = r \log a$.

Dimostrazione. Sia $r = m/n$ e poniamo $a^r = b$. Allora

$$n \log b = \log b^n = \log(a^{\frac{m}{n}})^n = \log a^m = m \log a,$$

da cui $\log a^r = r \log a$. □

Proposizione 5.6. *Esiste uno ed un solo numero reale e tale che* $\log e = 1$.

Dimostrazione. La funzione log è continua, essendo ovunque derivabile (dove è definita). Inoltre $\log 1 = 0$ e $\lim_{x \to +\infty} \log x = +\infty$. Perciò log deve assumere il valore 1 e, essendo crescente, lo assume una sola volta. □

In realtà la proposizione precedente può essere generalizzata facilmente.

Teorema 5.7. *Se $b \in \mathbf{R}$, esiste uno ed un solo numero reale $a > 0$ tale che*

$$\log a = b.$$

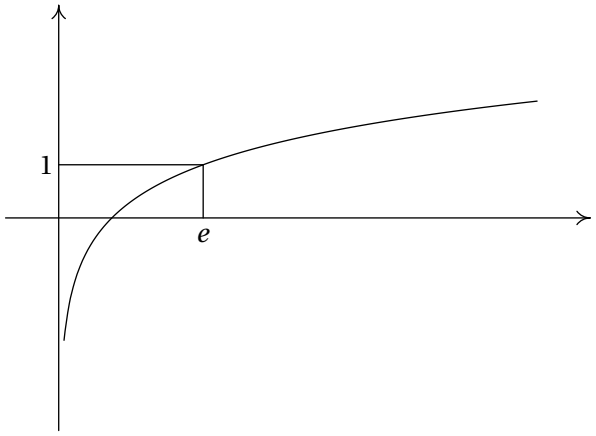

FIGURA 7. Grafico del logaritmo

Possiamo allora definire la funzione inversa di log e denotarla con exp, la funzione *esponenziale naturale*, il cui grafico è tracciato in figura 8 on the facing page.

Definizione 5.8. Per ogni numero reale x, $\exp x$ è l'unico numero reale $y > 0$ tale che

$$\log y = x.$$

In altre parole valgono le uguaglianze

$$\log \exp x = x, \quad \exp \log y = y,$$

per ogni $x \in \mathbf{R}$ e ogni $y \in \mathbf{R}$, $y > 0$.

La definizione della funzione esponenziale naturale come inversa del logaritmo naturale ha conseguenze immediate.

Proposizione 5.9. *Per ogni $x \in \mathbf{R}$,* $\exp' x = \exp x$*; inoltre*

$$\lim_{x \to -\infty} \exp x = 0 \quad e \quad \lim_{x \to +\infty} \exp x = +\infty.$$

La funzione esponenziale naturale è crescente.

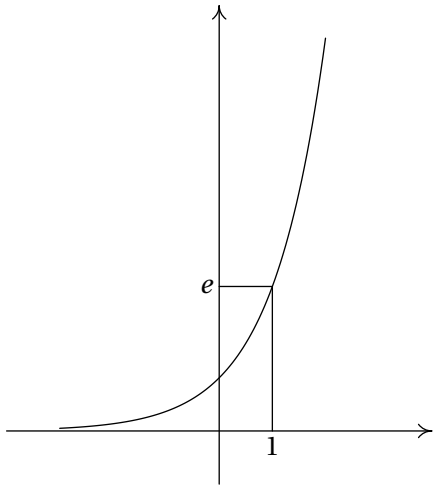

FIGURA 8. Grafico dell'esponenziale

Dimostrazione. Usiamo l'identità $\log \exp x = x$ e deriviamo:

$$1 = \log' \exp x \cdot \exp' x = \frac{1}{\exp x} \exp' x,$$

da cui $\exp' x = \exp x$ (stiamo usando la "formula per la derivata della funzione inversa"). Gli altri fatti sono immediate conseguenze della definizione. □

Vediamo ora le altre proprietà della funzione esponenziale naturale.

Proposizione 5.10. *Per ogni $a, b \in \mathbf{R}$,*

$$\exp(a+b) = \exp a \cdot \exp b.$$

Dimostrazione. Infatti $a + b = \log \exp(a+b)$ e

$$\log(\exp a \cdot \exp b) = \log \exp a + \log \exp b = a + b.$$ □

Allora, facilmente,

$$\exp(a-b) = \frac{\exp a}{\exp b}.$$

Abbiamo anche $\exp e = 1$, per definizione di e, e $\exp 0 = 1$.

Proposizione 5.11. *Se r è un numero razionale, allora*

$$\exp r = e^r.$$

Dimostrazione. Abbiamo $\log e^r = r \log e = r$, quindi $\exp r = \exp \log e^r = e^r$. □

Una volta che abbiamo questa relazione, viene spontaneo *definire* l'elevamento di e ad una qualunque potenza ad esponente reale x con

$$e^x = \exp x.$$

Possiamo anche fare di meglio.

Definizione 5.12. Se $a > 0$ e $x \in \mathbf{R}$, poniamo

$$a^x = e^{x \log a} = \exp(x \log a).$$

La notazione è giustificata dalla considerazione che, per $r \in \mathbf{Q}$, abbiamo

$$\exp(r \log a) = \exp \log(a^r) = a^r,$$

cioè a^x indica comunque la stessa cosa, se x è razionale. Possiamo anche definire, per $a > 0$, ma $a \neq 1$, e $y > 0$,

$$\log_a y = \frac{\log y}{\log a}.$$

Abbiamo allora le identità

$$a^{\log_a y} = y, \quad \log_a a^x = x,$$

valide per ogni x e ogni $y > 0$ (esercizio).

Osservazione 5.13. Non è possibile definire il logaritmo in base 1, in quanto $1^x = 1$, per ogni x.

Vogliamo calcolare la derivata della funzione $f(x) = a^x = \exp(x \log a)$:

$$f'(x) = \exp'(x \log a) \cdot \log a = \exp(x \log a) \cdot \log a = a^x \log a.$$

Per $x = 0$ ne consegue il limite notevole

$$\lim_{h \to 0} \frac{a^h - 1}{h} = \log a,$$

perché a sinistra c'è proprio il limite che definisce la derivata di $x \mapsto a^x$ in 0. Daremo più avanti un altro modo di calcolare numericamente il logaritmo di un numero reale positivo.

Esempio 5.14. Una volta definito l'esponenziale in base qualunque, possiamo, per ogni $s \in \mathbf{R}$, definire la funzione

$$f(x) = x^s,$$

che ha come dominio $(0, +\infty)$. Abbiamo cioè

$$f(x) = \exp(s \log x).$$

È facile allora calcolare la derivata di f:

$$f'(x) = \exp'(s \log x) \cdot s \log' x = \exp(s \log x) \cdot \frac{s}{x} = \frac{s x^s}{x} = s x^{s-1},$$

estendendo così la validità della nota formula di derivazione di x^n ad esponenti reali qualunque. Notiamo che però la funzione ad esponente reale $x \mapsto x^s$ è definita *solo* per $x > 0$, mentre $x \mapsto x^n$, per $n \in \mathbf{N}$, è definita ovunque.

Dalla definizione di derivata, abbiamo che, per $a > 0$,

$$\frac{1}{a} = \lim_{x \to a} \frac{\log x - \log a}{x - a}.$$

In particolare, prendendo $x = a + (1/n)$, con $n > 0$,

$$\frac{1}{a} = \lim_{n \to +\infty} \frac{\log\left(a + \dfrac{1}{n}\right) - \log a}{\dfrac{1}{n}} = \lim_{n \to +\infty} n \log\left(1 + \frac{1}{na}\right) = \lim_{n \to +\infty} \log\left(1 + \frac{1}{na}\right)^n.$$

Ponendo $1/a = z$ e l'esponenziale naturale di ambo i membri, otteniamo allora

$$\exp z = \lim_{n \to +\infty} \left(1 + \frac{z}{n}\right)^n$$

e, per $z = 1$, la nota formula

$$e = \lim_{n \to +\infty} \left(1 + \frac{1}{n}\right)^n.$$

Abbiamo tuttavia dato due definizioni del numero e: occorre dimostrare che sono equivalenti.

Consideriamo lo sviluppo, per $x \geq 0$,

$$\begin{aligned}\left(1+\frac{x}{n}\right)^n &= 1+\binom{n}{1}\frac{x}{n}+\binom{n}{2}\frac{x^2}{n^2}+\cdots+\binom{n}{n-1}\frac{x^{n-1}}{n-1}+\binom{n}{n}\frac{x^n}{n^n}\\ &= 1+\frac{x}{1!}+\frac{x^2}{2!}\left(1-\frac{1}{n}\right)+\cdots+\frac{x^n}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-2}{n}\right)\left(1-\frac{n-1}{n}\right)\\ &\leq 1+\frac{x}{1!}+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}.\end{aligned}$$

Possiamo dimostrare come in precedenza che

$$f_n(x)=\sum_{k=0}^{n}\frac{x^k}{k!}<3^x,$$

quindi che la successione $n\mapsto f_n(x)$ è convergente. Possiamo allora definire

$$f(x)=\lim_{n\to+\infty} f_n(x)$$

e non è troppo difficile dimostrare che vale

$$f'(x)=\lim_{n\to+\infty} f_n'(x);$$

Inoltre, per costruzione,

$$e^x=\lim_{n\to+\infty}\left(1+\frac{x}{n}\right)^n\leq\lim_{n\to+\infty} f_n(x)=f(x).$$

Ora, per $n>0$, $f_n'(x)=f_{n-1}(x)$, quindi $f'(x)=f(x)$, per ogni $x\geq 0$.

Osservazione 5.15. Analoghi calcoli mostrano che la stessa cosa vale per $x<0$: quindi la funzione f può essere definita per ogni $x\in\mathbf{R}$ e l'identità $f'(x)=f(x)$ vale ovunque.

Teorema 5.16. *Se $g\colon(a,b)\to\mathbf{R}$ è una funzione derivabile tale che, per ogni $x\in\mathbf{R}$ valga $g'(x)=g(x)$, allora esiste uno ed un solo $\alpha\in\mathbf{R}$ tale che, per ogni $x\in(a,b)$,*

$$g(x)=\alpha e^x.$$

Dimostrazione. Consideriamo la funzione $h(x)=g(x)e^{-x}$ e calcoliamone la derivata:

$$h'(x)=g'(x)e^{-x}+g(x)e^{-x}(-1)=g(x)e^{-x}-g(x)e^x=0.$$

Dunque la funzione h è costante su (a,b). □

Noi sappiamo che $e^x\leq f(x)$, per ogni $x\geq 0$ e che $f(x)=\alpha e^x$. Inoltre $1=f(0)=\alpha$, dunque $f(x)=e^x$.

Consideriamo ora la funzione

$$f(x)=\log\frac{1+x}{1-x},$$

definita in $(-1,1)$. Con facili calcoli, si vede che

$$f'(x)=\frac{2}{1-x^2}.$$

Sappiamo ora che

$$1+x^2+x^4+\cdots+x^{2n}=\frac{1-x^{2n+2}}{1-x^2}=\frac{1}{1-x^2}-\frac{x^{2n+2}}{1-x^2},$$

quindi

$$\frac{1}{1-x^2}=1+x^2+x^4+\cdots+x^{2n}+\frac{x^{2n+2}}{1-x^2},$$

da cui

$$\frac{2}{1-x^2} = 2 + 2x^2 + 2x^4 + \cdots + 2x^n + g_n(x),$$

dove

$$g_n(x) = \frac{2x^{2n+2}}{1-x^2}.$$

Abbiamo allora, integrando e ricordando che

$$\int_0^x 2t^m \, dt = \frac{2x^{m+1}}{m+1},$$

l'identità seguente:

$$\log \frac{1+x}{1-x} = \int_0^x \frac{-2t}{1-t^2} \, dt = 2x + 2\frac{x^3}{3} + 2\frac{x^5}{5} + \cdots + 2\frac{x^{2n+1}}{2n+1} + h_n(x),$$

dove, ancora,

$$h_n(x) = \int_0^x g_n(t) \, dt.$$

Il nostro scopo è di dimostrare che $\lim_{n\to+\infty} h_n(x) = 0$, così che la quantità

$$2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n+1}}{2n+1}\right)$$

fornisce un'approssimazione di $\log \frac{1+x}{1-x}$.

Abbiamo, per $x > 0$ e $0 \le t \le x$, che $1 - t^2 \ge 1 - x^2$. Se invece $x < 0$ e $x \le t \le 0$, ancora $1 - t^2 \ge 1 - x^2$. Perciò

$$|h_n(x)| = \left| \int_0^x g_n(t) \, dt \right| \le \int_0^x \left| \frac{2t^{2n+2}}{1-t^2} \right| dt \le \frac{1}{1-x^2} \int_0^x |2t^{2n+2}| \, dt = \frac{1}{1-x^2} \frac{|x^{2n+3}|}{2n+3} \le \frac{1}{(1-x^2)(2n+3)}$$

e l'ultima quantità tende a zero al crescere di n. Ad esempio, vogliamo calcolare $\log 3$. Si vede allora che occorre prendere $x = 1/2$; abbiamo

$$|h_n(1/2)| \le \frac{1}{2n+3} \frac{4}{3} \frac{1}{2^{2n+3}}$$

e, per ottenere un'approssimazione con quattro cifre decimali esatte, dobbiamo scegliere n in modo che l'espressione a destra sia minore di $0,00001$ e quindi occorre prendere $n = 6$. Allora il valore approssimato cercato è

$$2\left(\frac{1}{2} + \frac{1}{3 \cdot 2^3} + \frac{1}{5 \cdot 2^5} + \frac{1}{7 \cdot 2^7} + \frac{1}{9 \cdot 2^9} + \frac{1}{11 \cdot 2^{11}} + \frac{1}{13 \cdot 2^{13}}\right) = 1,0986\ldots$$

Questo sviluppo è molto efficiente; inoltre, per ogni $y \in \mathbf{R}$, $y > 0$, esiste un $x \in \mathbf{R}$, $-1 < x < 1$, tale che $y = \frac{1+x}{1-x}$.

Vale la pena di discutere brevemente la storia dei logaritmi. Fino al sedicesimo secolo i calcoli erano una delle attività che più impegnavano il tempo degli astronomi, distogliendoli dalle altre attività; il famoso astronomo Tycho Brahe (Knudstorp, Danimarca, 1546; Praga, 1601) fu il promotore dell'uso del metodo di *prostaferesi*, con il quale si trasformavano somme in prodotti e viceversa: le note formule di prostaferesi sono infatti:

$$\operatorname{sen}\alpha + \operatorname{sen}\beta = 2\operatorname{sen}\frac{\alpha+\beta}{2}\cos\frac{\alpha+\beta}{2}$$

e le altre analoghe. Abili riduzioni a questi casi permettevano di abbreviare i calcoli: Brahe ottenne varie misure molto precise usando la sua tecnica. Fu promotore anche di un terzo sistema del mondo, compromesso fra quello tolemaico e quello copernicano: nel sistema di Brahe, il sole e la luna girano intorno alla terra e i pianeti intorno al sole. Non occorre dire che questo sistema fu rifiutato da quasi

tutti; Brahe fu assunto al servizio dell'imperatore Rodolfo II ed è sepolto nella chiesa di Panna Marie před Týnem, la chiesa gotica nella Staroměstke Naměstí a Praga.

Nel 1614 venne pubblicato il trattato "Mirifici logarithmorum canonis descriptio", di John Napier (Edimburgo, 1550; 1617), barone di Merchiston, costato, secondo l'autore, vent'anni di lavoro. La prostaferesi poteva essere sostituita da una nuova tecnica, quella dei *numeri artificiali* o, appunto, logaritmi. L'idea fu data dal fatto già noto ad Archimede della additività degli esponenti: $a^{m+n} = a^m \cdot a^n$. Calcolando una base opportuna, Napier riuscì ad approssimare il calcolo di un prodotto con quello di una somma. Va detto subito che la teoria di Napier non è certo quella che conosciamo oggi, il modo di eseguire i calcoli era dettato dalla praticità di usare parti intere piuttosto grandi. Né Napier era consapevole di usare come base dei suoi logaritmi il numero e, che infatti prese il nome di numero di Eulero.

Di fatto Napier ebbe un predecessore, lo svizzero Jobst Bürgi (Lichtensteig, S. Gallo, 1552; Kassel, 1632) che pubblicò nel 1620 tavole analoghe a quelle di Napier, ma descritte in tedesco e quindi di scarsa circolazione fra gli studiosi: allora la scienza era scritta solamente in latino e Bürgi non lo conosceva.

La teoria dei logaritmi ebbe un contributo decisivo da Henry Briggs (Halifax, Inghilterra, 1556; Oxford, 1632), che ebbe tra le mani il trattato di Napier e giudicò l'idea originale e meravigliosa. Si mise in viaggio per raggiungere Napier e stette con lui un mese; l'anno successivo, il 1616 fece un'altra visita e non fece in tempo a compiere la terza per la morte dello scozzese.

Durante la prima visita Briggs suggerì di usare quelli che ora sono noti come logaritmi decimali, che resero molti calcoli veramente facili: già nel 1617 pubblicò le tavole *Logarithmorum chilias prima*, dove sono contenuti i logaritmi decimali dei numeri da 1 a 1000, con otto cifre decimali. Il lavoro di calcolo proseguì con la pubblicazione nel 1624 dell'*Arithmetica logarithmica*, dove si trovano i logaritmi decimali dei numeri da uno a 20000 e da 90000 a 100000 con *quattordici* cifre decimali, oltre alle regole di calcolo e ad esempi del loro uso.

Un detto comune all'epoca fu: "L'invenzione dei logaritmi ha raddoppiato la vita degli astronomi", ed era la verità. Eseguire moltiplicazioni di numeri grandi era un compito duro e non esente da errori; sostituire prodotti con somme rese il lavoro molto più spedito.

Naturalmente, al giorno d'oggi, l'importanza dei logaritmi è tutt'altra. La funzione logaritmo, con la sua inversa, compare in ogni ramo della matematica e della fisica, al pari delle funzioni trigonometriche. Non è più necessario consultare tavole ed eseguire calcoli, abbiamo potenti computer che lo fanno per noi. Ma le funzioni sono utili in moltissime applicazioni e hanno guadagnato la loro "dignità" di enti matematici, non solo di ausili al calcolo.

6. Serie

Fin dall'inizio dello sviluppo dell'Analisi, nel diciassettesimo secolo, ci si è imbattuti in procedimenti di somme infinite, provocando ampi dibattiti. È chiaro che la "somma di infiniti termini" non ha senso di per sé: si capì allora che questo richiede lo studio della convergenza di una particolare successione. Di fatto, lo studio delle serie o delle successioni è equivalente, come vedremo: tuttavia il formalismo delle serie è spesso più maneggevole.

Definizione 6.1. Se $s \in \mathscr{S}(\mathbf{R})$ è una successione, la *successione delle somme parziali* di s è la successione $\sigma(s)$ definita da

$$\sigma(s)\colon\ n \in \mathbf{N} \mapsto s_0 + s_1 + \cdots + s_n.$$

In altre parole, $\sigma(s)_0 = s_0$ e $\sigma(s)_{n+1} = \sigma(s)_n + s_{n+1}$.

Diremo che s è *convergente in serie* o, brevemente, che $\sum s$ esiste, se la successione $\sigma(s)$ converge ad un numero reale a; in tal caso scriveremo che

$$\sum s = a.$$

Esempio 6.2. Una delle prime serie di cui si seppe indicare la convergenza è quella di Mengoli:

$$s_n = \frac{1}{(n+1)(n+2)},$$

notando che

$$\sigma(s)_n = 1 - \frac{1}{n+2}.$$

La cosa è ovvia per $n = 0$. Supponiamo che valga per n; allora

$$\sigma(s)_{n+1} = \sigma(s)_n + s_{n+1} = 1 - \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} = 1 - \frac{1}{n+3},$$

come si voleva. Siccome questa successione converge ovviamente a 1, possiamo asserire che $\sum s = 1$.

Esempio 6.3. Poniamo $s_0 = 0$ e $s_n = 1/n$, per $n \geq 1$. Allora s non è convergente in serie. Infatti, se raggruppiamo opportunamente le somme parziali, abbiamo

$$\begin{aligned}
\sigma(s)_{2^k-1} &= 1 + \underbrace{\frac{1}{2} + \frac{1}{3}}_{2\text{ addendi}} + \underbrace{\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}}_{4\text{ addendi}} + \cdots + \underbrace{\frac{1}{2^{k-1}} + \frac{1}{2^{k-1}+1} + \cdots + \frac{1}{2^k - 1}}_{2^{k-1}\text{ addendi}} \\
&\geq 1 + \underbrace{\frac{1}{4} + \frac{1}{4}}_{2\text{ addendi}} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{4\text{ addendi}} + \cdots + \underbrace{\frac{1}{2^k} + \cdots + \frac{1}{2^k}}_{2^{k-1}\text{ addendi}} \\
&= 1 + \underbrace{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}}_{k-1\text{ addendi}} = \frac{k+1}{2}
\end{aligned}$$

e quindi la successione $\sigma(s)$ non è limitata, dunque non convergente. La serie di questa successione è detta *serie armonica*.

Esempio 6.4. Sia $c \in \mathbf{R}$ e consideriamo la successione *geometrica di ragione c* definita da $s\colon n \mapsto c^n$. Sappiamo già calcolare

$$\sigma(s)_n = \frac{1-c^{n+1}}{1-c}$$

per $c \neq 1$; per $c = 1$, ovviamente $\sigma(s)_n = n+1$. In particolare la successione non converge in serie per $c = 1$.

Se $|c| < 1$, abbiamo che $\lim_{n\to+\infty} c^n = 0$ e quindi

$$\sum s = \frac{1}{1-c}.$$

Se $|c| > 1$, il limite $\lim_{n\to+\infty} c^n$ non esiste in quanto la successione non è limitata (esercizio).

Dicevamo che lo studio delle successioni e delle serie è equivalente. Infatti, data la successione $s \in \mathscr{S}(\mathbf{R})$, consideriamo la successione $\tilde{s}$ definita da

$$\tilde{s}_0 = s_0, \qquad \tilde{s}_n = s_n - s_{n-1} \ (n > 0).$$

Allora $\sigma(\tilde{s}) = s$; quindi s converge se e solo se $\tilde{s}$ converge in serie.

Diamo un primo importante criterio di convergenza in serie: la condizione sarà solo necessaria, come mostra l'esempio della serie armonica.

Proposizione 6.5. *Se s converge in serie, allora s converge a* 0.

Dimostrazione. Fissiamo $\varepsilon > 0$; allora, dal fatto che $\sigma(s)$ è di Cauchy, segue che esiste n_0 tale che, per ogni $m, n \geq n_0$, $|\sigma(s)_m - \sigma(s)_n| < \varepsilon$. In particolare, per $n \geq n_0 + 1$,

$$|s_n| = |\sigma(s)_n - \sigma(s)_{n-1}| < \varepsilon,$$

cioè la tesi. □

Per successioni a termini non negativi, abbiamo un fondamentale criterio di convergenza in serie: infatti la successione delle somme parziali è crescente.

Proposizione 6.6. *Se $s \in \mathscr{S}(\mathbf{R})$ e $s_n \geq 0$, per ogni $n \in \mathbf{N}$, allora s converge in serie se e solo se $\sigma(s)$ è limitata.*

Corollario 6.7. *Siano $s, t \in \mathscr{S}(\mathbf{R})$ tali che, per ogni $n \in \mathbf{N}$, $0 \leq s_n \leq t_n$. Se t converge in serie, allora s converge in serie.*

Dimostrazione. È ovvio che $\sigma(s)_n \leq \sigma(t)_n \leq \sum t$. □

La condizione che $s_n \leq t_n$, per ogni $n \in \mathbf{N}$ può essere indebolita: la convergenza in serie non viene a cadere se un numero finito di termini di s è modificato.

Proposizione 6.8. *Se $s, t \in \mathscr{S}(\mathbf{R})$, s converge in serie e $t - s$ è definitivamente nulla, allora t converge in serie.*

Dimostrazione. Esercizio; si trovi un modo di esprimere $\sum t$. □

Se definiamo $\tau_k(s)$ con $\tau_k(s)\colon n \mapsto s_{n+k}$, dove $k \geq 0$, abbiamo che, se s converge in serie, allora $\tau_k(s)$ converge in serie e, per $k > 0$,

$$\sum \tau_k(s) = \sum s - \sigma(s)_{k-1}$$

(esercizio). Abbiamo già usato implicitamente questo fatto.

Se $s \in \mathscr{S}(\mathbf{R})$ e $a \in \mathbf{R}$, possiamo definire $as\colon n \mapsto as_n$. Abbiamo già definito la somma di due successioni.

Proposizione 6.9. *Se $s, t \in \mathscr{S}(\mathbf{R})$ convergono in serie e $a \in \mathbf{R}$, allora*

$$\sum as = a\sum s, \quad \sum(s+t) = \sum s + \sum t.$$

Dimostrazione. Esercizio. □

Esiste un altro criterio assai utile; data $s \in \mathscr{S}(\mathbf{R})$, la successione $|s|$ è definita da $n \mapsto |s_n|$.

Proposizione 6.10. *Se $s \in \mathscr{S}(\mathbf{R})$ e $|s|$ converge in serie, allora s converge in serie.*

Dimostrazione. Poniamo, per $n \in \mathbf{N}$,

$$s_n^+ = \begin{cases} s_n & \text{se } s_n \geq 0, \\ 0 & \text{se } s_n < 0, \end{cases} \qquad s_n^- = \begin{cases} 0 & \text{se } s_n \geq 0, \\ -s_n & \text{se } s_n < 0. \end{cases}$$

Le successioni s^+ e s^- sono a termini non negativi ed è ovvio che, per ogni $n \in \mathbf{N}$,

$$s_n^+ \leq |s_n|, \qquad s_n^- \leq |s_n|,$$

quindi entrambe convergono in serie. Ma allora $s^+ - s^-$ converge in serie a $\sum s^+ - \sum s^-$ e, per ogni $n \in \mathbf{N}$,

$$\sigma(s^+ - s^-)_n = \sigma(s)_n,$$

e quindi s converge in serie, proprio a $\sum s^+ - \sum s^-$. □

Esiste anche un criterio per le *successioni a segni alterni*: una successione s è così detta se $s_{2n} \geq 0$ e $s_{2n+1} \leq 0$.

Proposizione 6.11. *Se $s \in \mathscr{S}(\mathbf{R})$ è una successione a segni alterni, s converge a zero e $|s|$ è una successione decrescente, allora s converge in serie. Inoltre, per ogni $n \in \mathbf{N}$, abbiamo*

$$\left|\sigma(s)_n - \sum s\right| \leq |s_{n+1}|.$$

Dimostrazione. Abbiamo facilmente

$$\sigma(s)_0 \geq \sigma(s)_2 \geq \cdots \geq \sigma(s)_{2k} \geq \sigma(s)_{2k+2} \geq \cdots$$
$$\sigma(s)_1 \leq \sigma(s)_3 \leq \cdots \leq \sigma(s)_{2k+1} \leq \sigma(s)_{2k+3} \leq \cdots$$

e, inoltre, $s_{2k+1} = \sigma(s)_{2k+1} - \sigma(s)_{2k} \leq 0$. Perciò le successioni $t'_n = \sigma(s)_{2n}$ e $t''_n = \sigma(s)_{2n+1}$ sono una decrescente e l'altra crescente, ma entrambe limitate, quindi convergono, rispettivamente a a' e a'', con $a' \geq a''$. Ci basta allora verificare che $a' = a''$. Supponiamo per assurdo che $a' > a''$ e sia $\varepsilon = (a' - a'')/2$: esiste n_0 tale che, per $n \geq n_0$, $|s_n| < \varepsilon$. Ma

$$|s_{2n+1}| = -s_{2n+1} = t'_n - t''_n \geq a' - a'' :$$

assurdo. Il resto della dimostrazione si lascia come esercizio. □

Esempio 6.12. La successione s definita da $s_n = (-1)^n/(n+1)$ soddisfa le ipotesi della proposizione precedente, quindi converge in serie; si può dimostrare che $\sum s = \log 2$. Notiamo che allora una successione s può convergere in serie senza che $|s|$ converga in serie.

Definizione 6.13. Una successione $s \in \mathscr{S}(\mathbf{R})$ si dice *assolutamente convergente in serie* se $|s|$ è convergente in serie.

Osservazione 6.14. Una successione che converge assolutamente in serie converge in serie; non è vero il viceversa. Abel dimostrò che, se una successione s converge in serie ma non assolutamente, allora, fissato un qualunque numero reale r, esiste una permutazione dei numeri naturali tali che la successione "permutata" converge in serie a r. Più precisamente, esiste un'applicazione biiettiva $\varphi\colon \mathbf{N} \to \mathbf{N}$ tale che, posto $t_n = s_{\varphi(n)}$,

$$\sum t = r.$$

Questo non può avvenire per le successioni assolutamente convergenti in serie, anche se la dimostrazione non è molto semplice.

Un conveniente criterio per la convergenza assoluta in serie è quello del *rapporto*.

Proposizione 6.15. *Sia $s \in \mathscr{S}(\mathbf{R})$. Se esistono $n_0 \in \mathbf{N}$ e $c \in \mathbf{R}$, $0 < c < 1$, tali che, per ogni $n \geq n_0$ si abbia*

$$|s_{n+1}| \leq c|s_n|,$$

allora s converge assolutamente in serie.

Dimostrazione. Non è restrittivo supporre $s_n > 0$, per ogni n: infatti, possiamo porre $s_n = 1$, per $n < n_0$, che non modifica la convergenza in serie. Se poi $s_m = 0$, per qualche m, la condizione imposta dice che $s_{m+1} = s_{m+2} = \cdots = 0$, quindi la successione è ovviamente convergente in serie.

Abbiamo allora: $s_1 \leq cs_0$, $s_2 \leq cs_1 \leq c^2 s_0$, e così via. La successione geometrica $g\colon n \mapsto c^n$ converge in serie, come abbiamo già osservato, quindi $s_0 g$ converge in serie e anche s converge in serie. □

Osservazione 6.16. È abbastanza facile verificare che, se esistono $n_0 \in \mathbf{N}$ e $c \in \mathbf{R}$, $c > 1$, tali che, per ogni $n \geq n_0$, si abbia

$$|s_{n+1}| \geq c|s_n|$$

allora, escluso il caso banale in cui la successione s è definitivamente nulla, la successione s non converge in serie (esercizio).

Il risultato precedente viene spesso usato calcolando un certo limite; la condizione $s_n > 0$ non è molto restrittiva: come sappiamo possiamo modificare un numero finito di termini di una successione senza alterare l'eventuale convergenza o non convergenza.

Corollario 6.17. *Sia $s \in \mathscr{S}(\mathbf{R})$ con $s_n > 0$, per ogni $n \in \mathbf{N}$, e supponiamo che*

$$\lim_{n\to+\infty} \frac{s_{n+1}}{s_n} = c.$$

Se $c < 1$ allora s converge in serie; se $c > 1$, allora s non converge in serie.

Dimostrazione. Supponiamo $c < 1$ e sia $\varepsilon = (1-c)/2$; allora esiste n_0 tale che, per $n \geq n_0$,

$$\frac{s_{n+1}}{s_n} - c < \varepsilon, \quad \text{cioè} \quad s_{n+1} < (c+\varepsilon)s_n$$

e possiamo applicare il criterio precedente in quanto $c + \varepsilon < 1$.

Nel caso $c > 1$, è evidente come prima che s non converge a zero, quindi non converge in serie. □

Osservazione 6.18. Se, nelle ipotesi precedenti, $\lim_{n\to+\infty} \frac{s_{n+1}}{s_n} = 1$, il criterio non dà alcuna informazione; esistono infatti successioni con queste proprietà che convergono e altre che non convergono.

Esempio 6.19. La successione $s\colon n \mapsto 1/(n+1)^2$ converge in serie; tuttavia

$$\lim_{n\to+\infty} \frac{s_{n+1}}{s_n} = \lim_{n\to+\infty} \frac{(n+2)^2}{(n+1)^2} = 1.$$

La convergenza in serie può essere così dimostrata:

$$\sigma(s)_{2^k-1} = 1 + \underbrace{\frac{1}{4} + \frac{1}{9}}_{\text{2 addendi}} + \underbrace{\frac{1}{16} + \frac{1}{25} + \frac{1}{36} + \frac{1}{49}}_{\text{4 addendi}} + \cdots + \underbrace{\frac{1}{(2^{k-1})^2} + \frac{1}{(2^{k-1}+1)^2} + \cdots + \frac{1}{(2^k-1)^2}}_{2^{k-1}\text{ addendi}}$$

$$\leq 1 + \underbrace{\frac{1}{16} + \frac{1}{16}}_{\text{2 addendi}} + \underbrace{\frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64}}_{\text{4 addendi}} + \cdots < 2$$

Si può dimostrare che $\sum s = \pi^2/6$.

7. Serie di potenze

Abbiamo già dimostrato il fatto seguente: se $-1 < x < 1$, allora la successione $s(x)\colon n \mapsto x^n$ converge in serie, quindi possiamo definire una funzione $f\colon (-1,1) \to \mathbf{R}$ con $f(x) = \sum s(x)$.

Definizione 7.1. Data una successione $s \in \mathscr{S}(R)$, la *serie di potenze* indotta da s è la famiglia di successioni

$$\hat{s}(x)\colon n \mapsto s_n x^n,$$

per ogni $x \in \mathbf{R}$; la serie di potenze indotta da s sarà denotata con $\hat{s}$. Diremo *insieme di convergenza* di $\hat{s}$ l'insieme dei numeri reali x tali che $\hat{s}(x)$ converge in serie.

Proposizione 7.2. *Se esiste $r \in \mathbf{R}$, $r > 0$, tale che la successione $\hat{s}(r)$ converge assolutamente in serie, allora la successione $\hat{s}(x)$ converge assolutamente in serie, per ogni $x \in \mathbf{R}$ con $|x| < r$.*

Dimostrazione. È chiaro che, per $|x| \le r$, abbiamo, per ogni n,

$$|s_n x^n| \le |s_n r^n|,$$

quindi la tesi. □

Consideriamo allora l'insieme X dei valori r tali che $\hat{s}(r)$ converge assolutamente in serie: se X non è limitato, abbiamo dalla proposizione precedente che $\hat{s}(x)$ converge assolutamente in serie, per ogni $x \in \mathbf{R}$. Se invece X è limitato e ρ è il suo estremo superiore, abbiamo che $\hat{s}(x)$ converge assolutamente in serie per ogni x con $|x| < \rho$. Nel primo caso diremo che $\hat{s}$ ha raggio di convergenza infinito, nel secondo caso che $\hat{s}$ ha raggio di convergenza ρ. In questo caso è facile vedere (esercizio) che $\hat{s}(x)$ non converge in serie, se $|x| > \rho$.

L'insieme di convergenza di una serie di potenze è allora un intervallo con centro nell'origine; gli estremi possono appartenere o no all'insieme di convergenza. Nei punti interni dell'insieme di convergenza la convergenza è assoluta.

Esempio 7.3. Sia $s_n = 1$, per ogni n. Allora l'insieme di convergenza di $\hat{s}$ è $(-1,1)$.

Sia $s_0 = 0$, $s_n = 1/n$ $(n > 0)$; l'insieme di convergenza di $\hat{s}$ è $[-1,1)$.

Abbiamo già un esempio di funzione *sviluppabile in serie di potenze*: poniamo $s_n = 0$ se n è pari e $s_n = 2/n$ se n è dispari. Allora, per $-1 < x < 1$, abbiamo

$$\sum \hat{s}(x) = \log \frac{1+x}{1-x}.$$

Useremo d'ora in poi anche la notazione tradizionale, scrivendo ad esempio, in questo caso,

$$\log \frac{1+x}{1-x} = 2x + \frac{2x^3}{3} + \frac{2x^5}{5} + \cdots + \frac{2x^{2n+1}}{2n+1} + \cdots = \sum_{n=0}^{+\infty} \frac{2x^{2n+1}}{2n+1}.$$

Esempio 7.4. Le serie di potenze

$$\sum_{n=0}^{+\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!} \qquad \text{e} \qquad \sum_{n=0}^{+\infty}(-1)^n\frac{x^{2n}}{(2n)!}$$

hanno raggio di convergenza infinito. Infatti, per la prima, il criterio del rapporto dà:

$$\frac{|x^{2n+3}|}{(2n+3)!}\frac{(2n+1)!}{|x^{2n+1}|}=\frac{|x^2|}{(2n+2)(2n+3)}$$

il cui limite è zero, per ogni x. Analogamente per la seconda. Lo stesso si può dire della serie, che conosciamo già,

$$\sum_{n=0}^{+\infty}\frac{x^n}{n!}.$$

Porremo

$$\mathsf{S}(x)=\sum_{n=0}^{+\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!},\quad \mathsf{C}(x)=\sum_{n=0}^{+\infty}(-1)^n\frac{x^{2n}}{(2n)!},\quad \mathsf{E}(x)=\sum_{n=0}^{+\infty}\frac{x^n}{n!}.$$

Esistono naturalmente casi in cui il raggio di convergenza è 0: ad esempio, basta prendere $s_n = n!$ e verificare che $\hat{s}(x)$ converge solo per $x = 0$.

Le funzioni definite tramite serie di potenze con raggio di convergenza positivo hanno proprietà molto buone.

Proposizione 7.5. *Se $s \in \mathscr{S}(\mathbf{R})$ e il raggio di convergenza di $\hat{s}$ è $\rho > 0$ (o infinito), allora la serie di potenze definita dalla successione $s'\colon n \mapsto (n+1)s_{n+1}$ ha lo stesso raggio di convergenza.*

Dimostrazione. Supponiamo che la serie di potenze

$$\sum_{n=0}^{+\infty}s_nx^n$$

converga assolutamente per $x = r$, con $0 < r$. Consideriamo ora la nuova successione $t\colon n \mapsto ns_n$ e dimostriamo che la serie di potenze

$$\sum_{n=0}^{+\infty}t_nx^n$$

converge assolutamente per ogni x con $0 < |x| < r$. Sia infatti $0 < |x| < c < r$; affermiamo che esiste n_0 tale che $|ns_nx^n| < |s_nc^n|$: una volta dimostrato questo, abbiamo che la serie $\sum_n s_nc^n$ converge assolutamente e quindi anche la serie $\sum_n t_nx^n$ converge assolutamente.

Ma $\lim_{n\to+\infty} n^{1/n} = 1$, come si vede prendendo il logaritmo e provando che $\lim_{n\to+\infty}(\log n)/n = 0$. Questo si vede applicando il teorema di L'Hôpital alla funzione $g(x) = (\log x)/x$.

Ne segue che esiste n_0 tale che, per $n \geq n_0$, $n^{1/n}|x| < c$. Ma allora $|ns_nx^n| = |s_n(n^{1/n}x)^n| < |s_nc^n|$, come si voleva.

Ora, per $x \neq 0$, abbiamo

$$\sum\hat{s}'(x)=\frac{1}{x}\sum\hat{t}(x),$$

come si desiderava. La convergenza in 0 è ovvia. □

La definizione di $\hat{s}'$ è forse più chiara scrivendola esplicitamente come serie di potenze:

$$\sum\hat{s}'(x)=\sum_{n=1}^{+\infty}ns_nx^{n-1}.$$

In altre parole $\hat{s}'$ si ottiene derivando termine a termine la serie $\hat{s}$. Ci aspettiamo perciò che in ogni punto x interno all'intervallo di convergenza, $\sum\hat{s}'(x)$ sia la derivata della funzione $x \mapsto \sum\hat{s}(x)$.

Proposizione 7.6. *Sia $s \in \mathscr{S}(\mathbf{R})$ tale che il raggio di convergenza di $\hat{s}$ sia $\rho > 0$ (o infinito). Definiamo*

$$f(x) = \sum_{n=0}^{+\infty} s_n x^n$$

per x interno all'intervallo di convergenza. Allora f è derivabile e, per ogni x interno all'intervallo di convergenza, abbiamo

$$f'(x) = \sum_{n=1}^{+\infty} n s_n x^{n-1}.$$

Dimostrazione. Sia $r > 0$ tale che la serie di potenze converga per $x = r$. Prendiamo b tale che $0 < b < r$ e $\delta > 0$ tale che $b + \delta < r$. Considereremo solo valori x tali che $|x| < b$ e valori di h tali che $0 < |h| < \delta$. Abbiamo, con un facile calcolo,

$$f(x+h) - f(x) = \sum_{n=0}^{+\infty} s_n((x+h)^n - x^n)$$

e, per il teorema del valor medio, per ogni n, esiste x_n tra x e $x+h$ tale che

$$(x+h)^n - x^n = n x_n^{n-1} h,$$

quindi

$$\frac{f(x+h) - f(x)}{h} = \sum_{n=1}^{+\infty} n s_n x_n^{n-1}$$

(questa è una serie numerica, non una serie di potenze). Abbiamo allora anche

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} - \sum_{n=1}^{+\infty} n s_n x^{n-1} &= \sum_{n=1}^{+\infty} n s_n x_n^{n-1} - \sum_{n=1}^{+\infty} n s_n x^{n-1} \\ &= \sum_{n=1}^{+\infty} n s_n (x_n^{n-1} - x^{n-1}). \end{aligned}$$

Ancora per il teorema del valor medio, abbiamo y_n tra x e x_n ($n \geq 2$) tali che

$$x_n^{n-1} - x^{n-1} = (n-1) y_n^{n-2} (x_n - x)$$

e quindi

$$\frac{f(x+h) - f(x)}{h} - \sum_{n=1}^{+\infty} n s_n x^{n-1} = \sum_{n=2}^{+\infty} n(n-1) s_n y_n^{n-2} (x_n - x).$$

Ne segue che, essendo $|y_n| < b + \delta < r$ e $|x_n - x| < |h|$,

$$\begin{aligned} \left| \frac{f(x+h) - f(x)}{h} - \sum_{n=1}^{+\infty} n s_n x^{n-1} \right| &\leq \sum_{n=2}^{+\infty} n(n-1) |s_n| |y_n|^{n-2} |h| \\ &\leq |h| \sum_{n=2}^{+\infty} n(n-1) |s_n| (b+\delta)^{n-2}. \end{aligned}$$

Applicando la proposizione precedente due volte, otteniamo che l'ultima serie converge, diciamo ad a. Dunque

$$\left| \frac{f(x+h) - f(x)}{h} - \sum_{n=1}^{+\infty} n s_n x^{n-1} \right| \leq |h| a$$

e, prendendo il limite per h tendente a zero, abbiamo che il termine di sinistra tende a zero, cioè la tesi. □

Indichiamo con $f^{(k)}$ la derivata k-esima della funzione f (se esiste); in particolare $f^{(0)} = f$ e $f^{(1)} = f'$.

Definizione 7.7. Una funzione si dice *sviluppabile in serie di potenze attorno a zero tramite s* se esiste una successione $s \in \mathscr{S}(\mathbf{R})$ tale che il raggio di convergenza di $\hat{s}$ sia $\rho > 0$ (o infinito) e che

$$f(x) = \sum_{n=0}^{+\infty} s_n x^n$$

per x interno all'intervallo di convergenza. Diremo che s è una successione di coefficienti di Taylor per f attorno a zero.

Corollario 7.8. *Sia f una funzione sviluppabile in serie di potenze attorno a zero tramite s. Allora f è derivabile infinite volte e, per ogni x interno all'intervallo di convergenza ed ogni $k \in \mathbf{N}$, abbiamo*

$$f^{(k)}(x) = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} s_n x^{n-k}.$$

Dimostrazione. Non c'è che da verificare la formula: poniamo $g = f^{(k)}$; allora

$$f^{(k+1)}(x) = g'(x) = \sum_{n=k+1}^{+\infty} \frac{n!}{(n-k)!}(n-k) s_n x^{n-k-1}$$

ed è evidente che

$$\frac{n!}{(n-k)!}(n-k) = \frac{n!}{(n-k-1)!},$$

da cui la tesi. □

Corollario 7.9. *Sia f una funzione sviluppabile in serie di potenze attorno a zero tramite s. Allora, per ogni $n \in \mathbf{N}$,*

$$s_n = \frac{f^{(n)}(0)}{n!}.$$

Dimostrazione. Basta calcolare $f^{(n)}(0)$ con il corollario precedente. □

Esempio 7.10. Consideriamo la funzione $f : \mathbf{R} \to \mathbf{R}$ definita al modo seguente:

$$f(x) = \begin{cases} 0 & \text{se } x = 0, \\ \exp(-1/x^2) & \text{se } x \neq 0. \end{cases}$$

È immediato verificare che f è continua; anzi, f è infinitamente derivabile, per $x \neq 0$: infatti esistono polinomi P_n di grado $< 3n$ tali che, per $n \geq 1$,

$$f^{(n)}(x) = \frac{P_n(x)}{x^{3n}} \exp(-1/x^2).$$

Lo dimostriamo per induzione su n: la cosa è banalmente vera per $n = 1$, con $P_n = -2$. Calcoliamo ora la derivata di $f^{(n)}$, tenendo conto dell'ipotesi induttiva: abbiamo

$$f^{(n+1)}(x) = \frac{x^3 P_n'(x) - x^2 P_n(x) - 2P_n(x)}{x^{3n+3}} \exp(-1/x^2)$$

come si desiderava, perché il grado del numeratore è minore di $3n+3$. Non è allora difficile verificare che, per ogni n,

$$\lim_{x \to 0} f^{(n)}(x) = 0,$$

il che comporta che la funzione f è derivabile infinite volte anche in 0, e che ogni derivata successiva vale 0. Perciò f non è sviluppabile in serie di potenze attorno a zero.

Come applicazione delle formule trovate, possiamo verificare che, per ogni $x \in \mathbf{R}$,

$$\mathsf{E}'(x) = \mathsf{E}(x), \qquad \mathsf{S}'(x) = \mathsf{C}(x), \qquad \mathsf{C}'(x) = -\mathsf{S}(x).$$

Poiché $\mathsf{E}(0) = 1$, abbiamo che, per ogni x, $\mathsf{E}(x) = e^x$. Dunque abbiamo trovato lo sviluppo in serie di potenze della funzione esponenziale naturale.

8. Seno e coseno

La definizione e la dimostrazione delle proprietà analitiche delle funzioni trigonometriche sono piuttosto lacunose, facendo riferimento all'intuizione di fatti poco chiari: si pensi all'usuale dimostrazione che

$$\lim_{x\to 0}\frac{\operatorname{sen} x}{x}=1,$$

in cui si dà per nota la definizione di lunghezza di un arco di curva. Ora, nella sezione precedente abbiamo dimostrato che esistono due funzioni S e C che hanno proprietà analitiche simili a quelle di seno e coseno: $\mathsf{S}'=\mathsf{C}$, $\mathsf{C}'=\mathsf{S}$, $\mathsf{S}(0)=0$ e $\mathsf{C}(0)=1$.

Supponiamo ora di avere due funzioni $f\colon \mathbf{R}\to\mathbf{R}$ e $g\colon \mathbf{R}\to\mathbf{R}$ ovunque derivabili e tali che

$$f'=g,\quad g'=-f,\quad f(0)=0,\quad g(0)=1.$$

Lemma 8.1. *Per ogni $x\in\mathbf{R}$ vale la relazione*

$$f(x)^2+g(x)^2=1.$$

Dimostrazione. Poniamo $h(x)=f(x)^2+g(x)^2$. Allora $h(0)=1$ e

$$h'=2ff'+2gg'=2fg-2gf=0.$$

Ne segue che h è costante e vale ovunque 1. □

Lemma 8.2. *Se $f_1\colon \mathbf{R}\to\mathbf{R}$ e $g_1\colon \mathbf{R}\to\mathbf{R}$ sono funzioni ovunque derivabili tali che*

$$f_1'=g\qquad e\qquad g_1'=-f,$$

allora esistono e sono unici due numeri reali a e b tali che, per ogni $x\in\mathbf{R}$,

$$(*)\qquad \begin{aligned} f_1(x)&=af(x)-bg(x),\\ g_1(x)&=bf(x)+ag(x).\end{aligned}$$

Dimostrazione. Ponendo $h_1(x)=f(x)f_1(x)+g(x)g_1(x)$ e $h_2(x)=f(x)g_1(x)-f_1(x)g(x)$, possiamo derivare e verificare che $h_1'(x)=0$ e $h_2'(x)=0$, da cui discende che esistono $a,b\in\mathbf{R}$ tali che, per ogni $x\in\mathbf{R}$,

$$\begin{aligned} f(x)f_1(x)+g(x)g_1(x)&=a,\\ f(x)g_1(x)-f_1(x)g(x)&=b.\end{aligned}$$

Moltiplichiamo la prima uguaglianza per $f(x)$ e la seconda per $g(x)$, sottraiamo e ricordiamo il lemma precedente: otteniamo

$$f_1(x)=af(x)-bg(x).$$

Moltiplichiamo la prima uguaglianza per $g(x)$ e la seconda per $f(x)$, sommiamo e ricordiamo il lemma precedente: otteniamo

$$g_1(x)=ag(x)+bf(x),$$

come richiesto. □

Proposizione 8.3. *Esistono e sono uniche due funzioni* $\mathsf{S}\colon \mathbf{R} \to \mathbf{R}$ *e* $\mathsf{C}\colon \mathbf{R} \to \mathbf{R}$ *ovunque derivabili e tali che*

$$\mathsf{S}' = \mathsf{C}, \qquad \mathsf{C}' = \mathsf{S}, \qquad \mathsf{S}(0) = 0, \qquad \mathsf{C}(0) = 1.$$

Dimostrazione. L'esistenza è già stata dimostrata. Per l'unicità possiamo usare il lemma precedente: se $f = \mathsf{S}$, e $g = \mathsf{C}$ e supponiamo che f_1, g_1 siano funzioni soddisfacenti le nostre richieste, abbiamo necessariamente $a = 1$ e $b = 0$ nella $(*)$. □

Le relazioni $(*)$ possono essere usate anche per dimostrare le *formule di addizione*.

Proposizione 8.4. *Se* $x, y \in \mathbf{R}$*, abbiamo*

$$\mathsf{S}(x+y) = \mathsf{S}(x)\,\mathsf{C}(y) + \mathsf{C}(x)\,\mathsf{S}(y),$$
$$\mathsf{C}(x+y) = \mathsf{C}(x)\,\mathsf{C}(y) - \mathsf{S}(x)\,\mathsf{S}(y).$$

Dimostrazione. In quello che segue y è fissato. Poniamo $f(x) = \mathsf{S}(x)$, $g(x) = \mathsf{C}(x)$, $f_1(x) = \mathsf{S}(x+y)$ e $g_1(x) = \mathsf{C}(x+y)$. Per la regole della derivazione, è evidente che

$$f_1'(x) = \mathsf{C}(x+y) = g_1(x) \qquad \text{e} \qquad g_1'(x) = -\mathsf{S}(x+y) = -f_1(x),$$

quindi esistono $a, b \in \mathbf{R}$ come in $(*)$. È facile vedere, ponendo $x = 0$, che $a = \mathsf{C}(y)$ e $b = -\mathsf{S}(y)$, ciò che completa la dimostrazione. □

Corollario 8.5. *Se* $x \in \mathbf{R}$ *abbiamo*

$$\mathsf{S}(2x) = 2\,\mathsf{S}(x)\,\mathsf{C}(x) \qquad e \qquad \mathsf{C}(2x) = \mathsf{C}(x)^2 - \mathsf{S}(x)^2 = 2\,\mathsf{C}(x)^2 - 1$$

Le funzioni S e C sono una dispari e l'altra pari.

Proposizione 8.6. *Per ogni* $x \in \mathbf{R}$ *abbiamo*

$$\mathsf{S}(-x) = -\mathsf{S}(x) \qquad e \qquad \mathsf{C}(-x) = \mathsf{C}(x).$$

Dimostrazione. Poniamo $f(x) = \mathsf{S}(x)$, $g(x) = \mathsf{C}(x)$, $f_1(x) = \mathsf{C}(-x)$ e $g_1(x) = \mathsf{S}(-x)$. Per la regole della derivazione, è evidente che

$$f_1'(x) = \mathsf{S}(-x) = g_1(x) \qquad \text{e} \qquad g_1'(x) = -\mathsf{C}(-x) = -f_1(x),$$

quindi esistono $a, b \in \mathbf{R}$ come in $(*)$. È facile vedere, ponendo $x = 0$, che $a = 0$ e $b = -1$, ciò che completa la dimostrazione. □

Vogliamo ora dimostrare il fatto forse più importante su queste funzioni S e C.

Lemma 8.7. *Esiste* $c \in \mathbf{R}$ *tale che* $\mathsf{C}(c) = 0$.

Dimostrazione. Supponiamo che $\mathsf{C}(x) \neq 0$, per ogni $x \in \mathbf{R}$. Poiché sappiamo che $\mathsf{C}(0) = 1 > 0$ e che C è ovunque derivabile, quindi continua, otteniamo che $\mathsf{C}(x) > 0$, per ogni $x \in \mathbf{R}$. Di conseguenza $\mathsf{S}'(x) > 0$, per ogni x e perciò S è crescente; in particolare $\mathsf{S}(x) > 0$ se $x > 0$. Inoltre, dalla relazione $\mathsf{S}(x)^2 + \mathsf{C}(x)^2 = 1$, otteniamo

$$\mathsf{C}(x) = \sqrt{1 - \mathsf{S}(x)^2},$$

per ogni $x \in \mathbf{R}$, e quindi che C è decrescente in $[0, +\infty)$. Sia $d > 0$: allora

$$0 < \mathsf{C}(2d) = \mathsf{C}(d)^2 - \mathsf{S}(d)^2 < \mathsf{C}(d)^2.$$

Facendo induzione, per ogni $n \in \mathbf{N}$,

$$0 < \mathsf{C}(2^n d) < (\mathsf{C}(d))^{2n}$$

e, dal fatto che $\mathsf{C}(d) < 1$, segue che, per il teorema del confronto di limiti,

$$\lim_{n \to +\infty} \mathsf{C}(2^n d) = 0.$$

Di conseguenza, anche

$$\lim_{n\to+\infty} \mathsf{S}(2^n d) = 1.$$

Perciò possiamo trovare $\bar{d} > 0$ tale che

$$\mathsf{C}(\bar{d}) < \frac{1}{2}.$$

Abbiamo trovato allora la contraddizione: infatti

$$\mathsf{C}(2\bar{d}) = 2\,\mathsf{C}(\bar{d})^2 - 1 < 0,$$

contro l'ipotesi che $\mathsf{C}(\bar{d}) > 0$. □

Un risultato sulle funzioni continue che viene di rado menzionato è molto utile. Nell'enunciato seguente intendiamo che anche una semiretta chiusa o l'intera retta reale sono intervalli chiusi.

Teorema 8.8. *Siano $f_1, f_2, \dots, f_n$ funzioni continue su un intervallo chiuso I e sia $a \in I$. Se esiste $x_0 > a$ tale che*

$$f_1(x_0) = f_2(x_0) = \cdots = f_n(x_0) = 0,$$

e poniamo

$$b = \inf\{\, x \in I : x > a \text{ e } f_1(x) = f_2(x) = \cdots = f_n(x) = 0\,\},$$

allora $f_1(b) = f_2(b) = \cdots = f_n(b) = 0$.

Dimostrazione. Basta dimostrare la cosa per una sola funzione, ponendo

$$g(x) = f_1(x)^2 + f_2(x)^2 + \cdots + f_n(x)^2.$$

Esiste allora una successione s convergente a b tale che

$$s_n > a \qquad \text{e} \qquad g(s_n) = 0.$$

Per la continuità di g abbiamo concluso. □

Detto alla buona, l'enunciato afferma che, se le funzioni $f_1, f_2, \dots, f_n$ hanno uno zero in comune maggiore di a, allora l'insieme degli zeri comuni maggiori o uguali ad a ha minimo.

Possiamo allora applicare il teorema alle funzioni definite da $f_1(x) = |\mathsf{S}(x) - 1|$ e $f_2(x) = \mathsf{C}(x)$, che hanno uno zero comune, il c fornito dal lemma precedente. Sia allora b il minimo zero comune di queste funzioni con la proprietà che $b \geq 0$. Allora $b > 0$, perché $\mathsf{C}(0) = 1$. Poniamo

$$\pi = 2b.$$

Allora, per definizione $\mathsf{C}(\pi/2) = 0$ e $\mathsf{C}(x) > 0$, per $0 \leq x < \pi/2$. La funzione S è allora crescente nell'intervallo $[0, \pi/2]$ e quindi $\mathsf{S}(\pi/2) = 1$. Inoltre, per $0 < x < \pi/2$, $\mathsf{S}(x) < 1$.

Osservazione 8.9. Il numero π che abbiamo definito è esattamente il numero di Archimede, come vedremo più avanti. È una sfortunata coincidenza storica che sia stato scelto questo come numero fondamentale, e non la sua metà.

Con le formule di duplicazione abbiamo subito che

$$\mathsf{S}(\pi) = 0, \qquad \mathsf{C}(\pi) = -1, \qquad \mathsf{S}(2\pi) = 0, \mathsf{C}(2\pi) = 1$$

e, con le formule di addizione, vediamo che, per ogni $x \in \mathbf{R}$,

$$\mathsf{S}(x + 2\pi) = \mathsf{S}(x), \qquad \mathsf{C}(x + 2\pi) = \mathsf{C}(x).$$

Inoltre

$$\mathsf{S}\left(x + \frac{\pi}{2}\right) = \mathsf{C}(x).$$

Definizione 8.10. Sia $f\colon \mathbf{R} \to \mathbf{R}$ una funzione continua; diremo che f è *periodica* se esiste $T > 0$ tale che, per ogni $x \in \mathbf{R}$,

$$f(x+T) = f(x).$$

Un tale T si dirà un periodo di f.

Dunque S e C sono funzioni *periodiche* e il grafico di C si ottiene traslando quello di S. Abbiamo anche, facilmente, la tabella che mostra il comportamento delle due funzioni nell'intervallo $[0, 2\pi]$.

	$x=0$	$0<x<\frac{\pi}{2}$	$x=\frac{\pi}{2}$	$\frac{\pi}{2}<x<\pi$	$x=\pi$	$\pi<x<\frac{3}{2}\pi$	$x=\frac{3}{2}\pi$	$\frac{3}{2}\pi<x<2\pi$	$x=2\pi$
$\mathsf{S}(x)$	0	$\nearrow$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0
$\mathsf{C}(x)$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0	$\nearrow$	1

TABELLA 1. Crescenza e decrescenza delle funzioni S e C

Proposizione 8.11. *Sia $f\colon \mathbf{R} \to \mathbf{R}$ una funzione continua periodica e non costante. Allora esiste un minimo periodo di f.*

Dimostrazione. Consideriamo la funzione definita da $g(x) = f(x) - f(0)$. Allora, se $T > 0$ è un periodo di f, $g(T) = 0$. Sia T_0 l'estremo inferiore dell'insieme dei periodi di f; si dimostra, usando la stessa tecnica impiegata nella dimostrazione del teorema 8.8, che $f(x+T_0) = f(x)$, per ogni $x \in \mathbf{R}$ (esercizio). Non può essere $T_0 = 0$, altrimenti, per ogni $\varepsilon > 0$, troveremmo un periodo T_ε con $0 < T_\varepsilon < \varepsilon$.

Poiché f è per ipotesi non costante, esiste $\varepsilon > 0$ tale che $f(x) \neq f(0)$, per $x \in (-\varepsilon, \varepsilon)$ (teorema della permanenza del segno); ma $f(0 + T_\varepsilon) = f(0)$: assurdo. □

Corollario 8.12. *Il minimo periodo di S e C è 2π.*

Dimostrazione. Sappiamo che 2π è un periodo per entrambe le funzioni. Se T_0 è il minimo periodo, abbiamo $T_0 \leq 2\pi$, $\mathsf{S}(T_0) = 0$ e $\mathsf{C}(T_0) = 1$. La tabella sul comportamento delle funzioni S e C mostra allora che $T_0 = 2\pi$. □

Proposizione 8.13. *L'area di un cerchio di raggio 1 è uguale a π.*

Dimostrazione. Quello che dobbiamo dimostrare è che

$$\frac{\pi}{4} = \int_0^1 \sqrt{1-t^2}\, dt.$$

Eseguiamo la sostituzione $t = \mathsf{S}(u)$, che è giustificata dal fatto che la funzione S è crescente in $[0, \pi/2]$ ed assume in questo intervallo tutti i valori compresi fra 0 e 1. Abbiamo allora, ricordando che $\mathsf{C}(2u) = 2\,\mathsf{C}(u)^2 - 1$,

$$\int_0^1 \sqrt{1-t^2}\, dt = \int_0^{\pi/2} \sqrt{1-\mathsf{S}(u)^2} \cdot \mathsf{C}(u)\, du = \int_0^{\pi/2} \mathsf{C}(u)^2\, du = \int_0^{\pi/2} \frac{1+\mathsf{C}(2u)}{2} du.$$

Una primitiva di $(1 + \mathsf{C}(2u))/2$ è, chiaramente,

$$F(u) = \frac{1}{2}u + \frac{\mathsf{S}(2u)}{4},$$

quindi l'integrale è

$$\int_0^1 \sqrt{1-t^2}\, dt = F(\pi/2) - F(0) = \left(\frac{1}{2}\frac{\pi}{2} + \frac{\mathsf{S}(\pi)}{4}\right) - \left(\frac{1}{2}\cdot 0 + \frac{\mathsf{S}(0)}{4}\right) = \frac{\pi}{4},$$

come richiesto. □

Secondo la tradizione, possiamo allora porre, per $x \in \mathbf{R}$,

$$\operatorname{sen} x = \mathsf{S}(x) \qquad \text{e} \qquad \cos x = \mathsf{C}(x).$$

Queste sono proprio le usuali funzioni goniometriche.

Proposizione 8.14. *Dati $a, b \in \mathbf{R}$, con $a^2 + b^2 = 1$, esiste un unico $t \in \mathbf{R}$, $0 \le t < 2\pi$, tale che*

$$a = \cos t \qquad e \qquad b = \operatorname{sen} t.$$

Dimostrazione. È chiaro che, in questo caso, $|a| \le 1$ e $|b| \le 1$. Facciamo il caso in cui, ad esempio, $-1 \le a \le 0$ e $0 \le b \le 1$. Dalla tabella dei valori di seno e coseno, vediamo che è necessario scegliere $\pi/2 \le t \le \pi$. Per il teorema degli zeri e il fatto che S è decrescente nell'intervallo dato, esiste uno ed un sol valore di t tale che $\operatorname{sen} t = b$. Allora, da $(\mathsf{S}\, t)^2 + (\cos t)^2 = 1$ e $\cos t \le 0$, segue che

$$\cos t = -\sqrt{1 - b^2} = a.$$

Gli altri tre casi si trattano analogamente. □

Possiamo allora definire le altre funzioni goniometriche, e anche le inverse. Ad esempio,

$$\operatorname{tg} x = \frac{\operatorname{sen} x}{\cos x} \qquad \left(x \neq \frac{\pi}{2} + k\pi,\ k \text{ intero}\right)$$

è la tangente di x. La funzione "arcotangente" è definita su $\mathbf{R}$, a valori in $(-\pi/2, \pi/2)$: $\operatorname{arctg} x$ è l'unico numero reale t tale che $t \in (-\pi/2, \pi/2)$ e $\operatorname{tg} t = x$.

Dall'identità $x = \operatorname{tg}\operatorname{arctg} x$, segue che

$$\operatorname{arctg}' x = \frac{1}{1 + x^2}.$$

La serie di potenze seguente ha raggio di convergenza 1 e quindi definisce una funzione $f\colon (-1,1) \to \mathbf{R}$:

$$f(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Possiamo derivare f:

$$f'(x) = \sum_{n=0}^{+\infty} (-1)^n x^{2n} = \frac{1}{1 + x^2},$$

perché la serie è la serie geometrica. Perciò f e arctg differiscono in $(-1,1)$ per una costante; ma $f(0) = 0 = \operatorname{arctg} 0$, quindi abbiamo dimostrato che, per ogni $x \in (-1,1)$,

$$\operatorname{arctg} x = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Non è difficile giustificare che

$$\frac{\pi}{4} = \lim_{x \to 1^-} \operatorname{arctg} x = \sum_{n=0}^{+\infty} (-1)^n \frac{1}{2n+1},$$

poiché la serie indicata converge. Si tratta del famoso sviluppo in serie di π trovato da Eulero,

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$