GRAN VARIETA' MATEMATICO

Giorgio Mangioni

7 febbraio 2022

Perle di saggezza, miti e fatterelli dal corso di Analisi I 2018-19 del professor Pietro Majer

Introduzione

> Massì, tanto per analisi bastano
> le conoscenze del liceo...
>
> *Mia madre*

In questa (attualmente provvisoria, assolutamente non ufficiale) dispensa vi è parte delle stranezze che il buon vecchio Pierino Majer ha citato o mostrato contestualmente al corso di Analisi Matematica I, da lui tenuto nell'anno 2018-19 all'università di Pisa. Mi sono focalizzato su quegli argomenti che esulano dal programma di Analisi I standard (leggasi: Gobbino) ma che ritenevo interessanti e/o potenzialmente parte del materiale d'esame. Non sarà molto, ma è un lavoro onesto.
Data l'assenza di un buon ordinamento all'interno del mio quaderno, ho cercato di risistemare i capitoli in modo che la loro lettura avesse un certo senso. Ognuno di essi ha una sua coerenza interna, ma non è necessariamente sequenziale ai precedenti, in quanto talvolta richiede conoscenze citate in seguito. Alcune sezioni, indicate con (*), sono state aggiunte per completezza, oppure sono argomenti a mio avviso meno pregnanti di significato. Può capitare che le dimostrazioni vengano sostituite con altre più eleganti o semplici, mie o di altri (i quali verranno esplicitamente citati).
Ringrazio in anticipo chi vorrà darmi una mano a sistemare eventuali typo e imprecisioni. Per queste e altre questioni di vitale importanza scrivetemi a `giorgio.mangioni@gmail.com`.
N.B.: Questa dispensa è open source, quindi puoi farci quello che ti pare. L'importante è che, se mi rubi il lavoro, dopo non te ne vanti con gli amici.

Indice

1 Teoria degli insiemi e funzioni **7**
1.1 Funzione caratteristica 7
1.2 Teorema di Cantor-Bernstein-Schroeder 7
1.3 Esponenziazione di zero 7
1.4 Inclusione-Esclusione 7
1.5 Applicazioni di Inclusione-Esclusione 8
1.6 Formula delle fibre 9

2 Combinatoria **10**
2.1 Teorema dei tre fattoriali 10
2.2 Combinazioni con ripetizione 10
2.3 (*)Fattori primi del binomiale 10
2.4 Stima del binomiale centrale 11
2.5 Formula di Stirling 12

3 Campi **15**
3.1 Proprietà archimedea 15
3.2 Costruzione di $\mathbb{R}$ 15
3.3 Proprietà del sup 16
3.4 Esistenza della radice n-esima 16
3.5 Definizione di $\mathbb{C}$ 17
3.6 Teorema fondamentale dell'algebra 17

4 Polinomi e equazioni polinomiali **19**
4.1 Stima degli zeri di un polinomio 19
4.2 Radici complesse di $(1+z)^n - (1-z)^n$ 19
4.3 Numero di soluzioni di un'equazione di terzo grado 20
4.4 Formula di Cardano per le equazioni di terzo grado 20
4.5 Formula di Viéte per il "casus impossibilis" 21

5 Sviluppi polinomiali **22**
5.1 Proprietà degli sviluppi polinomiali 22
5.2 Polinomio di Taylor 23
5.3 Resto di Lagrange 23
5.4 Resto integrale 24
5.5 Derivate del polinomio di Taylor 25
5.6 (*)Quanto è buono il polinomio di Taylor? 25
5.7 Funzioni analitiche 27
5.8 Caratterizzazione delle funzioni C^n 28
5.9 Interpolazione polinomiale di Hermite 30

6 Disuguaglianze e identità **31**
6.1 Disuguaglianza delle medie 31
6.2 (*)Teorema di Nicomaco 31
6.3 Partial Fraction Decomposition - Esistenza 31
6.4 Partial Fraction Decomposition - Unicità 32

7 Spazi topologici e spazi metrici **34**
7.1 (*) Spazi topologici 34
7.2 Spazi metrici 34
7.3 (*) La distanza è una funzione lipschitziana 34
7.4 Lunghezza di una curva 35
7.5 Proprietà della lunghezza 35
7.6 Riparametrizzazioni 37
7.7 Compattezza: Definizioni equivalenti 37

7.8 Teorema di Heine-Borel . . . 37
7.9 Teorema di Weierstrass . . . 38
7.10 Teorema di Heine-Cantor (HeineKen) . . . 39

8 $\mathbb{R}^n$ **40**
8.1 Volume della palla euclidea . . . 40
8.2 Volume del simplesso con gli assi ortogonali . . . 41

9 Cenni di teoria della misura su $\mathbb{R}$ **42**
9.1 Insieme trascurabile . . . 42

10 Spazi normati **43**
10.1 Disuguaglianza di Cauchy-Schwarz . . . 43
10.2 Teorema del valore medio per curve a valori in $(V, \|\cdot\|)$ . . . 44
10.3 Variante integrale del TVM . . . 45
10.4 Calcolo della lunghezza tramite un integrale . . . 45
10.5 Polinomio di Taylor per curve . . . 46
10.6 Parametrizzazione in lunghezza d'arco . . . 47

11 Funzioni e continuità **49**
11.1 Definizione topologica di continuità . . . 49
11.2 Connessione . . . 49
11.3 Continuità per successioni . . . 50
11.4 Continuità uniforme . . . 51
11.5 Discontinuità di una funzione monotona . . . 52
11.6 Funzioni convesse . . . 52
11.7 Proprietà delle funzioni convesse di variabile reale . . . 52
11.8 Derivabilità di una funzione convessa . . . 53

12 Successioni **54**
12.1 Convergenza di sotto-successioni . . . 54
12.2 Ogni successione in $\mathbb{R}$ ammette una sotto-successione monotona (debolmente) . . . 54
12.3 Teorema di Bolzano-Weierstrass . . . 55
12.4 (*)Teorema delle contrazioni . . . 55
12.5 (*)Cenni al Teorema di Stolz-Cesaro . . . 56

13 Funzione esponenziale complessa **57**
13.1 La successione $(1 + x/n)^n$ . . . 57
13.2 $\exp(x)$ come isomorfismo di gruppi . . . 58
13.3 Serie esponenziale . . . 58
13.4 Stima del resto . . . 59
13.5 Irrazionalità di e . . . 60
13.6 Esponenziale complessa . . . 60
13.7 $\left|e^{ix} - e^{iy}\right| \leq |x - y|$ . . . 60

14 Funzioni trigonometriche **61**
14.1 Definizione di seno e coseno . . . 61
14.2 Definizione di π . . . 61
14.3 Sviluppo alternativo del seno . . . 61
14.4 Prodotto di Wallis . . . 62
14.5 Problema di Basilea . . . 62
14.6 (*) Somma delle potenze quarte . . . 63

15 Sommatorie e serie **65**
15.1 Somme di reali positivi . . . 65
15.2 Termini non nulli di una sommatoria convergente . . . 66
15.3 Somme di famiglie generiche di reali . . . 66
15.4 Prodotti di famiglie di reali positivi . . . 67

16 Calcolo di serie e produttorie **68**
16.1 Criterio di condensazione di Cauchy . . . 68
16.2 Criterio di Abel-Dirichlet . . . 68
16.3 (*)Serie binomiale . . . 69
16.4 (*)$\sum_{k\in\mathbb{N}} kc^k = \frac{c}{(c-1)^2}$ . . . 71
16.5 (*)Generalizzazione del caso precedente . . . 71
16.6 (*)Serie di Mengoli generalizzata . . . 71
16.7 Prodotto di $(1+a_i)$ . . . 72
16.8 (*) Stima del resto dell'armonica generalizzata . . . 73

17 Serie generatrici e applicazioni in Combinatoria **74**
17.1 Serie generatrice di una successione per ricorrenza . . . 74
17.2 Cardinalità delle partizioni di $[n]$ . . . 75
17.3 Numero di partizioni in un numero fissato di classi . . . 76
17.4 (*) Stima del resto della formula di Dobinski . . . 77
17.5 Problema del cambio della moneta . . . 78
17.6 Numeri di Catalan . . . 80

18 Integrale di Riemann-Darboux **82**
18.1 Integrabilità e definizioni equivalenti . . . 83
18.2 Proprietà dell'integrale . . . 84
18.3 Integrabilità delle funzioni continue e/o monotone . . . 84
18.4 Integrabilità della funzione composta . . . 85
18.5 Passaggio al limite sotto il segno di integrale . . . 86
18.6 Funzione integrale . . . 87
18.7 Teorema fondamentale del calcolo integrale . . . 87
18.8 Variante del TFCI . . . 88
18.9 Integrazione per parti . . . 89
18.10 Integrazione per sostituzione . . . 89
18.11 Sostituzioni notevoli . . . 89
18.12 Caratterizzazione delle funzioni integrabili . . . 90

19 Calcolo di integrali **93**
19.1 (*)$\int_0^\pi \sin^n x\,dx$ . . . 93
19.2 Teorema del prosciutto (Principio di Cavalieri) . . . 93
19.3 Test integrale per serie . . . 94
19.4 Criterio di Abel-Dirichlet integrale . . . 94

20 Funzione Gamma **95**
20.1 Primi cenni sulla funzione fattoriale . . . 95
20.2 Funzione Digamma . . . 96
20.3 Primitiva di g e costante di Eulero-Mascheroni . . . 98
20.4 Teorema di Artin-Bohr-Mollerup . . . 99
20.5 Rappresentazioni della funzione fattoriale . . . 100
20.6 Fattoriale di $-1/2$ e integrale della gaussiana. . . . 101
20.7 Sviluppi asintotici . . . 102
20.8 Formula di Stirling . . . 103
20.9 Formule di moltiplicazione . . . 103
20.10 Formula di riflessione . . . 105
20.11 Funzione Beta . . . 105

21 Successioni e serie di funzioni **106**
21.1 Convergenza puntuale e uniforme 106
21.2 Norma infinito 106
21.3 (*)Completezza delle funzioni limitate 106
21.4 Continuità del limite uniforme 107
21.5 Convergenza di funzioni equicontinue 107
21.6 Teorema del ConDom 107
21.7 Lemma di Abel per serie di potenze 109

22 Equazioni differenziali ordinarie **112**
22.1 Riscaldamento 112
22.2 Equazioni lineari omogenee a coefficienti costanti 113
22.3 EDOLCC non omogenee 114
22.4 Problema di Cauchy (indovinate un po'? Per le EDOLCC) 114
22.5 Principio di Duhamel 115
22.6 Equazioni lineari alle differenze finite 117
22.7 Equazioni lineari del primo ordine 117
22.8 (*)Equazione di Riccati 118
22.9 Come sfondare le equazioni differenziali in 16 minuti 119

23 (*)Esercizi e altre amenità irrilevanti **120**
23.1 Formica su una banda elastica 120
23.2 $\prod_{\text{p primo}} \left(1 - \frac{1}{p^k}\right)^{-1}$ 120
23.3 Postulato di Bertrand 120
23.4 Algoritmo di Erone per la radice quadrata 122
23.5 Algoritmo di Erone generalizzato 123
23.6 Teorema di Borel 123
23.7 Sviluppo asintotico di $c = x + \log x$ 124
23.8 Formula di inversione di Moebius 124
23.9 Funzione a spuntoni 124
23.10Curva del Pudding 125
23.11Randomiche osservazioni sulla funzione precedente 126
23.12Variante troll della funzione del pudding 127
23.13(*)Feynman's Trick e applicazione alla gaussiana 128
23.14Insieme di Cantor 129
23.15Funzione di Cantor-Vitali 130
23.16Funzione analitica modellabile a piacere 132

1 Teoria degli insiemi e funzioni

1.1 Funzione caratteristica

Dato un insieme Y, si consideri un suo sottoinsieme A. Si definisce funzione caratteristica χ_A la funzione

$$\begin{aligned}\chi_A :& Y \longrightarrow \mathbb{Z}_2 \\ y \longrightarrow & \begin{cases} 1 \text{ se } x \in A \\ 0 \text{ se } x \notin A \end{cases}\end{aligned}$$

La funzione caratteristica dell'intersezione è il prodotto delle funzioni caratteristiche:

$$\chi_{\bigcap_{i\in I} A_i}(x) = \prod_{i\in I} \chi_{A_i}(x)$$

1.2 Teorema di Cantor-Bernstein-Schroeder

Dati due insiemi X e Y, se esistono due funzioni iniettive $f : X \hookrightarrow Y$ e $g : Y \hookrightarrow X$, allora esiste una bigezione tra i due insiemi, Detto in altri termini, detta $|X|$ la cardinalità dell'insieme, si ha che

$$|X| \le |Y| \le |X| \implies |X| = |Y|$$

Dimostrazione Si definiscono per ricorrenza queste due successioni di insiemi:

$$\begin{cases} X_0 = X \\ X_{n+1} = g(Y_n) \end{cases} \qquad \begin{cases} Y_0 = Y \\ Y_{n+1} = f(X_n) \end{cases}$$

É possibile mostrare che gli insiemi $X_{2n} \setminus X_{2n+1}$ sono in bigezione tramite f con gli insiemi $Y_{2n+1} \setminus Y_{2n+2} = f(X_{2n}) \setminus f(X_{2n+1})$, e analogamente che gli insiemi $Y_{2n} \setminus Y_{2n+1}$ e $X_{2n+1} \setminus X_{2n+2}$ sono in bigezione tramite g. Manca solo da definire una bigezione per i punti che stanno in tutti gli X_n, e anche in questo caso si osserva che $\bigcap_{n\in\mathbb{N}} X_n$ è in bigezione secondo f con $\bigcap_{n\in\mathbb{N}} Y_n$. Infatti, se la restrizione di f non fosse suriettiva, esisterebbe un certo $y \in \bigcap_{n\in\mathbb{N}} Y_n$ privo di controimmagine, e ciò è impossibile per come sono stati definiti gli Y_n.

□

1.3 Esponenziazione di zero

Vale che $0^0 = \left|\emptyset^\emptyset\right| = 1$ perché esiste un'unica funzione dall'insieme vuoto in se stesso.

1.4 Inclusione-Esclusione

L'analisi si occupa spesso di questioni sociali. Più avanti tratteremo problemi di integrazione.

Pietro Majer

Detto $[n] = \{1, 2, \dots, n\}$, la cardinalità dell'unione di $n \in \mathbb{N}$ insiemi è data da

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{\emptyset\neq J\subseteq[n]} (-1)^{|J|+1} \left|\bigcap_{k\in J} A_k\right| \tag{1}$$

Dimostrazione Si procede per induzione su n. La tesi è banalmente vera per $n = 1$. Inoltre, è facile mostrarla anche per $n = 2$. Per quanto riguarda il passo induttivo, si ha che

$$\left|\bigcup_{i\leq k+1} A_i\right| = \left|\bigcup_{i\leq n} A_i \cup A_{k+1}\right|$$

Per ipotesi induttiva, applicando prima il caso $n = 2$ e poi il caso $n = k$, si ha

$$\left|\bigcup_{i\leq n} A_i \cup A_{k+1}\right| = \left|\bigcup_{i\leq n} A_i\right| + |A_{k+1}| - \left|\bigcup_{i\leq n} (A_i \cap A_{k+1})\right| =$$

$$= \sum_{\emptyset\neq J\subseteq[n]} (-1)^{|J|+1} \left|\bigcap_{i\in J} A_i\right| + |A_{k+1}| - \sum_{\emptyset\neq J\subseteq[n]} (-1)^{|J|+1} \left|\bigcap_{k\in J} (A_i \cap A_{k+1})\right|$$

Raccogliendo i due termini a sinistra, si ottiene la somma sui sottoinsiemi di indici $J \subseteq [k+1]$ che contengono $k + 1$, mentre il termine a destra è la somma sugli insiemi di indici che non lo contengono. Sommando i due pezzi si ottiene la tesi.

□

Assumendo che gli insiemi A_i siano contenuti in un universo Ω, dal teorema precedente segue una formula per la cardinalità del complementare dell'unione

$$\left|\left(\bigcup_{i=1}^{n} A_i\right)^C\right| = \sum_{J\subseteq[n]} (-1)^{|J|} \left|\bigcap_{k\in J} A_k\right|$$

Dove si assume che $\Omega = \bigcap_{i\in\emptyset} A_i$

1.5 Applicazioni di Inclusione-Esclusione

Definizione Una famiglia di insiemi si dice a intersezione regolare se la cardinalità dell'intersezione di un numero n finito di essi dipende solo da n, e non dagli insiemi considerati.

Funzioni suriettive Si mostra che la cardinalità delle suriezioni tra due insiemi finiti $[k]$ e $[m]$, con $k \geq m$, è data da

$$|S_{k,m}| = \sum_{h=0}^{m} \binom{m}{h} (-1)^h (m-h)^k \tag{2}$$

Dimostrazione Consideriamo prima le applicazioni non surgettive $NS_{k,m}$. Detto $N_i, \quad i \leq n$ l'insieme delle funzioni nella cui immagine non compare i, si ha

$$NS_{k,m} = \bigcup_{i\leq n} N_i$$

Per formula 1.4, la cardinalità delle funzioni surgettive risulta

$$|S_{k,m}| = \left|\left(\bigcup_{i\leq n} N_i\right)^C\right| = \sum_{J\subseteq[n]} (-1)^{|J|} \left|\bigcap_{h\in J} N_h\right|$$

Ora, gli insiemi N_i sono a intersezione regolare, perché la cardinalità delle funzioni la cui immagine non interseca un certo insieme J di indici dipende solamente dalla cardinalità h dello stesso, e corrisponde al numero di funzioni da $[k]$ in $[m] \setminus J$, che sono $(m - h)^k$. I modi di scegliere h elementi tra m, ossia il numero di volte in cui il termine $(m - h)^k$ compare nella somma, sono $\binom{m}{h}$. Si ha allora che

$$|S_{k,m}| = \sum_{h=0}^{m} \binom{m}{h} (-1)^h (m-h)^k$$

□

Punti fissi Si vuole ora determinare la cardinalità dell'insieme NF delle permutazioni di $[n]$ prive di punti fissi. Analogamente a quanto visto nell'applicazione precedente, l'insieme delle permutazioni F_J che fissano un sottoinsieme $J \subseteq [n]$ sono tante quante le permutazioni degli elementi restanti, e dunque sono $(n - |J|)!$. Dunque, si ha che

$$|NF| = \left| \left(\bigcup_{\emptyset \neq J \subseteq [n]} F_J \right)^C \right| = \sum_{J \subseteq [n]} (-1)^{|J|} \left| \bigcap_{k \in J} F_k \right| =$$

$$= \sum_{J \subseteq [n]} (-1)^{|J|} (n - |J|)! = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)!$$

□

1.6 Formula delle fibre

Data una funzione $f : X \longrightarrow Y$, si dice fibra di $y \in Y$ l'insieme $f^{-1}(y)$ delle sue controimmagini. Vale che $X = \bigcup_{y \in f(X)} f^{-1}(y)$ e che le fibre sono tutte disgiunte, e cioè le fibre determinano una partizione di X. Allora[1]

$$|X| = \sum_{y \in f(X)} \left| f^{-1}(y) \right| \tag{3}$$

[1] Si sottintende che la somma è quella tra cardinali (che corrisponde a quella usuale per cardinalità finite, ma fa cose molto buffe per insiemi e/o indici infiniti).

2 Combinatoria

2.1 Teorema dei tre fattoriali

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{4}$$

Dimostrazione Posseggo una mirabile dimostrazione combinatoria, hanc marginis exiguitas non caperet.. Diamo qui un'altra dimostrazione, di tipo analitico.
Per formula del binomio di Newton

$$(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$$

Derivando k volte ambo i membri

$$n(n-1)\dots(n-k+1)(1+x)^{n-k} = \sum_{i=0}^{n} \binom{n}{i} x^{i-k} i(i-1)\dots(i-k+1)$$

Eliminando a destra i termini di esponente negativo

$$\frac{n!}{(n-k)!}(1+x)^{n-k} = \sum_{i=k}^{n} \binom{n}{i} x^{i-k} \frac{i!}{(i-k)!}$$

Si impone $x = 0$. A destra si annullano tutti i termini eccetto il termine in cui $i = k$, poiché $0^0 = 1$.

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

Da cui segue la tesi.

□

2.2 Combinazioni con ripetizione

Le combinazioni con ripetizione di n oggetti in k posti sono date da $\binom{n+k-1}{k}$.

Dimostrazione Si ringrazia Sirio Resteghini per la seguente dimostrazione. Le combinazioni con ripetizione sono i modi di disporre n oggetti, indistinguibili tra loro, in k spazi. Supponiamo di mettere in fila gli n oggetti, e di aggiungere $k-1$ "separatori". Ciascuna combinazione è determinata dalla posizione dei separatori, in modo che gli elementi tra due separatori successivi stanno nello stesso spazio. Allora, è sufficiente contare i modi di permutare gli n oggetti, tutti uguali tra loro, e i $k-1$ separatori, tutti uguali tra loro. Le permutazioni con ripetizione di questi $n+k-1$ oggetti, di cui n ripetuti e $k-1$ ripetuti, sono esattamente $\binom{n+k-1}{k}$.

□

2.3 (*)Fattori primi del binomiale

Sia $n, p \in \mathbb{N}$, con p primo, $k \in \mathbb{N} \setminus \{0\}$. Allora

$$p^k \mid \binom{2n}{n} \implies p^k \leq 2n$$

Dimostrazione Indicheremo con $\lfloor x \rfloor$ la parte intera (inferiore) di x. Sia $\nu_p(a)$ l'esponente di p nella fattorizzazione di a. Essendo un esponente, valgono le proprietà dei logaritmi, perciò la quantità da considerare è

$$\nu_p\left(\binom{2n}{n}\right) = \nu_p\left(\frac{(2n)!}{(n!)^2}\right) = \nu_p(2n!) - 2\nu_p(n!)$$

Vale, inoltre, la formula di Legendre

$$\nu_p(n!) = \sum_{i=1}^{+\infty} \left\lfloor \frac{n}{p^i} \right\rfloor$$

Infatti, p compare con esponente almeno una volta nei numeri minori o uguali a n divisibili per p, che sono $\left\lfloor \frac{n}{p} \right\rfloor$, almeno due volte nei numeri divisibili per p^2...
Ora, se $p^k \mid \binom{2n}{n}$, vale che, detto $M = max(m \in \mathbb{N} \quad | \quad p^m \leq 2n)$, si ha

$$k \leq \nu_p(2n!) - 2\nu_p(n!) = \sum_{i=1}^{M} \left\lfloor \frac{2n}{p^i} \right\rfloor - 2\nu_p(n!)$$

Vogliamo mostrare che l'esponente k vale al più M. Distinguiamo ora due casi

- Se $p^M \leq n$, allora $\left\lfloor \frac{n}{p^M} \right\rfloor$ compare nella scrittura di $\nu_p(n!)$. Allora

  $$k \leq \sum_{i=1}^{M} \left\lfloor \frac{2n}{p^i} \right\rfloor - 2\left\lfloor \frac{n}{p^i} \right\rfloor$$

  Il termine generale della serie vale al più 1. Dunque $k \leq M$, come voluto.

- se $p^M > n$, con un ragionamento analogo si ha che i primi $M-1$ termini valgono al più $M-1$. Inoltre, il termine $\left\lfloor \frac{2n}{p^M} \right\rfloor$ vale 1, perché $n < p^M \leq 2n$. Perciò anche in questo caso si conclude che $k \leq M$.

$\square$

2.4 Stima del binomiale centrale

Per $n \to +\infty$ vale la seguente formula:

$$\binom{2n}{n} \approx \frac{2^{2n}}{\sqrt{\pi n}} \tag{5}$$

Dimostrazione Facciamo uso della formula del prodotto di Wallis, ottenuta in 14.4.

$$\pi = 2 \prod_{k \in \mathbb{N}} \frac{2k}{2k-1} \frac{2k}{2k+1}$$

Riscriviamola moltiplicando e dividendo per $(2k)^2$ e esplicitando il processo di passaggio al limite sui prodotti parziali:

$$1 = \pi/2 \lim_{n \to +\infty} \prod_{k=0}^{n} \frac{2k-1}{2k} \frac{2k+1}{2k} \frac{(2k)^2}{(2k)^2} = \pi/2 \lim_{n \to +\infty} \prod_{k=0}^{n} \frac{(2k+1)(2k)\,(2k)(2k-1)}{(2k)^4}$$

Osserviamo che il prodotto $\prod_{k=0}^{n}(2k+1)(2k)$ vale $(2n+1)!$, e analogamente per gli altri fattoriali. Otteniamo dunque

$$1 = \pi/2 \lim_{n \to +\infty} \frac{(2n+1)!\,(2n)!}{2^{4n}(n!)^4} = \pi/2 \lim_{n \to +\infty} \frac{(2n+1)\,((2n)!)^2}{2^{4n}(n!)^4} = \pi/2 \lim_{n \to +\infty} \frac{2n+1}{2^{4n}} \binom{2n}{n}^2$$

Da cui
$$\lim_{n\to+\infty}\binom{2n}{n}^2=\lim_{n\to+\infty}\frac{2^{4n}}{\pi(n+1/2)}$$
Estraendo la radice segue la tesi:
$$\binom{2n}{n}=\frac{2^{2n}}{\sqrt{\pi(n+1/2)}}(1+o(1))=\frac{2^{2n}}{\sqrt{\pi n}}(1+o(1))$$
□

Volendo, si può ottenere una formula più precisa:
$$\frac{2^{2n}}{\sqrt{\pi n}}\left(1-\frac{1}{4n}+o\left(\frac{1}{n}\right)\right)\le\binom{2n}{n}\le\frac{2^{2n}}{\sqrt{\pi n}}$$
Infatti, se si estrae il logaritmo della formula del prodotto di Wallis si ottiene che
$$\log\pi=\log 2+\sum_{k\in\mathbb{N}}(-1)^{k+1}\log\left(1+\frac{1}{k}\right)$$
La serie è convergente per criterio di Leibniz, e in particolare il limite è compreso tra ciascuna somma pari e la successiva somma dispari. Poiché il logaritmo è un omeomorfismo le stesse considerazioni valgono per i prodotti. Allora vale che
$$2\prod_{k=1}^{n}\frac{2k}{2k-1}\frac{2k}{2k+1}\le\pi\le 2\left(\prod_{k=1}^{n}\frac{2k}{2k-1}\frac{2k}{2k+1}\right)\frac{2k}{2k-1}$$
Ora si procede analogamente a quanto già visto, moltiplicando per $4n^2$ sopra e sotto:
$$\frac{4^{2n}(n!)^4}{((2n)!)^2(2n+1)}\le\pi/2\le\frac{4^{2n}(n!)^4}{((2n)!)^2(2n)}$$
Estraendo la radice quadrata
$$\frac{4^n(n!)^2}{(2n)!\sqrt{n+1/2}}\le\sqrt{\pi}\le\frac{4^n(n!)^2}{(2n)!\sqrt{n}}$$
Da cui
$$\frac{2^{2n}}{\sqrt{\pi n(1+1/2n)}}\le\binom{2n}{n}\le\frac{2^{2n}}{\sqrt{\pi n}}$$
Sviluppando il denominatore in serie di Taylor otteniamo
$$\frac{2^{2n}}{\sqrt{\pi n}}\left(1-\frac{1}{4n}+o\left(\frac{1}{n}\right)\right)\le\binom{2n}{n}\le\frac{2^{2n}}{\sqrt{\pi n}}$$
□

2.5 Formula di Stirling

Per $n\to+\infty$, è possibile stimare il fattoriale come
$$n!\approx\sqrt{2\pi n}\,\frac{n^n}{e^n}\tag{6}$$

Dimostrazione Iniziamo a fare stime grezze. Mostriamo innanzitutto, per induzione su n, che

$$n! \geq \left(\frac{n}{e}\right)^n$$

Per $n = 1$ la tesi è banalmente vera. Ora, se è vera per $n = k$, vale che

$$(k+1)! \geq (k+1)\left(\frac{k}{e}\right)^k \overset{?}{\geq} \left(\frac{k+1}{e}\right)^{k+1}$$

Semplificando la formula a destra

$$ek^k \overset{?}{\geq} (k+1)^k \quad \longrightarrow \quad e \geq \left(1 + \frac{1}{k}\right)^k$$

Che è vero perché la successione $\left(1+\frac{1}{k}\right)^k$ tende ad e dal basso.
Cerchiamo ora di essere più quantitativi. Ci chiediamo se sia possibile trovare due costanti a e C reali positive tali che

$$n! = C\left(\frac{n}{e}\right)^n n^a \left(1 + o(1)\right)$$

In altri termini, ci chiediamo se sia possibile costruire una successione $x_n = \left(\frac{n}{e}\right)^n n^a/n!$ che converga a un numero reale non nullo $1/C$. Ogni termine della successione è evidentemente non nullo, dunque possiamo esprimerlo come prodotto "telescopico" dei precedenti.

$$x_n = \frac{x_n}{x_{n-1}} \frac{x_{n-1}}{x_{n-2}} \dots x_1$$

Estraiamo il logaritmo di x_n. Infatti, la successione converge a un numero reale positivo sse converge la successione dei logaritmi.

$$\log x_n = \left(\sum_{k=1}^{n-1} \log\left(\frac{x_{k+1}}{x_k}\right)\right) + \log x_1$$

Il termine x_1 vale $1/e$, dunque il suo logaritmo fa -1. Calcoliamo ora il rapporto tra due termini:

$$\frac{x_{n+1}}{x_n} = \frac{\left(\frac{n+1}{e}\right)^{n+1}(n+1)^a}{(n+1)!} \frac{n!}{\left(\frac{n}{e}\right)^n n^a} = \frac{1}{e(n+1)n^{n+a}}(n+1)\,(n+1)^{n+a} = \frac{1}{e}\left(1+\frac{1}{n}\right)^{n+a} \qquad (\star)$$

Sostituendo nella formula otteniamo

$$\log x_n = -1 + \sum_{k=1}^{n-1} \log\left(\frac{1}{e}\left(1+\frac{1}{k}\right)^{k+a}\right) = -1 + \sum_{k=1}^{n-1}\left((k+a)\log\left(1+\frac{1}{k}\right) - 1\right)$$

Per $k \gg 1$ vale lo sviluppo in $x = 0$ del logaritmo:

$$\log\left(1+\frac{1}{k}\right) = \frac{1}{k} - \frac{1}{2k^2} + \frac{1}{3k^3} + \mathcal{O}\left(\frac{1}{k^4}\right)$$

Il termine generale diventa allora

$$(k+a)\log\left(1+\frac{1}{k}\right) - 1 = (k+a)\left(\frac{1}{k} - \frac{1}{2k^2} + \frac{1}{3k^3} + \mathcal{O}\left(\frac{1}{k^4}\right)\right) - 1 =$$

$$= \left(a - \frac{1}{2}\right)\frac{1}{k} + \left(\frac{1}{3} - \frac{a}{2}\right)\frac{1}{k^2} + \mathcal{O}\left(\frac{1}{k^3}\right)$$

Se $a = 1/2$, il termine generale è un multiplo della serie dei reciproci dei quadrati, e dunque la serie converge. In ogni altro caso non si annulla il termine di grado 1, e la serie diverge per confronto con la serie armonica. Abbiamo allora che

$$n! = C\left(\frac{n}{e}\right)^n \sqrt{n}\left(1 + o(1)\right)$$

Per determinare C utilizziamo lo sviluppo asintotico del binomiale centrale, mostrato in 2.4, sostituendo i fattoriali con la formula appena trovata.

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} \approx \frac{C\left(\frac{2n}{e}\right)^{2n}\sqrt{2n}}{C^2\left(\frac{n}{e}\right)^{2n} n} = \frac{2^{2n}\sqrt{2}}{C\sqrt{n}}$$

Poiché vale $\binom{2n}{n} \approx \frac{2^{2n}}{\sqrt{\pi n}}$, comparando i due sviluppi si deve avere che $C = \sqrt{2\pi}$. Dunque segue la tesi:

$$n! \approx \sqrt{2\pi n}\,\frac{n^n}{e^n}$$

□

Ci si può chiedere ora se la stima sia dall'alto o dal basso, o se la successione $x_n = \left(\frac{n}{e}\right)^n \frac{\sqrt{n}}{n!}$ oscilli attorno al limite. Si mostra che la successione è crescente, e dunque l'approssimazione tende al fattoriale dal basso. Consideriamo infatti il rapporto tra due termini successivi e mostriamo che è definitivamente maggiore di 1, ossia che il suo logaritmo è positivo. Come mostrato in ($\star$) vale

$$\log\left(\frac{x_{n+1}}{x_n}\right) = \left(n + \frac{1}{2}\right)\log\left(1 + \frac{1}{n}\right) - 1$$

Sviluppando il logaritmo in serie di Taylor otteniamo

$$\log\left(\frac{x_{n+1}}{x_n}\right) = \left(n + \frac{1}{2}\right)\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{6n^3} + o\left(\frac{1}{n^3}\right)\right) - 1 = \frac{1}{12n^2} + o\left(\frac{1}{n^2}\right) > 0 \text{ definitivamente.}$$

3 Campi

Si tralasciano tutte le semplici proprietà dei campi, che seguono dalla definizione (e/o dal corso di Aritmetica). Si riassumono solo le proprietà più importanti:

- In un campo ordinato K vale che, $\forall x \in K \quad x^2 > 0$. Ciò implica in particolare che $\mathbb{C}$ non ammette un ordinamento compatibile con le operazioni (sebbene sia possibile definire un ordinamento, ad esempio lessicografico).
- Ogni campo ordinato è infinito e contiene un sottocampo isomorfo a $\mathbb{Q}$, che si costruisce esattamente come i razionali a partire dai naturali.

3.1 Proprietà archimedea

Dato un campo ordinato K, è equivalente dire che:

- K gode della proprietà archimedea;
- $\forall x \in K$, l'insieme dei multipli di x, $\{nx, \quad n \in \mathbb{N}\}$ è superiormente illimitato. In particolare, l'insieme dei "numeri naturali" è superiormente illimitato;
- $\mathbb{Q}_K$, il sottoinsieme di K isomorfo a $\mathbb{Q}$, è denso in K, ossia

$$\forall x < y \in K \quad \exists q \in \mathbb{Q}_K \quad | \quad x < q < y$$

3.2 Costruzione di $\mathbb{R}$

Def Un campo ordinato si dice completo se ogni insieme superiormente limitato ammette estremo superiore. É equivalente dire che ogni successione di Cauchy converge nel campo (cosa che non richiede l'ordinamento).

Teorema Esiste ed è unico (a meno di isomorfismi) un campo ordinato completo. Tale campo si identifica con l'insieme dei numeri reali.

Esistenza

- Si costruisce un campo ordinato completo a partire dal campo ordinato $\mathbb{Q}$. Per fare ciò, si definisce sezione di Dedekind un sottoinsieme $\alpha \subset \mathbb{Q}$ che soddisfa le seguenti proprietà:
  - α è inferiormente illimitato;
  - α è privo di massimo.

  Si definisce $\mathbb{R}$ come l'insieme delle sezioni di Dedekind. Osserviamo innanzitutto che esiste un'immersione iniettiva di $\mathbb{Q}$ in $\mathbb{R}$ che manda ciascun razionale q nella sezione $\{x \in \mathbb{Q} \quad | \quad x < q\}$.
- Si vuole dotare $\mathbb{R}$ di una struttura di campo ordinato. Per fare ciò, si definisce un ordinamento tra sezioni, che corrisponde all'inclusione insiemistica:

  $$\alpha < \beta \iff \alpha \subset \beta$$

  Si definisce la somma tra sezioni come

  $$\alpha + \beta = \{x + y \quad | \quad x \in \alpha \wedge y \in \beta\}$$

  La definizione del prodotto è più complessa e richiede un po' di casistiche. Si verifica che $\mathbb{R}$ dotato delle operazioni e dell'ordinamento di cui sopra è un campo ordinato. Si sottolinea che, nel processo di definizione delle operazioni, è fondamentale la proprietà archimedea di $\mathbb{Q}$.

- Mostriamo che $\mathbb{R}$ è completo. Sia A un insieme non vuoto di sezioni. Escludiamo il caso in cui l'insieme è superiormente illimitato, in cui si ha banalmente che $\sup(A) = +\infty$. Sia allora A superiormente limitato. Allora esso ammette estremo superiore, che corrisponde all'unione di tutte le sezioni di A
$$\aleph = \sup A = \bigcup_{\alpha \in A} \alpha$$
É evidente che $\aleph$ è un maggiorante per l'insieme. Inoltre, ogni maggiorante deve essere più grande di ogni sezione, e dunque maggiore o uguale della loro unione. Perciò $\aleph$ è il minimo tra i maggioranti, e dunque è l'estremo superiore. Ora, basta osservare che $\aleph \neq +\infty$, dato che l'insieme era superiormente limitato.

Prima di passare all'unicità forniamo il seguente lemma:

Lemma Un campo completo è archimedeo.

Dimostrazione Sia M l'estremo superiore dell'insieme dei numeri naturali. Per assurdo, M non sia infinito. Allora vale che $M > n+1 \quad \forall n \in \mathbb{N}$. Spostando a sinistra 1 si osserva che anche $M-1$ è un maggiorante dell'insieme dei numeri naturali, cosa che contraddice l'ipotesi che M fosse il sup.

Unicità Sia K un campo ordinato completo e sia $\mathbb{Q}_K$ il sottoinsieme di K isomorfo a $\mathbb{Q}$. Si completa $\mathbb{Q}_K$ allo stesso modo con cui si è completato $\mathbb{Q}$. Ciò induce un isomorfismo tra $\mathbb{R}$ e il completamento $\bar{K}$. Quest'ultimo coincide con K, in soldoni perché ogni elemento di K può essere ottenuto come sup di una successione di elementi di $\mathbb{Q}_K$, vicina a piacere per proprietà archimedea. (*) Più precisamente, presi comunque due elementi $x, y \in K$, $\exists q \in \mathbb{Q}$ tale che $x < q < y$. Ciò significa che è sempre possibile separare due punti in due diverse sezioni, da cui discende che ogni punto di K è di accumulazione per $\mathbb{Q}_K$ ed è quindi contenuto nel completamento. Dal momento che K è già completo, deve coincidere con il suo completamento.

□

3.3 Proprietà del sup

Tutte le proprietà seguenti sono di facile dimostrazione, e seguono ciascuna dalle precedenti:

1. $\sup A \leq \sup B \iff M(B) \subset M(A)$, dove $M(A)$ è l'insieme dei maggioranti di A;
2. $\sup A \leq \sup B \iff \forall a \in A \exists b \in B (a \leq b)$;
3. $\sup \bigcup_{i \in I} A_i = \sup(\sup A_i, \quad i \in I)$;
4. $\sup(A+B) = \sup\{a+b, \quad a \in A \wedge b \in B\} = \sup A + \sup B$

Dimostrazione di 3 Poiché $\forall a \in \bigcup_{i \in I} \exists \sup A_i \quad | \quad a \leq \sup A_i$, si ha, per la proprietà 2, $\sup \bigcup_{i \in I} A_i \subset \sup(\sup A_i, \quad i \in I)$.
Per l'altra inclusione, si osserva che $\sup \bigcup_{i \in I} A_i \geq \sup A_i$. Allora passando al sup a destra si ha $\sup \bigcup_{i \in I} A_i \supset \sup(\sup A_i, \quad i \in I)$. Si conclude per estensionalità che i due insiemi coincidono.

□

3.4 Esistenza della radice n-esima

Si fornisce qui una dimostrazione che fa uso di alcune proprietà delle funzioni continue. Per una dimostrazione alternativa e più "con le mani" si rimanda al paragrafo sull'algoritmo di Erone.
Consideriamo la funzione $y = x^n$. Essa è continua in $\mathbb{R}_0^+$ perché si ottiene moltiplicando la funzione $y = x$, continua, con se stessa. Dunque l'immagine di un intervallo è un intervallo, per teorema dei valori intermedi. In particolare, l'immagine di $\mathbb{R}_0^+$ è $\mathbb{R}_0^+$ stesso, in quanto deve essere illimitata superiormente e deve contenere lo zero. Da ciò segue che f è surgettiva in $\mathbb{R}_0^+$. Inoltre, la funzione è strettamente crescente, e dunque iniettiva. Ciò implica che è biettiva, e dunque ammette inversa.

□

3.5 Definizione di $\mathbb{C}$

Definizione algebrica Il campo dei complessi è il campo di spezzamento del polinomio $x^2 + 1$ su $\mathbb{R}$, ossia $\mathbb{C} = \mathbb{R}[x]_{/(x^2+1)}$. Gli elementi del campo sono dunque della forma $a + bi$, con $a, b \in \mathbb{R}$.

Definizione operazionale Il campo dei complessi è $\mathbb{R}^2$ munito delle seguenti operazioni:

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) * (c, d) = (ac - bd, ad + bc)$$

Si osserva che in generale non è possibile estendere il concetto di prodotto a un qualsiasi $\mathbb{R}^n$, ossia in generale $\mathbb{R}^n$ non è un campo.

Definizione geometrica Il campo dei complessi è un sottocampo dell'anello degli endomorfismi di $\mathbb{R}^2$, i cui elementi sono rappresentati da matrici $2x2$ della forma

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

Una matrice di questo tipo si dice speciale conforme ed è associata univocamente al numero complesso $a + bi$.

- Una matrice speciale ha determinante positivo (a meno che $a = b = 0$). Di fatto, il determinante corrisponde al modulo quadro del numero complesso $a + bi$.
- Una matrice è conforme se l'inversa è un multiplo della trasposta. In particolare, Il fattore di scala è il determinante della matrice.

  $$A^T A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} det(A) & 0 \\ 0 & det(A) \end{bmatrix}$$

  Si osserva che la matrice trasposta è associata al numero complesso coniugato, in quanto scambia b con $-b$. Si ritrova allora la formula nota $z\bar{z} = |z|^2$.
  Sviluppando i conti con una matrice generica $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ si può mostrare che essa è conforme se e solo se è della forma $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.
- La condizione di conformità è equivalente a dire che la matrice commuta con $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, che rappresenta l'unità immaginaria. Ciò segue dal fatto che, sempre sviluppando i conti, si mostra che una matrice commuta con J se e solo se é della forma $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$
- Si osserva che l'endomorfismo associato a un numero complesso $z = \rho e^{i\theta}$ corrisponde alla moltiplicazione di un numero complesso per z, che si può vedere come la composizione di una dilatazione di costante $\rho = |z|$ e di una rotazione di angolo θ, entrambe centrate nell'origine.
- (*) L'endomorfismo che manda un numero complesso nel coniugato è un isomorfismo (perché conserva il modulo) involutorio (composto con se stesso dà l'identità).

3.6 Teorema fondamentale dell'algebra

$\mathbb{C}$ è algebricamente chiuso, ossia ogni polinomio a coefficienti complessi di grado non nullo ammette almeno una radice. In altri termini, esso si fattorizza completamente in termini di grado 1.

Dimostrazione Esistono un'infinità di dimostrazioni di questo fatto (ce n'è una di algebra lineare molto carina, che vi invito a cercare, *n.d.A.*). Diamo qui una dimostrazione di tipo variazionale. Consideriamo un polinomio $p(z)$ e associamo ad esso la funzione

$$\begin{aligned} f : \mathbb{C} &\longrightarrow \mathbb{R}^+_{\not 0} \\ z &\longrightarrow |p(z)| \end{aligned}$$

La funzione ammette un minimo reale m. Infatti,poiché $|p(z)| \to +\infty$ per $|z| \to +\infty$, il minimo, se esiste, viene assunto in un punto che sta in una regione limitata e chiusa del piano, attorno all'origine. Un sottoinsieme chiuso e limitato di $\mathbb{R}^n$ è compatto, e il teorema di Weierstrass per gli spazi metrici, mostrato in 7.9, afferma che una funzione continua da un compatto a $\mathbb{R}$ ha massimo e minimo. Dunque il minimo esiste.
Per assurdo, supponiamo che il minimo non sia zero, ossia che il polinomio non si annulli mai, e mostriamo che è possibile discostarsi dal minimo per assumere valori più piccoli, ossia che quello considerato non era il minimo. WLOG, assumiamo che il minimo venga assunto in $z = 0$. Allora il polinomio sarà della forma $p(z) = a_n z^n + \ldots + a_0$, con $m = a_0 \neq 0$. Detto b il primo coefficiente non nullo dopo a_0, che esiste dato che $\deg p \neq 0$. Allora, per $z \to 0$, $p(z) = a_0 + bz^r + o(z^r) = b\left(\frac{a_0}{b} + z^r + o(z^r)\right)$.
Sia $v = \sqrt[r]{-\frac{a_0}{b}}$. Tale valore esiste (sebbene non sia unico) perché siamo in $\mathbb{C}$, e in particolare è non nullo poiché a_0 non lo è. Allora v è rappresentato nel piano di Argand-Gauss da un vettore non nullo, con una sua direzione e un suo verso. Spostiamoci lungo la direzione di v, ossia lungo la retta $y = tv$, con $t \in \mathbb{R}$, e calcoliamo i valori assunti dalla funzione.

$$|p(tv)| = \left| b\left(\frac{a_0}{b} + (tv)^r + o(t^r)\right) \right| = |a_0(1 - t^r) + o(t^r)| =$$

$$= m\,|1 - t^r + o(t^r)| < m \text{ per } t \text{ sufficientemente piccolo.}$$

Dunque si è trovata una direzione in cui il polinomio assume valori minori del minimo, e ciò è assurdo.

□

4 Polinomi e equazioni polinomiali

4.1 Stima degli zeri di un polinomio

Sia $p(x) = \sum_{k\leq N} a_k x^k$ e sia $\bar{x}$ un suo zero. Allora vale che, detto a_N il termine di grado massimo,

$$|\bar{x}| \leq 1 + \max_{k\leq N}\left(\left|\frac{a_k}{a_N}\right|\right)$$

Dimostrazione Escludiamo innanzitutto il caso in cui $|\bar{x}| \leq 1$, che è ovvio. Sia allora $|\bar{x}| > 1$. Dunque

$$p(\bar{x}) = \sum_{k\leq N} a_k \bar{x}^k = 0$$

$$\left|\bar{x}^N\right| = \left|\sum_{k\leq N-1} a_k \bar{x}^k\right|$$

Per disuguaglianza triangolare segue che

$$\left|\bar{x}^N\right| \leq \sum_{k\leq N-1}\left|a_k \bar{x}^k\right| \leq \max_{k\leq N}\left(\left|\frac{a_k}{a_N}\right|\right)\sum_{k\leq N-1}|\bar{x}|^k = \max_{k\leq N}\left(\left|\frac{a_k}{a_N}\right|\right)\frac{|\bar{x}|^N - 1}{|\bar{x}| - 1}$$

Detto $M = \max_{k\leq N}\left(\left|\frac{a_k}{a_N}\right|\right)$, si ha

$$|\bar{x}|^N \leq M\frac{|\bar{x}|^N - 1}{|\bar{x}| - 1} \leq M\frac{|\bar{x}|^N}{|\bar{x}| - 1}$$

$$1 \leq \frac{M}{|\bar{x}| - 1} \longrightarrow |\bar{x}| \leq M + 1$$

$\square$

4.2 Radici complesse di $(1 + z)^n - (1 - z)^n$

La formula sarà utile per alcuni risultati derivanti dallo sviluppo del seno, come lo sviluppo di Wallis e la formula asintotica del binomiale centrale.
Vogliamo risolvere

$$(1 + z)^n = (1 - z)^n$$

Osserviamo subito che 1 non è una radice, cosa che si vede banalmente sostituendo. Allora possiamo dividere per $(1-z)^n$ e osservare che $(1+z)/(1-z)$ deve essere una radice n-esima dell'unità. Dunque

$$\frac{1 + z}{1 - z} = e^{\frac{2i\pi k}{n}}, \text{ con } k = 0, \ldots, n - 1$$

Da cui, sviluppando i conti,

$$z = \frac{e^{\frac{2i\pi k}{n}} - 1}{e^{\frac{2i\pi k}{n}} + 1}$$

Si osserva che, se n è pari, bisogna escludere il valore $k = n/2$, per il quale il denominatore è nullo. Moltiplichiamo sopra e sotto per $\frac{e^{-\frac{i\pi k}{n}}}{2i}$, in modo da ottenere il rapporto tra seno e coseno, visti rispettivamente come parte immaginaria e parte reale di $e^{\frac{i\pi k}{n}}$.

$$z = i\frac{e^{\frac{i\pi k}{n}} - e^{-\frac{i\pi k}{n}}}{2i}\frac{2}{e^{\frac{i\pi k}{n}} + e^{-\frac{i\pi k}{n}}} = i\frac{\sin\frac{\pi k}{n}}{\cos\frac{\pi k}{n}} = i\tan\frac{\pi k}{n}, \text{ con } k = 0, \ldots, n - 1, \quad k \neq n/2$$

Osserviamo che le radici sono esattamente n se n è dispari, $n-1$ se n è pari (caso in cui l'equazione si abbassa di grado). Una radice è $z=0$ e possiamo separarla dal resto. —'E possibile scrivere il polinomio originale come

$$(1+z)^n-(1-z)^n = az\prod_{0<k<n}^{k\neq n/2}\left(1-\frac{z}{i\tan\frac{\pi k}{n}}\right)$$

Per un certo coefficiente reale a. Dato che quest'ultimo risulta essere il coefficiente del termine di grado 1, si deve avere che $a=2n$, in quanto $(1+z)^n-(1-z)^n = 2nz+\ldots$. Dunque

$$(1+z)^n-(1-z)^n = 2nz\prod_{0<k<n}^{k\neq n/2}\left(1-\frac{z}{i\tan\frac{\pi k}{n}}\right) \tag{7}$$

Dato che le radici sono coniugate a coppie, essendo il polinomio a coefficienti reali, la fattorizzazione in $\mathbb{R}[x]$ del polinomio sarà la seguente:

$$(1+z)^n-(1-z)^n = 2nz\prod_{0<k<n/2}\left(1+\frac{z^2}{\tan^2\frac{\pi k}{n}}\right) \tag{8}$$

□

4.3 Numero di soluzioni di un'equazione di terzo grado

Ci chiediamo per quali valori di p e q un'equazione del tipo $p(x)=x^3+px+q=0$ ammetta tre soluzioni reali. Consideriamo la derivata $p'(x)=3x^2+p$.
Se $p\geq 0$ la derivata è sempre positiva o nulla. Dunque il polinomio è strettamente crescente, e non può avere più di una intersezione con l'asse x.
Se invece $p<0$, affinché il polinomio si annulli i punti stazionari devono avere ordinate discordi. Siano x_1, x_2 tali punti. Vale che $p'(x_1)=p'(x_2)=0$, e dunque $x_1=-x_2=\sqrt{-p/3}$. Ora, imponiamo che il prodotto delle loro ordinate sia negativo.

$$\left(x_1^3+px_1+q\right)\left(-x_1^3-px_1+q\right) = \left(q^2-(x_1^3+px_1)^2\right)<0$$

Ossia

$$q^2<(x_1^3+px_1)^2 = \left(\sqrt{-p^3}\left(\frac{1}{\sqrt{27}}+\frac{1}{\sqrt{3}}\right)\right)^2 = -p^3\left(\frac{1}{27}+\frac{1}{3}+\frac{2}{9}\right)$$

Quindi, la condizione affinché si avviano tre radici positive distinte è che

$$q^2+\frac{4p^3}{27}<0$$

4.4 Formula di Cardano per le equazioni di terzo grado

Consideriamo un'equazione generica di grado 3, del tipo

$$x^3+ax^2+bx+c=0$$

Con un cambio di variabile, è possibile eliminare il termine di grado 2, ponendo $y=x+a/3$. Così facendo si ottiene un'equazione del tipo

$$y^3+py+q=0$$

Supponiamo che sia $y=u+v$. Allora

$$y^3=(u+v)^3=u^3+v^3+3uv(u+v)=u^3+v^3+3uvy$$

Sostituendo nell'equazione originaria si ha

$$y^3+py+q=u^3+v^3+3uvy+py+q=y(3uv+p)+(u^3+v^3+q)$$

Allora, la soluzione del sistema

$$\begin{cases} 3uv = -p \\ u^3 + v^3 = -q \end{cases}$$

è una soluzione dell'equazione, perché annulla entrambi i termini. Elevando la prima equazione alla terza si ottiene

$$\begin{cases} u^3 v^3 = -\frac{p^3}{27} \\ u^3 + v^3 = -q \end{cases}$$

Il sistema è ora risolubile facilmente ponendo $U = u^3$ e $V = v^3$. Sviluppando i conti si giunge a

$$U = -\frac{q + \sqrt{q^2 + \frac{4}{27}p^3}}{2} \qquad V = -\frac{q - \sqrt{q^2 + \frac{4}{27}p^3}}{2}$$

Da cui si ottiene

$$Y = \sqrt[3]{U} + \sqrt[3]{V} = \sqrt[3]{-\frac{q + \sqrt{q^2 + \frac{4}{27}p^3}}{2}} + \sqrt[3]{-\frac{q - \sqrt{q^2 + \frac{4}{27}p^3}}{2}}$$

Il problema è che la formula risolutiva funziona solo se l'equazione ha una sola soluzione reale, perché richiede l'estrazione di una radice quadrata di una quantità, $q^2 + \frac{4}{27}p^3$, che deve essere positiva. Paradossalmente, come mostrato nel paragrafo 4.3, il radicando è negativo se e solo se l'equazione ammette tre soluzioni reali. Questa stranezza, definita dai matematici il "casus impossibilis", ha portato di fatto allo sviluppo della teoria dei numeri complessi.

4.5 Formula di Viéte per il "casus impossibilis"

Consideriamo la formula di triplicazione del coseno, che si dimostra facilmente scrivendo il coseno come parte reale dell'esponenziale complessa.

$$\cos(3t) = 4\cos^3 t - 3\cos t$$

Moltiplicando per $2r^3$, dove r è un parametro reale non nullo, si ottiene

$$(2r\cos t)^3 - 3r^2(2r\cos t) - 2r^3\cos(3t) = 0$$

In altri termini, $x = 2r\cos t$ risolve l'equazione di terzo grado parametrica

$$x^3 - 3r^2 x - 2r^3\cos(3t) = 0$$

Ora, è possibile ricondurre un'equazione di terzo grado del tipo $x^3 + px + q = 0$ al caso precedente se esistono $r \neq 0$ e t reali tali che

$$\begin{cases} p = -3r^2 \\ q = 2r^3\cos(3t) \end{cases} \longrightarrow \begin{cases} r = \sqrt{-p/3} \\ \cos(3t) = q/2r^3 = -\frac{q}{2}\left(\frac{p}{3}\right)^{3/2} \end{cases}$$

La seconda equazione è risolubile sse la RHS è minore in modulo di 1, in modo che possa essere un argomento assunto da un coseno. Deve valere allora

$$\frac{q}{2}\left(\frac{p}{3}\right)^{-3/2} \leq 1$$

Elevando tutto al quadrato

$$\frac{q^2}{4}\left(\frac{p}{3}\right)^{-3} \leq 1 \quad \longrightarrow \quad \frac{q^2}{4} + \left(\frac{p}{3}\right)^3 \leq 0$$

La condizione trovata è complementare alla condizione della formula di Cardano.
Si è dunque trovata una formula risolutiva delle equazioni di terzo grado, che si differenzia a seconda del segno di $\frac{q^2}{4} + \left(\frac{p}{3}\right)^3$.

$\square$

5 Sviluppi polinomiali

Si dice sviluppo polinomiale di grado n di una funzione $f(x)$ in un punto x_0 un polinomio p(x) tale che

$$f(x) = p(x) + o\left((x - x_0)^n\right)$$

Non è necessario che il polinomio abbia effettivamente grado n, basta che la differenza sia un "o piccolo" di grado n.
Da ciò segue una definizione alternativa della differenziabilità: una funzione ammette derivata pari a m in un punto x_0 se e solo se esiste uno sviluppo polinomiale di grado 1 della funzione,[2] del tipo

$$f(x) = f(x_0) + m(x - x_0) + o(x - x_0)$$

L'analogia porta, senza sforzi eccessivi, a dire che per gli sviluppo polinomiali valgono tutte le regole note per le derivate (somma, prodotto, funzione inversa, composizione. . .).
Generalizzando ulteriormente la notazione, si può dire che due funzioni hanno in x_0 un "contatto" di ordine n se la loro differenza è un "o piccolo" di $(x - x_0)^n$. In tal senso, una funzione è differenziabile se ammette una funzione lineare affine $f(x_0) + m(x - x_0)$ con cui ha un contatto di ordine 1.

5.1 Proprietà degli sviluppi polinomiali

Lemma 1 Dato un polinomio $p(x)$ e un punto x_0, è equivalente dire che

1. $(x - x_0)^{n+1} | p(x)$
2. $p(x) = \mathcal{O}((x - x_0)^{n+1})$ per $x \to x_0$
3. $p(x) = o((x - x_0)^n)$ per $x \to x_0$

Questo perché, se un polinomio è della forma $p(x) = (x - x_0)^{n+1}q(x)$, allora per $x \to x_0$ il suo ordine di infinitesimo è almeno $(x - x_0)^{n+1}$, e quindi tende a zero più rapidamente di $(x - x_0)^n$. Viceversa, se vale che$p(x) = o((x - x_0)^n)$, allora tutti i termini fino all'n-esimo compreso devono annullarsi, altrimenti andrebbe a zero più lentamente. Dunque deve valere che $(x - x_0)^{n+1} | p(x)$.

Corollario Un polinomio di grado n è uno sviluppo polinomiale di grado minore o uguale a n.

Unicità dello sviluppo polinomiale Per ogni $n \in \mathbb{N}$, lo sviluppo polinomiale di ordine n di una funzione $f(x)$ in un punto x_0, se esiste, è unico. In altri termini, una funzione può ammettere una sola scrittura come

$$f(x) = p(x) + o\left((x - x_0)^n\right)$$

Dove p è un polinomio di grado al massimo n. Infatti, siano p_1 e p_2 due sviluppi polinomiali di ordine n. Allora

$$f(x) = p_1(x) + o\left((x - x_0)^n\right) = p_2(x) + o\left((x - x_0)^n\right)$$

Separando i termini si ha che

$$p_1(x) - p_2(x) = o\left((x - x_0)^n\right)$$

Per il lemma 1, il polinomio $p_1(x) - p_2(x)$ è diviso da $(x - x_0)^{n+1}$. Tuttavia, il suo grado è minore o uguale del massimo tra i gradi dei due polinomi, cioè n. Ciò significa che $p_1(x) - p_2(x)$ deve essere identicamente nullo, e cioè che i due sviluppi coincidono.

□

Corollario Lo sviluppo in serie di potenze di una funzione, se esiste, è unico, perché la sua restrizione all'ordine n è unica $\forall\, n \in \mathbb{N}$.

[2] Sottolineo che ciò non vale per le derivate successive: ammettere uno sviluppo polinomiale di ordine n non significa per forza essere derivabile n volte nel punto, perché è possibile che le derivate non esistano. Ad esempio, la funzione $y = x^2 \sin\left(1/x^3\right)$ ammette uno sviluppo di ordine 2, ma non ha derivata seconda in zero perché la derivata prima non è continua. Questa osservazione è fondamentale per capire perché sia necessario il teorema 5.8.

Lemma 2 Se una funzione f è derivabile n volte in $x = x_0$ e tutte le derivate fino all'n-esima sono nulle, allora $f = f(x_0) + o((x - x_0)^n)$ per $x \to x_0$. La dimostrazione procede per induzione su n.

- Se $n = 1$, basta considerare la definizione della derivata di f come uno sviluppo polinomiale di grado 1:
$$f(x) = f(x_0) + 0(x - x_0) + o(x - x_0) = f(x_0) + o(x - x_0)$$

- Passiamo ora al passo induttivo, supponendo la tesi vera per $n = k$. Se la funzione è derivabile $k+1$ volte, la derivata prima, che indicheremo con $f'(x)$, è derivabile k volte, con k derivate nulle. Applicando l'ipotesi induttiva alla derivata prima, otteniamo che
$$f'(x) = f'(x_0) + o((x - x_0)^k) = o((x - x_0)^k)$$
Assumiamo WLOG che $x < x_0$, tanto il ragionamento è simmetrico. Ora, per teorema di Lagrange applicato a f esiste un punto $c \in (x, x_0)$ tale che
$$f(x) = f(x_0) + f'(c)(x - x_0) = f(x_0) + o((c - x_0)^k)(x - x_0)$$
Da cui, poiché $x < c < x_0$, segue la tesi.
$$f(x) = f(x_0) + o((x - x_0)^{k+1})$$

□

5.2 Polinomio di Taylor

Data una funzione $f(x)$, derivabile n volte in $x = x_0$, si definisce polinomio di Taylor di ordine n e centro x_0 un polinomio di grado minore o uguale a n che assume le stesse derivate di $f(x)$ in x_0. Il polinomio è della forma

$$T[x_0,\, f,\, n](x) = \sum_{k=0}^{n} \frac{f^k(x_0)(x - x_0)^k}{k!}$$

Si osserva che il polinomio è uno sviluppo polinomiale di $f(x)$ di ordine $n+1$, cosa che segue applicando il lemma 2 del paragrafo 5.1 alla differenza $f(x) - T[x_0,\, f,\, n](x)$, che per ipotesi ha le prime n derivate nulle. Da ciò segue anche la formula del resto di Peano:

$$f(x) = T[x_0,\, f,\, n](x) + o((x - x_0)^n) \quad \text{per } x \to x_0$$

Essendo uno sviluppo polinomiale, è anche l'unico sviluppo di ordine minore o uguale a $n+1$, per quanto osservato in precedenza. Rileggendo quanto detto in altri termini, esiste un unico polinomio di Taylor di ordine n, per unicità dello sviluppo polinomiale.
I coefficienti dello sviluppo dipendono dal centro, ma in generale possono variare di molto e in modo discontinuo. Invece, si può dire che lo sviluppo di Taylor è lineare rispetto all'argomento, ossia
$$T[x_0,\, af + bg,\, n] = aT[x_0,\, f,\, n] + bT[x_0,\, g,\, n]$$

5.3 Resto di Lagrange

Data una funzione f derivabile $n+1$ volte in un intorno di x_0, la differenza tra la funzione e il suo polinomio di Taylor di ordine n vale

$$f(x) - T[x_0,\, f,\, n](x) = \frac{f^{(nk+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$$

Dove $c \in (x,\, x_0)$. Ciò significa che è possibile stimare il resto con il massimo della funzione $f^{(n+1)}(c)$ tra x e x_0.
Si osservi che dal teorema discende come corollario il teorema di Lagrange, che equivale al caso $n = 1$.

Dimostrazione Denotiamo con γ il coefficiente numerico del resto di Lagrange, ossia

$$f(x) = T[x_0,\, f,\, n](x) + \gamma(x)\,(x - x_0)^{n+1} \quad (\star)$$

L'obiettivo è mostrare che $\gamma(x) = \frac{f^{(n+1)}(c)}{(n+1)!}$ per un certo $c \in (x, x_0)$. Esplicitando otteniamo

$$\gamma(x) = \frac{f(x) - T[x_0,\, f,\, n](x)}{(x - x_0)^{n+1}}$$

Definiamo ora una nuova funzione

$$g(t) = T[t,\, f,\, n](x) + \gamma(x)\,(x - t)^{n+1}$$

Praticamente stiamo facendo variare il centro dello sviluppo del polinomio di Taylor. La nuova funzione è continua, perché $\gamma(x)$ è costante se si fissa x. Si osserva che $g(x) = f(x)$, in quanto tutti i termini del polinomio calcolato in x sono nulli (perché si annulla il termine $(x - x)$), eccetto il primo che vale $f(x)$. Inoltre, $g(x_0) = f(x)$, per quanto visto in $(\star)$. Allora $g(x) = g(x_0)$. Dunque vale il teorema di Rolle: esiste un punto $c \in (x, x_0)$ in cui la derivata della funzione, rispetto a t, si annulla. Ora, calcoliamo separatamente le derivate delle due parti di $g(t)$.
Deriviamo prima $\gamma(x)\,(x - t)^{n+1}$.

$$\frac{d}{dt}\gamma(x)\,(x - t)^{n+1} = -(n+1)\gamma(x)\,(x - t)^n$$

Ora, passiamo a derivare il polinomio di Taylor rispetto al centro.

$$\frac{d}{dt}T[t,\, f,\, n](x) = \frac{d}{dt}\sum_{k=0}^{n} \frac{f^{(k)}(t)\,(x - t)^k}{k!} = \sum_{k=0}^{n} \frac{f^{(k+1)}(t)\,(x - t)^k}{k!} - \frac{f^{(k)}(t)\,(x - t)^{k-1}}{(k-1)!}$$

Dato che la somma si telescopizza, resta solo il termine finale. Il termine iniziale non c'è perché, quando l'esponente di $(x - t)$ è nullo, si sta derivando una costante, ottenendo zero. Allora

$$\frac{d}{dt}T[t,\, f,\, n](x) = \frac{f^{(n+1)}(t)\,(x - t)^n}{n!}$$

Per teorema di Rolle, esiste c tale che

$$g'(c) = T'[c,\, f,\, n](x) - \gamma(x)\,(n+1)(x - c)^n = \frac{f^{(n+1)}(c)\,(x - c)^n}{n!} - \gamma(x)\,(n+1)(x - c)^n = 0$$

Allora otteniamo

$$\gamma(x) = \frac{f^{n+1}(c)}{(n+1)!}$$

Che è quanto volevamo dimostrare.

□

5.4 Resto integrale

Un'altra formula per il resto dello sviluppo n-esimo, che fa uso degli integrali, mi sembra assai graziosa. Si ha che

$$f(x) - T[x_0,\, f,\, n](x) = \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)\,(x - t)^n\,dt$$

Dimostrazione Si procede per induzione.

- Per $n = 0$ la formula segue dal teorema fondamentale del calcolo integrale:

$$f(x) - T[x_0,\, f,\, 0](x) = f(x) - f(x_0) = \int_{x_0}^{x} f'(t)\,dt$$

- Se è vero per $n = k$, sviluppando il resto $k+1$-esimo come somma del resto precedente e di un fattore in più otteniamo:

$$f(x) - T[x_0,\, f,\, k+1](x) = f(x) - \left(\sum_{i=0}^{k} \frac{f^{(i)}}{i!}(x-x_0)^i \right) - \frac{f^{(k+1)}}{(k+1)!}(x-x_0)^{k+1} =$$
$$= \frac{1}{k!} \int_{x_0}^{x} f^{(k+1)}(t)\,(x-t)^k\, dt - \frac{f^{(k+1)}}{(k+1)!}(x-x_0)^{k+1}$$

Osserviamo che la formula ottenuta è esattamente quella cercata. Infatti, integrando per parti (derivando $(x-t)^{k+1}$ e integrando $f^{(k+2)}$) otteniamo

$$\frac{1}{(k+1)!} \int_{x_0}^{x} f^{(k+2)}(t)\,(x-t)^{k+1}\, dt = \frac{1}{k!} \int_{x_0}^{x} f^{(k+1)}(t)\,(x-t)^k\, dt + \left(\frac{1}{(k+1)!} f^{(k+1)}(t)\,(x-t)^{k+1} \right)_{x_0}^{x}$$
$$= \frac{1}{k!} \int_{x_0}^{x} f^{(k+1)}(t)\,(x-t)^k\, dt - \frac{f^{(k+1)}}{(k+1)!}(x_0)\,(x-x_0)^{k+1}$$

Che è esattamente quanto trovato.

□

La formula integrale del resto, sebbene richieda che il prodotto $f^{(n+1)}(t)\,(x-t)^n$ sia integrabile (cosa che comunque è generalmente vera), è strettamente più potente del resto di Lagrange, perché non fa riferimento a un terzo punto c la cui posizione è ignota. Di fatto, se il resto integrale fornisce un valore esatto, quello di Lagrange permette solo una sovrastima con $\max_{c \in (x, x_0)} f^{(n+1)}(c)$, che è poi quella che rimane nella generalizzazione a più variabili (vedasi 10.2). Questa pecca verrà ribadita nel paragrafo 5.6.

5.5 Derivate del polinomio di Taylor

Dalla dimostrazione del resto di Lagrange segue che la derivata del polinomio di Taylor rispetto al centro vale

$$\frac{d}{dt} T[t,\, f,\, n](x) = \frac{f^{(n+1)}(t)\,(x-t)^n}{n!}$$

É inoltre facile osservare che la derivata rispetto a x del polinomio di Taylor di grado n è il polinomio di Taylor di grado $n-1$ della derivata della funzione.

$$\frac{d}{dx} T[x_0,\, f,\, n](x) = T[x_0,\, f',\, n-1](x)$$

5.6 (*)Quanto è buono il polinomio di Taylor?

Detto meglio, se una funzione è derivabile infinite volte (ossia è di classe C^∞), è possibile approssimarla a piacere con un suo sviluppo di Taylor di centro un qualche x_0?[3] In generale non su tutto $\mathbb{R}$. Distinguiamo un po' di casi rilevanti.

Esponenziale, seno, coseno Queste funzioni sono proprio definite come delle serie di potenze, ossia di fatto come il loro sviluppo in serie di Taylor nel punto $x = 0$. Quindi l'approssimazione è perfetta ovunque.

[3] Una funzione che in ogni punto ammette un intorno in cui coincide col suo sviluppo di Taylor si dice analitica. Le serie di potenze, l'esponenziale e $(1-x)^{-1}$ ne sono degli esempi. Tuttavia, non tutte le funzioni C^∞ sono analitiche. Si rimanda al paragrafo5.7.

$(1-x)^{-1}$ Osservando che, per $x < 1$, la funzione è il limite della serie geometrica, si deve avere che il suo sviluppo di Taylor è la serie geometrica stessa. Il resto di Lagrange della ridotta n-esima vale

$$\left|(1-x)^{-1} - \sum_{k=0}^{n} x^k\right| = \left|\left(\frac{d^{n+1}}{dc^{n+1}}(1-c)^{-1}\right)\frac{x^{n+1}}{(n+1)!}\right| = \left|(1-c)^{-(n+2)}x^{n+1}\right|$$

Ora, dato che stiamo in sviluppando in $x_0 = 0$, avremo $0 < |c| < |x| < 1$. Dunque, stimando dall'alto il resto, avremo

$$\left|\frac{x^{n+1}}{(1-c)^{n+2}}\right| = \left|\frac{1}{1-c}\frac{x^{n+1}}{(1-c)^{n+1}}\right| \le \left|\frac{1}{1-c}\right|\left(\frac{|x|}{1-|x|}\right)^{n+1}$$

Ora, il resto diminuisce a piacere al crescere di n se $\frac{|x|}{1-|x|} < 1$, cosa che avviene sse $|x| < 1/2$. Ciò significa che l'approssimazione della funzione con il suo sviluppo di Taylor in $x_0 = 0$ è piccola a piacere solo nell'intervallo $(-1/2, 1/2)$. La cosa buffa è che, se avessimo sviluppato il resto come la coda della serie geometrica, avremmo ottenuto che

$$\left|(1-x)^{-1} - \sum_{k=0}^{n} x^k\right| = \left|\frac{1}{1-x} - \frac{1-x^{n+1}}{1-x}\right| = \left|\frac{x^{n+1}}{1-x}\right|$$

Il quale è piccolo a piacere già se $|x| < 1$. Mai fidarsi del resto di Lagrange, che come si è detto in precedenza è solo una sovrastima, e anche abbastanza rozza.

Arcotangente Calcoliamo lo sviluppo di Taylor dell'arcotangente. La derivata prima vale

$$\frac{d\arctan x}{dx} = \frac{1}{1+x^2} = \frac{i}{2}\left(\frac{1}{x+i} - \frac{1}{x-i}\right)$$

Questo rende tutto molto più semplice perché è possibile derivare i due pezzi separatamente. Calcoliamo le derivate successive:

$$D^K(\arctan x) = D^{K-1}\left(\frac{i}{2}\left(\frac{1}{x+i} - \frac{1}{x-i}\right)\right) = (-1)^{K-1}\frac{i(K-1)!}{2}\left(\frac{1}{(x+i)^{K-1}} - \frac{1}{(x-i)^{K-1}}\right)$$

Calcoliamo lo sviluppo di Taylor in $x_0 = 0$. Per come si sviluppano le potenze di i, i termini di grado pari si annullano. Avremo allora che

$$T[0, \arctan, 2n+1](x) = \sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{2k+1}$$

Ora, mostriamo che per $|x| \le 1$ l'approssimazione dell'arcotangente col suo polinomio di Taylor è piccola a piacere. WLOG, consideriamo il caso $x > 0$.Sia $R(x)$ il resto. Ovviamente vale $R(0) = 0$. Allora, applicando al resto il teorema di Lagrange, abbiamo:

$$R(x) = R(x) - R(0) = xR'(c) \quad \text{con } 0 < c < x \le 1$$

Ora, il resto vale

$$R(x) = \arctan x - \sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{2k+1}$$

Dunque, derivando, otteniamo

$$R(x) = xR'(c) = x\left(\frac{1}{1+c^2} - \sum_{k=0}^{n}(-1)^k c^{2k}\right) = x\left(\frac{1}{1+c^2} - \frac{1+(-c)^{2n+2}}{1+c^2}\right) = x\frac{(-c)^{2n+2}}{1+c^2}$$

Se $|c| < 1$, il resto è arbitrariamente piccolo.

□

Si osserva che, dallo sviluppo dell'arcotangente per $x = 1$, otteniamo una nuova e più rapida approssimazione di π:[4]

$$4 \arctan 1 = \pi = 4 \sum_{k=0}^{+\infty} (-1)^k \frac{1}{2k+1} = \left(4 - \frac{4}{3} + \frac{4}{5} - \frac{4}{7} \dots\right)$$

□

5.7 Funzioni analitiche

Funzione analitica Una funzione si dice analitica se in ogni punto x_0 ammette un intorno in cui coincide con la sua serie di Taylor. In altri termini, esiste un intorno in cui il resto $R_n(x)$ è infinitesimo per $n \to \infty$ uniformemente in x.
Esempi di funzioni analitiche sono l'esponenziale e le serie di potenze. Per queste ultime sembra un'ovvietà, ma sottolineo che essere analitica significa poter essere sviluppata come serie di potenze rispetto a qualsiasi centro x_0, non solo rispetto a $x = 0$. In altri termini, le serie di potenze sono analitiche perché possono essere ricentrate in un qualsiasi punto x_0.
Le funzioni analitiche formano un'algebra, ossia sono chiuse per somma, prodotto e moltiplicazione per uno scalare. La verifica è abbastanza semplice, basta scrivere tutto come serie di potenze. É inoltre ovvio che se $f(x)$ è analitica lo è anche ogni suo traslato $f(x + \lambda)$ e lo sono le sue riflessioni rispetto agli assi cartesiani $-f(x)$ e $f(-x)$.

Le funzioni razionali sono analitiche. Per dimostrare la proposizione basta far vedere che $1/(1 - x)$ è una funzione analitica. Infatti, ogni funzione razionale si ottiene moltiplicando un polinomio (evidentemente analitico) e un denominatore fatto di monomi del tipo $(\lambda - x)$ che corrispondono a traslati della funzione $1/(1 - x)$. Inoltre, senza perdita di generalità ci limitiamo al caso in cui $x < 1$, dato che un ramo dell'iperbole si ottiene dall'altro tramite due riflessioni.
Sia $x_0 = 1 - R$, con $R > 0$. Sia $I = (x_0 - R, 1)$ l'intorno in cui mostreremo la validità della tesi, ossia che $\forall x \in I$ vale

$$f(x) = \frac{1}{1-x} \stackrel{?}{=} \sum_{k=0}^{+\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

Consideriamo le derivate successive della funzione. Abbiamo che

$$f^{(k)}(x) = \frac{d^k}{dx^k}(1-x)^{-1} = k!(1-x)^{-k-1}$$

Dunque possiamo sostituire e ottenere

$$\sum_{k=0}^{+\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \sum_{k=0}^{+\infty} \frac{k!(1-x_0)^{-k-1}}{k!}(x - x_0)^k = \frac{1}{1-x_0} \sum_{k=0}^{+\infty} \left(\frac{x - x_0}{1 - x_0}\right)^k$$

Dato che, per come si è scelto I, $|x - x_0| < |1 - x_0|$, possiamo sviluppare come serie geometrica

$$\sum_{k=0}^{+\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \frac{1}{1-x_0} \frac{1}{1 - \frac{x-x_0}{1-x_0}} = \frac{1}{1-x}$$

Come volevasi dimostrare.

□

[4] In realtà, non sarebbe lecito assumere che il limite per $x \to 1$ dello sviluppo coincida con il valore della funzione in 1, perché finora si è visto solo che lo sviluppo converge assolutamente per $|x| < 1$. Per concludere servono un paio di risultati sulle serie di potenze, tra cui il lemma di Abel. Esso afferma che se una serie di potenze converge in un punto $x = r$ (nel nostro caso, nel punto $x = 1$, dove sappiamo che l'arcotangente fa $\pi/4$), allora la successione delle somme parziali, come funzioni di x, converge uniformemente sul compatto $[0, r]$. Questo significa che la serie è continua, dato che è limite uniforme di funzioni continue, e dunque che il limite della serie è il valore assunto nel limite.

5.8 Caratterizzazione delle funzioni C^n

Una funzione $f : [a, b] \in \mathbb{R}$ è di classe C^n, ossia è derivabile n volte con derivate continue, se e solo se ammette in ogni punto t uno sviluppo polinomiale di ordine n, i cui coefficienti e il cui resto sono funzioni continue nel centro dello sviluppo t, ossia

$$f(t+h) = \sum_{k=0}^{n} \frac{a_k(t)}{k!} h^k + h^n R(t, h)$$

Con $a_k(t)$ continua in t, $R(t, h)$ continua in entrambe le variabili e infinitesima per $h \to 0$.

Dimostrazione

$\Longrightarrow$ Se f è di classe C^n, ammette uno sviluppo di Taylor in x di ordine $n-1$ della forma

$$f(t+h) = \sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!} h^k + \frac{f^{(n)}(t+c)}{n!} h^n$$

Per un certo $0 < c < h$. A questo sviluppo sommiamo e sottraiamo $\frac{f^{(n)}(t)}{n!} h^n$ in modo da completare lo sviluppo. Otteniamo:

$$f(t+h) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!} h^k + \frac{f^{(n)}(t+c) - f^{(n)}(t)}{n!} h^n$$

I termini dello sviluppo sono continui per ipotesi. Per continuità della derivata n-esima, inoltre, l'ultimo termine è infinitesimo per $h \to 0$. Resta solo da mostrare che il resto è continuo sia in t, cosa ovvia, che in h. Questo è vero, in quanto basta applicare il teorema di Lagrange e ottenere $f^{(n)}(t+c) = \frac{f^{(n-1)}(t+h) - f^{(n-1)}(t)}{h}$, che è una funzione continua in entrambe le variabili.

$\Longleftarrow$ Sia f una funzione che ammette uno sviluppo continuo del tipo

$$f(t+h) = \sum_{k=0}^{n} \frac{a_k(t)}{k!} h^k + h^n R(t, h)$$

Osserviamo subito che $a_0(t) = f(t)$ banalmente. Il nostro obiettivo è mostrare che le derivate della funzione sono i coefficienti, e per farlo basta mostrare che $a_k = a'_{k-1}$ per ogni k.
I coefficienti, essendo continui, ammettono delle primitive per TFCI. Sia allora $F'_k = a_k$. Fissiamo ora due punti x, y e consideriamo la funzione

$$\phi(h) = F_0(y+h) - F_0(x+h)$$

Osserviamo che $F_0 = \int a_0 = \int f$. Dato che la scrittura $f(t+h)$ è perfettamente simmetrica rispetto allo scambio di t e h, integrando in dt e in dh otterremo due scritture diverse, che dovranno però essere sostanzialmente uguali. Dunque, otterremo due diversi sviluppi polinomiali per ϕ in $h = 0$, che dovranno coincidere per unicità dello sviluppo polinomiale.

$$\phi(h) = F_0(y+h) - F_0(x+h) = \int_x^y f(t+h)\, dt = \int_x^y \sum_{k=0}^{n} \frac{a_k(t)}{k!} h^k + h^n R(t, h)\, dt =$$

$$= \sum_{k=0}^{n} \frac{F_k(y) - F_k(x)}{k!} h^k + h^n \int_x^y R(t, h)\, dt$$

Ora, la funzione $R(t, h)$ è uniformemente continua su un dominio compatto sufficientemente grande. Dunque ammette un modulo di continuità $\omega(\delta)$. Allora vale che

$$|R(t, h)| = |R(t, h) - R(t, 0)| \leq \omega(h)$$

Ossia i valori di $R(t+h)$ sono infinitesimi per $h \to 0$ in modo uniforme. Allora è possibile stimar l'ultimo integrale con il suo massimo e osservare che questo è un infinitesimo per $h \to 0$. Otteniamo dunque

$$\phi(h) = \sum_{k=0}^{n} \frac{F_k(y) - F_k(x)}{k!} h^k + h^n o(1)$$

Ora, partiamo dalla scrittura di ϕ e cerchiamo un altro sviluppo, questa volta integrando rispetto a h.

$$\phi(h) = \phi(0) + \int_0^h \phi'(s)\, ds = F_0(y) - F_0(x) + \int_0^h (f(y+s) - f(x+s))\, ds =$$

$$= F_0(y) - F_0(x) + \sum_{k=0}^{n} \frac{a_k(y) - a_k(x)}{k!} \int_0^h s^k\, ds + \int_0^h s^n (R(y,s) - R(x,s))\, ds$$

Ora, invece di integrare subito la parte di sinistra, isoliamo il termine di grado n dalla sommatoria e definiamo la seguente funzione, continua in entrambe le variabili e nulla in $s = 0$:

$$S(t,s) = s\left(\frac{a_n(t)}{n!} + R(t,s)\right)$$

Sostituendo otteniamo

$$\phi(h) = F_0(y) - F_0(x) + \sum_{k=0}^{n-1} \frac{a_k(y) - a_k(x)}{(k+1)!} h^{k+1} + \int_0^h s^{n-1} s\left(\frac{a_n(y) - a_n(x)}{n!} + R(y,s) - R(x,s)\right) ds =$$

$$= F_0(y) - F_0(x) + \sum_{k=0}^{n-1} \frac{a_k(y) - a_k(x)}{(k+1)!} h^{k+1} + \int_0^h s^{n-1} (S(y,s) - S(x,t))\, ds$$

Ora, stimiamo il modulo dell'integrale con il massimo dell'integrando, tenendo x e y fissati:

$$\left| \int_0^h s^{n-1} (S(y,s) - S(x,t))\, ds \right| \leq h^n \left(\max_{s \leq h} S(y,s) + \max_{s \leq h} S(x,s) \right)$$

La scrittura tra parentesi è infinitesima per $h \to 0$, sempre perché il resto R era uniformemente continuo e infinitesimo. Allora otteniamo

$$\phi(h) = F_0(y) - F_0(x) + \sum_{k=0}^{n-1} \frac{a_k(y) - a_k(x)}{(k+1)!} h^{k+1} + h^n o(1)$$

Scalando gli indici otteniamo:

$$\phi(h) = F_0(y) - F_0(x) + \sum_{k=1}^{n} \frac{a_{k-1}(y) - a_{k-1}(x)}{k!} h^k + h^n o(1)$$

Compariamo i coefficienti dei due sviluppi ottenuti. Abbiamo che

$$\phi(h) = \sum_{k=0}^{n} \frac{F_k(y) - F_k(x)}{k!} h^k + h^n o(1) = F_0(y) - F_0(x) + \sum_{k=1}^{n} \frac{a_{k-1}(y) - a_{k-1}(x)}{k!} h^k + h^n o(1)$$

Da cui si ricava

$$F_k(y) - F_k(x) = a_{k-1}(y) - a_{k-1}(x)$$

Ora, dividendo per $y - x$ e mandando al limite $y \to x$ otteniamo la derivata di entrambe le scritture, ossia

$$a_k(x) = a'_{k-1}(x)$$

Che è quanto volevamo dimostrare.

□

Corollario Una funzione è di classe C^∞, ossia di classe C^n per ogni n, se e solo se ammette in ogni punto uno sviluppo in serie di Taylor, i cui coefficienti sono continui nel centro dello sviluppo.

5.9 Interpolazione polinomiale di Hermite

Supponiamo che siano assegnati n punti $(x_i)_{i=1,\dots,n}$ e, per ciascuno di essi, una stringa di numeri $(a_{ij})_{j=0,\dots,k_i-1}$, non per forza tutte della stessa lunghezza. Vogliamo dimostrare che esiste un unico polinomio $P(x)$ di grado strettamente minore di $\sum_{i=1}^n k_i$ tale che

$$\frac{d^j}{dx^j}P(x_i) = a_{ij}$$

Ossia, dati dei punti e dei valori a caso, vogliamo mostrare che esiste un unico polinomio che ha quei valori come derivate nei corrispondenti punti.
Innanzitutto, $P_i(x) = \sum_{j=0}^{k_i-1} \frac{a_{ij}}{j!}x^j$ deve essere il polinomio di Taylor di grado $k_i - 1$ di $P(x)$ nel punto x_i. In altri termini, si deve avere che

$$P(x) - P_i(x) = o(x - x_i)^{k_i-1}$$

Ma la differenza tra due polinomi è un polinomio. Allora, per quanto visto in 5.1, la differenza deve essere un multiplo di $(x - x_i)^{k_i}$. In altri termini, possiamo dire che

$$P(x) \equiv P_i(x)\,(mod(x - x_i)^{k_i})$$

Dove la notazione è quella delle congruenze su un anello euclideo (in questo caso, $\mathbb{R}[x]$). Abbiamo così un sistema di n equazioni modulari del tipo

$$\begin{cases} P(x) \equiv P_1(x)\,(mod(x - x_1)^{k_1}) \\ \dots \\ P(x) \equiv P_n(x)\,(mod(x - x_n)^{k_n}) \end{cases}$$

Il teorema cinese del resto, nella sua versione per gli anelli euclidei, asserisce che se i moduli sono coprimi a coppie la soluzione esiste ed è unica, modulo il prodotto $\prod_{i=1}^n (x - x_i)^{k_i}$. In particolare, esiste un'unica soluzione di grado minore della somma dei gradi dei moduli, in questo caso $\sum_{i=1}^n k_i$. La tesi risulta dunque dimostrata.[5]

$\square$

[5] Per chi non fosse convinto, rimando a https://en.wikipedia.org/wiki/Chinese_remainder_theorem.

6 Disuguaglianze e identità

6.1 Disuguaglianza delle medie

La media aritmetica di n numeri reali positivi è maggiore della loro media geometrica. L'uguaglianza vale se e solo se tutti i termini sono uguali.

$$\frac{1}{n}\sum_{k=1}^{n} a_n \geq \sqrt[n]{\prod_{k=1}^{n} a_n}$$

Dimostrazione Si ringrazia Giulio Grammatica per la seguente dimostrazione alternativa.
É facile osservare che se tutti i termini sono uguali vale l'uguaglianza. Se non sono tutti uguali, esiste almeno un termine minore della media aritmetica e almeno uno maggiore. Sia $a_i < AM$ e $a_j > AM$. Sostituiamo a_i con AM, a_j con $a_j + a_i - AM$. Si osserva che sotto questa trasformazione la media aritmetica rimane invariata. Invece, la media geometrica cresce. in quanto

$$a_i * a_j < AM * (a_j + a_i - AM) \quad \longrightarrow \quad a_i * (a_j - AM) < AM * (a_j - AM)$$

Dunque la media geometrica cresce fino a che i termini sono tutti uguali alla media aritmetica. Perciò la media geometrica doveva essere minore.

□

6.2 (*)Teorema di Nicomaco

Il teorema afferma che la somma dei primi n cubi è il quadrato della somma dei primi n interi.

$$\sum_{k=1}^{n} k^3 = \left(\sum_{k=1}^{n} k\right)^2$$

Dimostrazione Si procede per induzione. Il caso $n = 1$ è banalmente vero. Per quanto riguarda il passo induttivo, abbiamo che

$$\sum_{k=1}^{n} k^3 = \sum_{k=1}^{n-1} k^3 + n^3 = \left(\sum_{k=1}^{n-1} k\right)^2 + n^3 = \frac{n^2(n-1)^2}{4} + n^3 = \frac{n^2(n+1)^2}{4} = \left(\sum_{k=1}^{n} k\right)^2$$

□

6.3 Partial Fraction Decomposition - Esistenza

Anche nota col nome grossolano di "decomposizione in fratte semplici", è il procedimento per spezzare una funzione razionale $\frac{p(x)}{q(x)}$ in una somma di termini di grado inferiore. Innanzitutto, ci riduciamo al caso $\deg p < \deg q$, dato che è possibile eseguire la divisione euclidea tra numeratore e denominatore.
Supponiamo ora che il denominatore si scriva come $q = ab$, con $\gcd(a, b) = 1$. Allora vale l'identità di Bézout $ac + bd = 1$, per certi polinomi c, d. Moltiplicando per $\frac{p}{ab}$ otteniamo una scomposizione in due frazioni distinte.

$$\frac{p}{ab} = \frac{pc}{b} + \frac{pd}{a}$$

Consideriamo ora il caso in cui il denominatore è della forma $q = (x - \lambda)^n$. Una possibile e ovvia identità è del tipo:

$$\frac{p}{(x-\lambda)^n} = \frac{c}{(x-\lambda)^n} + \frac{p-c}{(x-\lambda)^n}$$

Dove c è una costante. Scegliamo c in modo che al termine di sinistra si semplifichi un fattore $(x - \lambda)$. In altri termini, imponiamo che $(x - \lambda)|p(x) - c$. Per teorema di Ruffini, è equivalente dire che $p(\lambda) - c = 0$. Dunque otteniamo che esiste un certo polinomio $q(x) = (p(x) - c)/(x - \lambda)$. Possiamo allora scrivere

$$\frac{p}{(x-\lambda)^n} = \frac{p(\lambda)}{(x-\lambda)^n} + \frac{q(x)}{(x-\lambda)^{n-1}}$$

Nel caso di un campo algebricamente chiuso, questo conclude la trattazione. Abbiamo allora che in $\mathbb{C}$ una frazione razionale si scompone come

$$\frac{p}{q} = \sum_{i,j} \frac{a_{ij}}{(x-\lambda_i)^j} \quad \text{con } a_{ij} \in \mathbb{C}$$

Per comodità, possiamo sommare tra loro i termini con lo stesso λ_i e ottenere

$$\frac{p}{q} = \sum_i \frac{A_i(x)}{(x-\lambda_i)^{n_i}} \quad \text{con } A_i \in \mathbb{C}[x]$$

Nel caso reale, ci sono due modi di procedere. Si può prima ottenere una decomposizione in fattori complessi, e poi sommare tra loro i termini coniugati (cosa che è abbastanza intuitiva, ma andrebbe dimostrata), oppure si può definire un nuovo algoritmo per un denominatore del tipo $(x^2 + rx + q) = (x - \lambda)(x - \bar{\lambda})$, con $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Scegliamo la seconda via. Come prima, sia

$$\frac{p}{(x^2+rx+q)^n} = \frac{a+bx}{(x^2+rx+q)^n} + \frac{p-a-bx}{(x-\lambda)^n(x-\bar{\lambda})^n}$$

Si è sviluppato il secondo denominatore perché si vuole ora imporre che il numeratore sia divisibile per $(x - \lambda)$. Basta questo per concludere, dato che essendo un polinomio a coefficienti reali sarà anche divisibile per $(x - \bar{\lambda})$, con la stessa molteplicità. Abbiamo allora due equazioni, che determinano univocamente a e b in quanto sono evidentemente linearmente indipendenti:

$$\begin{cases} \Re(p(\lambda)) = b\Re(\lambda) + a \\ \Im(p(\lambda)) = b\Im(\lambda) \end{cases}$$

Da cui segue una decomposizione anche per il caso reale.

$$\frac{p}{(x^2+rx+q)^n} = \frac{a+bx}{(x^2+rx+q)^n} + \frac{q(x)}{(x^2+rx+q)^{n-1}}$$

Si potrebbe mostrare che la decomposizione ottenuta è univocamente determinata dall'algoritmo, ma non è così ovvio.

6.4 Partial Fraction Decomposition - Unicità

Sia dato un polinomio a coefficienti complessi $\frac{p(x)}{q(x)}$. Sia $q(x) = \prod_i (x-\lambda_i)^{n_i}$ e sia $q_i(x) = q(x)/(x-\lambda_i)^{n_i}$. Allora la decomposizione in fratte semplici

$$\frac{p}{q} = \sum_i \frac{A_i(x)}{(x-\lambda_i)^{n_i}} \quad \text{con } A_i \in \mathbb{C}[x]$$

è unica, e i coefficienti sono dei polinomi di Taylor di grado $n_i - 1$ delle funzioni $\frac{p}{q_i}$.

$$A_i(x) = T_{n_i-1}\left(\frac{p}{q_i}, x, \lambda_i\right)$$

Dimostrazione

$$\frac{p}{q} = \sum_i \frac{A_i(x)}{(x-\lambda_i)^{n_i}} \quad \longrightarrow \quad p - \sum_i q_i A_i = 0$$

Ora, poiché $p - \sum_i q_i A_i$ è il polinomio identicamente nullo, esso è diviso da ogni fattore di q $(x-\lambda_i)^{n_i}$, il quale divide sicuramente tutti i q_j con $j \neq i$. Allora l'unica cosa da imporre è che divida la somma dei termini in cui non compare, cioè

$$(x-\lambda_i)^{n_i} | (p - q_i A_i) = 0$$

Per quanto si è mostrato in 5.1 la condizione precedente è equivalente[6] a

$$\frac{p}{q_i} = A_i + o((x - \lambda_i)^{n_i - 1}) \quad \text{per } x \to \lambda_i$$

Da cui segue la tesi, essendo unico lo sviluppo polinomiale di una funzione di ordine strettamente minore di n.

$$A_i(x) = T_{n_i - 1}\left(\frac{p}{q_i}, x, \lambda_i\right)$$

□

[6] Piccola imprecisione: stiamo parlando di polinomi a coefficienti complessi, e tutta la teoria degli sviluppi polinomiali è stata fatta con i polinomi reali. Di fatto, tuttavia, non cambia praticamente nulla nella dissertazione.

7 Spazi topologici e spazi metrici

7.1 (*) Spazi topologici

La seguente sezione serve a familiarizzare con alcuni concetti topologici che sono sempre stati lasciati impliciti a lezione, ma che ritengo utili alla comprensione di alcune proprietà delle funzioni e degli spazi metrici.

Topologia Dato un insieme X, si dice topologia su X una famiglia di sottoinsiemi $\tau \in P(X)$ con le seguenti proprietà:

- $\emptyset \in \tau$ e $X \in \tau$;
- τ è chiusa per unione arbitraria, ossia se $A_i \in \tau \ \forall i \in I$, allora $\bigcup_{i \in I} A_i \in \tau$;
- τ è chiusa per intersezione finita, ossia se $A_k \in \tau$ per $k = 1, 2, \ldots, n$, allora $\bigcap_{k=1}^{n} A_k \in \tau$.

Uno spazio topologico si indica spesso con (X, τ). Gli elementi della topologia si dicono aperti. Il complementare di un aperto si dice chiuso.

7.2 Spazi metrici

Uno spazio metrico (X, d) è un insieme munito di una funzione distanza (detta anche, impropriamente, metrica) con le seguenti proprietà, valide $\forall x, y, z \in X$:

- $d : X \times X \longrightarrow [0, +\infty]$;
- Non degenerazione: $d(x, y) = 0 \iff x = y$;
- Simmetria: $d(x, y) = d(y, x)$;
- Disuguaglianza triangolare: $d(x, y) \leq d(x, z) + d(y, z)$.

Palla Si definisce palla di centro x e raggio $r > 0$, e si indica con $B(x, r)$, l'insieme dei punti a distanza minore di r dal punto x. In simboli

$$B(x, r) = \{y \in X \text{ t.c. } d(x, y) < r\}$$

Questa definizione induce su X una topologia in cui gli aperti sono le palle. In altri termini, uno spazio metrico è un particolare spazio topologico, e dunque tutte le definizioni topologiche hanno un corrispondente metrico. Non vale il viceversa: esistono spazi topologici in cui non è possibile definire una distanza.

Aperto Un sottoinsieme di uno spazio metrico si dice aperto se contiene una palla di ogni suo punto. Si osservi che ciò equivale a dire che è un unione arbitraria di palle, cioè di aperti della topologia dello spazio metrico.

7.3 (*) La distanza è una funzione lipschitziana

Consideriamo la distanza da un punto fissato $\bar{x}$. Questa è una funzione $d_x : X \longrightarrow [0, +\infty]$. Dalla disuguaglianza triangolare segue che

$$d(x, y) \leq d(x, z) + d(y, z) \quad \longrightarrow \quad |d(x, y) - d(x, z)| = |d_x(y) - d_x(z)| \leq d(y, z)$$

Il modulo segue dalla simmetria della disuguaglianza triangolare, che rimane vera se si scambiano $d(x, y)$ e $d(x, z)$. L'ultima disuguaglianza dice che la distanza è una funzione lipschitziana di costante 1, e dunque in particolare è uniformemente continua.

7.4 Lunghezza di una curva

Curva Si definisce (arco di) curva una funzione da un intervallo chiuso e limitato $[a,b]$ in uno spazio metrico (X,d).

Suddivisione Si dice suddivisione, o partizione, di un intervallo $[a,b]$ una n-upla finita di numeri reali $a = t_1 < t_2 < \ldots < t_n = b$.

Lunghezza Si definisce la lunghezza di una curva γ rispetto a una suddivisione t_n come

$$L(\gamma, t_n) = \sum_{k=1}^{n-1} d\left(\gamma\left(t_k\right), \gamma\left(t_{k+1}\right)\right)$$

Si definisce la lunghezza di una curva come l'estremo superiore delle lunghezze delle sue partizioni:

$$L(\gamma, [a,b]) = \sup_{t_n} \{L(\gamma, t_n)\}$$

7.5 Proprietà della lunghezza

Diamo qui una sfilza di proprietà, più o meno interessanti, della lunghezza.

- Se $a < c < b$, allora $L(\gamma, [a,b]) = L(\gamma, [a,c]) + L(\gamma, [c,b])$

  Dimostrazione Sia C l'insieme delle spezzate che contengono c. Ovviamente C è un sottoinsieme dell'insieme delle spezzate su [a,b]. Perciò, passando ai sup sui due insiemi, otteniamo che $L(\gamma, [a,b]) \geq \sup C = L(\gamma, [a,c]) + L(\gamma, [c,b])$. Viceversa, osserviamo che data una spezzata, è possibile sempre aggiungerle il punto c, cosa che, per disuguaglianza triangolare, genera una spezzata più lunga (o al massimo della stessa lunghezza). Dunque vale anche $L(\gamma, [a,b]) \leq \sup C = L(\gamma, [a,c]) + L(\gamma, [c,b])$. Dalle due disuguaglianze segue l'uguaglianza.

  □

- Se γ è lipschitziana di costante K, allora la lunghezza della curva è limitata da K volte l'ampiezza dell'intervallo di partenza. In simboli

  $$L(\gamma, [a,b]) \leq K(b-a)$$

  Questo perché ciascuna spezzata ha lunghezza limitata dalla costante K, in quanto

  $$L(\gamma, t_n) = \sum_{k=1}^{n-1} d\left(\gamma\left(t_k\right), \gamma\left(t_{k+1}\right)\right) \leq K \sum_{k=1}^{n-1} t_{k+1} - t_k = K(t_n - t_0) \leq K(b-a)$$

- Se γ è continua, la sua lunghezza può essere calcolata usando solo delle spezzate di punti equispaziati. Cioè:

  $$L(\gamma, [a,b]) = \sup L\left(\gamma, \left\{a + k\frac{b-a}{n},\ \text{con } k = 0, \ldots, n\right\}\right)$$

  Dimostrazione Basta mostrare che è possibile scegliere le spezzate con i vertici in un insieme denso D. Infatti, evidentemente $L(\gamma, [a,b]) \geq \sup L(\gamma, T \subseteq D)$ perché si considerano meno spezzate. D'altra parte, per ogni $\varepsilon > 0$, una qualsiasi spezzata può essere approssimata da una spezzata in D a meno di ε, perché ogni elemento di $[a,b]$ è un punto di accumulazione di D. Dunque i due sup coincidono.

  □

- Se la funzione γ è continua e a variazione limitata, ossia $L(\gamma, [a,b]) < +\infty$, allora anche la funzione lunghezza $L(\gamma, [a,x])$, con $x \in [a,b]$, è continua.

Dimostrazione Mostriamo che $\forall\varepsilon\,\exists\delta$ tc. $t_2 - t_1 < \delta \implies L(\gamma, [t_1, t_2]) \leq 4\varepsilon$. Ciò proverà la continuità uniforme di $L(t)$. Fissato $\varepsilon > 0$, sia $N \in \mathbb{N}$ per cui valgano le seguenti due proprietà:

1. $\exists x_k = a + k\frac{b-a}{N}$ t.c $L\left(\gamma, \{x_k\}_{k\leq N}\right) \geq L - \varepsilon$. Una spezzata siffatta esiste perché, per lemma precedente, la lunghezza di tutta la curva è il sup delle lunghezze delle spezzate equispaziate.
2. $d(\gamma(x_k), \gamma(x_{k+1})) < \varepsilon$. Scegliere gli x_k in tal modo è possibile per la continuità uniforme di γ.

Ora, consideriamo due punti t_1, t_2 a distanza minore di $\frac{b-a}{N}$. Ci sono due possibilità.

1. Se esiste $k \leq N$ tale che $x_k \leq t_1 < t_2 \leq x_{k+1}$, sostituiamo il segmento $\gamma(x_k), \gamma(x_{k+1})$ con i segmenti $\gamma(x_k), \gamma(t_1)$ e $\gamma(t_2), \gamma(x_{k+1})$ e con la lunghezza della curva tra t_1 e t_2. La nuova curva misura al massimo ε in più, dato che la spezzata era già lunga più di $L - \varepsilon$. Allora deve valere che

$$L(\gamma, [t_1, t_2]) + d(\gamma(x_k), \gamma(t_1)) + d(\gamma(t_2), \gamma(x_{k+1})) \leq d(\gamma(x_k), \gamma(x_{k+1})) + \varepsilon$$

In particolare vale che

$$L(\gamma, [t_1, t_2]) \leq d(\gamma(x_k), \gamma(x_{k+1})) + \varepsilon \leq 2\varepsilon$$

2. Supponiamo invece che i due t siano distanziati da qualche x. Poiché la loro distanza è minore di $\frac{b-a}{N}$, si deve avere che $x_k \leq t_1 \leq x_{k+1} \leq t_2 \leq x_{k+2}$. Si ragiona come nel caso precedente, considerando come punto $\bar{t}$ il punto x_{k+1}, e si ottiene che

$$L(\gamma, [t_1, t_2]) = L(\gamma, [t_1, \bar{t}]) + L(\gamma, [\bar{t}, t_2]) \leq 4\varepsilon$$

□

- Una curva $\gamma : [a, b] \longrightarrow \mathbb{R}$ è a variazione limitata se e solo se è la differenza di due funzioni debolmente crescenti. Infatti possiamo esprimere γ rispetto alla sua funzione lunghezza come segue:

$$\gamma(x) = L(\gamma, [a, x]) - (L(\gamma, [a, x]) - \gamma(x))$$

La funzione lunghezza è evidentemente crescente, per quanto già visto, mentre la funzione tra parentesi è crescente perché, presi due punti $x < y$, la lunghezza della curva tra di essi è almeno la differenza in modulo delle loro immagini. Allora

$$L(\gamma, [x, y]) \geq |\gamma(y) - \gamma(x)| \quad \implies \quad L(\gamma, [a, y]) - \gamma(y) \geq L(\gamma, [a, x]) - \gamma(x)$$

Viceversa, siano f e g funzioni debolmente crescenti. La lunghezza della curva $\gamma = f - g$ si stima come

$$L(\gamma, [a, b]) = \sup_{p\in P} \sum_{x_k\in p} |f(x_{k+1}) - f(x_k) - g(x_{k+1}) + g(x_k)| \leq \sup_{p\in P} \sum_{x_k\in p} f(x_{k+1}) - f(x_k) + g(x_{k+1}) - g(x_k) =$$

$$= f(b) - f(a) + g(b) - g(a) < +\infty$$

La proprietà appena vista prende il nome di decomposizione di Jordan.

- La lunghezza di una curva è una seminorma sullo spazio delle funzioni da $[a, b]$ in (X, d), ossia gode delle seguenti due proprietà, di facile verifica:

1. Disugugaglianza triangolare: $L(f + g) \leq L(f) + L(g)$
2. 1-Omogeneità: $L(\lambda f) = |\lambda| L(f)$

Non vale invece la non degenerazione: esistono curve di lunghezza nulla che non sono identicamente nulle. Per esempio, la funzione costante $y = c$ a valori reali ha lunghezza nulla.

7.6 Riparametrizzazioni

- Una funzione $\sigma : [c,d] \longrightarrow [a,b]$ è una parametrizzazione di $[a,b]$ se è bigettiva, continua e strettamente crescente. Vale che la lunghezza è invariante per riparametrizzazioni, ossia $L(\gamma,[a,b]) = L(\gamma\circ\sigma,[c,d])$. Ciò segue dal fatto che si sta facendo il sup sullo stesso insieme.
- Una curva continua $\gamma : [a,b] \longrightarrow \mathbb{R}$ di lunghezza minore di $\mathcal{L}$ può essere riparametrizzata in modo che risulti lipschitziana di costante $\mathcal{L}$.

  Dimostrazione Dobbiamo mostrare che esiste un omeomorfismo $\sigma : [0,1] \longrightarrow [a,b]$ strettamente crescente tale che $\gamma\circ\sigma$ sia lipschitziana di costante $\mathcal{L}$. Osserviamo innanzitutto che è sufficiente riparametrizzare la curva in modo che la funzione lunghezza sia lineare, ossia che $L(\gamma\circ\sigma,[0,x]) = \mathcal{L}x$. Infatti, per definizione di lunghezza, si ha che

  $$|\gamma\circ\sigma(x) - \gamma\circ\sigma(y)| \le L(\gamma\circ\sigma,[x,y]) = \mathcal{L}|x-y|$$

  Ora, la funzione lunghezza originaria è continua (dato che la curva γ è continua e a variazione limitata) e strettamente crescente per costruzione. L'immagine della funzione è l'intervallo $[0,\mathcal{L}]$, dunque basta moltiplicare per $1/\mathcal{L}$ per ottenere una funzione strettamente monotona e continua a valori in $[0,1]$. Tale funzione ammette un'inversa σ, anch'essa continua e strettamente crescente. Si ha dunque che esiste un omeomorfismo strettamente crescente $\sigma : [0,1] \longrightarrow [a,b]$ per cui vale

  $$L(\gamma,[a,y])\circ\sigma(x) = \mathcal{L}x$$

  Allora, noto che $a = \sigma(0)$, otteniamo:

  $$L(\gamma,[\sigma(0),\sigma(x)]) = \mathcal{L}x$$

  Per un lemma precedente, la lunghezza si riparametrizza come

  $$L(\gamma,[\sigma(0),\sigma(x)]) = L(\gamma\circ\sigma,[0,x]) = \mathcal{L}x$$

  Da cui, per quanto detto, segue la tesi.

  □

7.7 Compattezza: Definizioni equivalenti

Insieme compatto Uno spazio topologico (X,τ) si dice compatto (per ricoprimenti) se ogni ricoprimento aperto ammette un sottoricoprimento finito. In altri termini, se una famiglia di aperti A_i è tale che $\bigcup_{i\in I} A_i \supseteq X$, allora è possibile estrarne un numero finito che continua a ricoprire lo spazio.

Insieme compatto per successioni Un sottoinsieme K di uno spazio metrico (X,d)[7] si dice sequenzialmente compatto, o compatto per successioni, se ogni successione a valori in K ammette una sotto-successione convergente in K.
Le due definizioni sono equivalenti negli spazi metrici, anche se per mostrarlo serve introdurre un po' di concetti in più. Nella sezione seguente mostreremo l'equivalenza nel caso reale.

7.8 Teorema di Heine-Borel

In $\mathbb{R}$, è equivalente dire che:

1. Un insieme K è compatto;
2. K è compatto per successioni;
3. K è chiuso e limitato.[8]

[7] Sebbene sia possibile definire le successioni anche in spazi topologici, in generale la mancanza di sovrastrutture come la metrica rende tutto immaneggiabile, quindi ci limiteremo sempre agli spazi metrici a meno che non sia strettamente necessario generalizzare.

[8] In generale, non è vero in spazi generici che un insieme chiuso e limitato è compatto. Si può, ad esempio, mostrare che la palla di raggio 1 di uno spazio vettoriale normato infinito-dimensionale non è compatta.

$(1 \implies 3)$ Per assurdo, non sia limitato. Allora dal ricoprimento aperto $(-n, n)$ non si può estrarre alcun sottoricoprimento finito, perché ci sarà sempre un punto al di fuori di qualsiasi $(-n, n)$. Se invece per assurdo non fosse chiuso, sia x un punto di accumulazione dell'insieme non contenuto in K. consideriamo una successione x_n di elementi di K che tende a x, e senza perdita di generalità supponiamo che sia strettamente crescente. Allora non si può estrarre alcun ricoprimento finito dal seguente ricoprimento aperto:

$$\{(x, +\infty)\} \cup \left\{ \left(-\infty, \frac{x_n + x_{n+1}}{2} \right) \right\}$$

Infatti, comunque se ne prenda un numero finito, ci sarà sempre un x_n escluso.

$(3 \implies 1)$ Dimostriamo prima il caso in cui K è un intervallo del tipo $[a, b]$. Sia $\{U_\lambda\}_{\lambda \in \Lambda}$ un ricoprimento aperto di K. Sia

$$S = \{x \in [a, b] \text{ t.c. } \exists \Lambda_0 \text{ finito t.c. } \bigcup_{\lambda \in \Lambda_0} U_\lambda \supseteq [a, x]$$

Osserviamo innanzitutto che $a \in S$, perché è sempre possibile trovare un aperto U contenente a. Dato che è un aperto, deve contenere una palla (supponiamola chiusa, ossia contenente anche gli estremi) di centro a e raggio un certo δ, da cui anche $a + \delta \in S$. Ciò mostra che esiste almeno un altro elemento in S diverso da a.
Ora, sia $s = \sup S$. Per quanto osservato $a < s \leq b$. Sia U_s un aperto del ricoprimento che contiene s. Per lo stesso motivo, esiste un certo ε tale che $[s - \varepsilon, s + \varepsilon] \subset U_s$. Ora, dato che s è il sup, esiste almeno un elemento di S in $(c - \varepsilon, c)$, cioè è possibile raggiungere un elemento di $(c - \varepsilon, c)$ con un numero finito di aperti. Allora aggiungendo U_s è possibile raggiungere i punti in $(c, c + \varepsilon)$. Da ciò segue che $c = b$, altrimenti non sarebbe l'estremo superiore.
Consideriamo adesso un insieme K chiuso e limitato generico. Essendo limitato, esiste un certo $[a, b] \supseteq K$. Ora, dato un ricoprimento aperto di K, lo si può estendere a un ricoprimento aperto di $[a, b]$ aggiungendo il complementare di K, aperto in quanto complementare di un chiuso. Ciò permette di ricondursi al caso precedente.

$(3 \implies 2)$ Consideriamo una successione di elementi di K. Per teorema di Bolzano-Weierstrass, una successione limitata ammette una sottosuccessione convergente a un certo limite reale L. Dato che l'insieme è chiuso, esso contiene tutti i suoi punti di accumulazione, tra cui L dato che una successione gli converge. Allora la successione ammette una sottosuccessione convergente in K, che è dunque sequenzialmente compatto.

$(2 \implies 3)$ Se K non è limitato, si può prendere una successione che tende a infinito, e hai voglia a estrarne una convergente a un valore reale. Se non è chiuso, si può prendere una successione convergente a un punto di accumulazione non contenuto in K, e si conclude parimenti.

□

7.9 Teorema di Weierstrass

Una funzione f continua da un insieme sequenzialmente compatto K a $\mathbb{R}$ ammette massimo e minimo.

Dimostrazione Mostriamo che la funzione ha massimo, l'altra tesi è simmetrica.
Sicuramente la funzione ammette un sup, dato che l'immagine è un sottoinsieme di $\mathbb{R}$. Per definizione di sup esiste una successione $y_n \in Im(f)$ che vi tende[9]. Per ciascun y_n, scegliamo una sua controimmagine x_n nel compatto. Dalla successione x_n si estrae una sotto-successione convergente a un certo $\bar{x}$, per ipotesi di compattezza. Ora, la funzione è continua, e quindi sequenzialmente continua. Perciò $f(\bar{x}) = \lim_{n \to +\infty} f(x_n) = \sup Im(f)$, dato che ogni sotto-successione della successione y_n convergente converge allo stesso limite. Allora il sup viene assunto, ed è dunque un massimo.

[9] La successione prende il nome di successione massimizzante.

□

Osservazione Vedremo in seguito che, per richiedere solo l'esistenza del massimo e non del minimo, è sufficiente la semicontinuità superiore, una proprietà più debole della continuità.

7.10 Teorema di Heine-Cantor (HeineKen)

Una funzione continua a valori in uno spazio metrico compatto è uniformemente continua.

Dimostrazione Per assurdo, sia continua ma non uniformemente continua. Cioè esiste un certo ε positivo tale che, comunque si scelga δ, ci sono due punti x, y a distanza minore di δ le cui immagini distano più di ε. In particolare, consideriamo le due successioni x_n e y_n costruite utilizzando $\delta_n = 1/n$. Per ipotesi di compattezza, è possibile estrarre da x_n una successione $x_{n(k)}$ convergente a un certo $z \in K$. Consideriamo la sotto-successione $y_{n(k)}$ con gli stessi indici, che per costruzione tende anch'essa a z essendo infinitesima la distanza da $x_{n(k)}$. Per continuità, $\lim_{h\to+\infty} f(x_{n(k_h)}) = \lim_{h\to+\infty} f(y_{n(k_h)}) = f(z)$. Dunque anche le successioni delle immagini convergono allo stesso limite, contro l'ipotesi che la loro distanza fosse sempre almeno ε.

□

8 $\mathbb{R}^n$

8.1 Volume della palla euclidea

> "Il volume della sfera qual è? Quattro terzi pigreco R tre. Il volume della palla euclidea qual è?..." Serve qualcosa che concluda la rima.
>
> *Filastrocca per bambini over 20.*

Vogliamo determinare il volume della palla euclidea n-dimensionale $B_n(0,r)$.
Osserviamo innanzitutto che, per ragioni di scala, $B_n(0,r) = r^n B_n(0,1)$, per cui basta determinare il volume della palla di raggio 1. Sia ω_n tale volume. Per calcolarlo consideriamo la funzione $f(x) = \sqrt{1-x^2}$, che rappresenta una semicirconferenza tra $x=-1$ e $x=1$. Ora, la palla euclidea può essere pensata come il solido ottenuto ruotando il grafico della circonferenza attorno all'asse X, in n dimensioni.[10] Infatti, se la distanza di un punto della palla dall'origine è al più 1, la distanza dall'asse X deve essere al più $\sqrt{1-x^2}$ per teorema di Pitagora. Allora, indicato con V_n il volume n-dimensionale, otteniamo

$$\omega_n = \left|\int_{-1}^{1} V_{n-1}(B_{n-1}(0,\sqrt{1-x^2}))\,dx\right| = \left|\int_{-1}^{1} V_{n-1}(B_{n-1}(0,1))(1-x^2)^{\frac{n-1}{2}}\,dx\right| =$$

$$= \omega_{n-1}\left|\int_{-1}^{1}(1-x^2)^{\frac{n-1}{2}}\,dx\right|$$

Ora, eseguiamo il cambio di variabile $x = \cos t$.

$$\omega_n = \omega_{n-1}\int_0^{\pi}\sin^{n-1}x\sin x\,dx = \omega_{n-1}\int_0^{\pi}\sin^n x\,dx = \omega_{n-1}I(n)$$

Si è già mostrato in 19.1 che l'integrale $I(n)$ vale:

$$\begin{cases} I(2n) = \frac{(2n-1)!!}{(2n)!!}I(0) = \frac{(2n-1)!!}{(2n)!!}\pi \\ I(2n+1) = \frac{(2n)!!}{(2n+1)!!}I(1) = \frac{(2n)!!}{(2n+1)!!}2 \end{cases}$$

Consideriamo prima il caso in cui n è pari. Noto che $\omega_1 = 2$ essendo un segmento di raggio 1, otteniamo

$$\omega_n = \omega_1\prod_{k=2}^{n}I(k) = \frac{2}{I(1)}\prod_{k=1}^{\frac{n}{2}}I(2k)I(2k-1)$$

Abbiamo fatto partire la produttoria da 1 per poter accoppiare i termini, e soprattutto perché, dato che $I(1) = 2$, ci si semplifica il coefficiente davanti.

$$\omega_n = \prod_{k=1}^{\frac{n}{2}}\frac{(2n-1)!!}{(2n)!!}\frac{(2n-2)!!}{(2n-1)!!}2\pi = \prod_{k=1}^{\frac{n}{2}}\frac{\pi}{n} = \frac{\pi^{n/2}}{(n/2)!}$$

Consideriamo ora il caso in cui n è dispari. Abbiamo che

$$\omega_n = I(n)\omega_{n-1} = 2\frac{(n-1)!!}{(n)!!}\frac{\pi^{\frac{n-1}{2}}}{\left(\frac{n-1}{2}\right)!}$$

$n-1$ è pari, quindi posso tirare fuori un $2^{(n-1)/2}$, ottenendo

$$\omega_n = 2\frac{2^{(n-1)/2}\left(\frac{n-1}{2}\right)!}{(n)!!}\frac{\pi^{\frac{n-1}{2}}}{\left(\frac{n-1}{2}\right)!} = \frac{2^{(n+1)/2}}{(n)!!}\pi^{\frac{n-1}{2}}$$

[10] Vedasi 19.2

Ora, osserviamo che, se n è dispari, vale la seguente formula per il suo fattoriale:

$$(n/2)! = \frac{n!!}{2^{(n-1)/2}}(1/2)! = \frac{\sqrt{\pi}n!!}{2^{(n+1)/2}}$$

Dove abbiamo usato che $(1/2)! = 1/2(-1/2)! = \sqrt{\pi}/2$. Allora possiamo sostituire e ottenere anche in questo caso

$$\omega_n = \frac{\pi^{n/2}}{(n/2)!}$$

□

8.2 Volume del simplesso con gli assi ortogonali

Consideriamo un piano n-dimensionale in $(\mathbb{R}^+)^n$, di equazione

$$x_1p_i + \ldots + x_np_n = m$$

Esso interseca l'asse x_i nel punto $(0, \ldots, m/p_i, 0, \ldots, 0)$. Allora il volume del simplesso che esso delimita vale

$$V(n) = \frac{m^n}{n! \prod_{i=1}^n p_i}$$

Dimostrazione Procediamo per induzione su n. Il caso $n = 2$ è l'area di un triangolo rettangolo, e quindi ok. Per il caso induttivo sviluppiamo per principio di cavalieri lungo l'asse x_n. Otteniamo allora

$$V(n) = \int_0^{p_n} V(n-1)(x_n)\, dx_n$$

Il volume del solido di dimensione inferiore è quello delimitato dal piano $n-1$-dimensionale di equazione

$$x_1p_i + \ldots + x_{n-1}p_{n-1} = m - x_np_n$$

Da cui otteniamo, per ipotesi induttiva:

$$V(n) = \int_0^{p_n} V(n-1)(x_n)\, dx_n = \int_0^{p_n} \frac{(m - x_np_n)^{n-1}}{(n-1)! \prod_{i=1}^{n-1} p_i}\, dx_n = \frac{m^n}{n! \prod_{i=1}^n p_i}$$

□

9 Cenni di teoria della misura su $\mathbb{R}$

9.1 Insieme trascurabile

Un sottoinsieme $S \subseteq \mathbb{R}$ si dice trascurabile, o di misura nulla, se, per ogni $\varepsilon > 0$ esiste una famiglia numerabile di intervalli $\{I_k\}_{k\in\mathbb{N}}$ che ricopre S e la cui sommatoria delle lunghezze è minore di ε. In simboli:

$$\forall \varepsilon \, \exists \, \{I_k\}_{k\in\mathbb{N}} \text{ intervalli tc. } \bigcup_{k\in\mathbb{N}} I_k, \; \sum_{k\in\mathbb{N}} |I_k| \leq \varepsilon$$

Non si richiede nulla sugli intervalli, che possono essere aperti, chiusi o totalmente a caso.

- Se un insieme è trascurabile, lo è ogni suo sottoinsieme.
- Unione numerabile di insiemi trascurabili è trascurabile. Infatti, fissato un certo ε, scegliamo per ogni insieme S_n un insieme di intervalli la cui somma delle lunghezze sia minore di $2^{-n}\varepsilon$. L'insieme di tutti gli intervalli è ancora numerabile, ricopre l'unione e ha somma delle lunghezze $\varepsilon \sum_{n\in\mathbb{N}} 2^{-n} = 2\varepsilon$.
- Un insieme numerabile è sempre trascurabile. Basta applicare l'osservazione agli insiemi $S_n = \{x_n\}$, sicuramente trascurabili.
- Sia S trascurabile e $f : \mathbb{R} \longrightarrow \mathbb{R}$ lipschitziana di costante k. Allora l'immagine $f(S)$ è trascurabile. Questo perché gli intervalli che ricoprono S vengono mandati in intervalli di lunghezza al massimo $k|I_n|$, dunque la somma è sempre infinitesima.
- Se S è compatto, si può estrarre un ricoprimento con un numero finito di intervalli. Volendo, è possibile prenderli disgiunti (basta unire quelli che si intersecano).

10 Spazi normati

Definizione Dato uno spazio vettoriale V, si definisce norma una funzione $\|\cdot\| : V \longrightarrow [0,+\infty[$ munita delle seguenti proprietà:

1. Non degenerazione: $\|v\| = 0 \iff v = 0$;
2. 1-Omogeneità: $\|\lambda v\| = |\lambda|\|v\|$ con $\lambda \in \mathbb{R}$;
3. Disuguaglianza triangolare: $\|v+w\| \leq \|v\| + \|w\|$.

La norma induce una distanza $d(v,w) = \|v-w\|$, ossia uno spazio normato è anche uno spazio metrico. In generale, non è vero il contrario: una distanza induce una norma $\|v\| = d(v,0)$ solo se lo spazio di partenza è uno spazio vettoriale (altrimenti non ha senso definire la proprietà triangolare) e se la distanza è invariante per traslazioni ($d(x+z,y+z) = d(x,y)$) e 1-omogenea ($d(\lambda x, \lambda y) = |\lambda| d(x,y)$).
Si può dimostrare che in uno spazio vettoriale su $\mathbb{R}$ finito dimensionale [11] tutte le norme possibili sono equivalenti, nel senso che la topologia indotta dalle relative distanze ha esattamente gli stessi aperti. In altri termini, una successione converge nella distanza indotta da una norma se e solo se converge in ogni altra distanza indotta. In generale, questo non è vero per spazi infinito dimensionali.

10.1 Disuguaglianza di Cauchy-Schwarz

Sia $\langle\cdot,\cdot\rangle$ un prodotto scalare definito positivo. Esso induce una norma data dalla radice della corrispondente forma quadratica:

$$\|v\| = \sqrt{\langle v,v\rangle}$$

Infatti, la funzione così definita è evidentemente non degenere e omogenea. Inoltre, vale la disuguaglianza triangolare:

$$\|v+w\| = \sqrt{\langle v+w, v+w\rangle} = \sqrt{\langle v,v\rangle + \langle w,w\rangle + 2\langle v,w\rangle} \leq \sqrt{\langle v,v\rangle + \langle w,w\rangle + 2\sqrt{\langle v,v\rangle\langle w,w\rangle}} = \|v\|+\|w\|$$

L'ultimo passaggio segue dalla disuguaglianza di Cauchy-Schwarz:

$$|\langle v,w\rangle| \leq \|v\|\|w\|$$

Dimostrazione Escludiamo il caso banale in cui almeno uno dei due vettori è nullo. Allora anche il loro prodotto scalare è non nullo, essendo definito positivo. Sia allora

$$z = w - v\frac{\langle v,w\rangle}{\langle v,v\rangle}$$

Facciamo il prodotto scalare di questo coso con il vettore v. Per linearità otteniamo

$$\langle z,v\rangle = \left\langle w - v\frac{\langle v,w\rangle}{\langle v,v\rangle}, v\right\rangle = \langle w,v\rangle - \langle w,v\rangle = 0$$

Dunque z è ortogonale sia a v che al suo multiplo $v\frac{\langle v,w\rangle}{\langle v,v\rangle}$. Allora vale che

$$\|w\| = \left\| z + v\frac{\langle v,w\rangle}{\langle v,v\rangle}\right\|$$

Per una generalizzazione del teorema di Pitagora, se due vettori sono ortogonali vale che

$$\|w\|^2 = \|z\|^2 + \frac{\langle v,w\rangle^2}{\langle v,v\rangle^2}\|v\|^2$$

[11] cioè in $\mathbb{R}^n$, a meno dell'isomorfismo di passaggio alle coordinate

Ricordando che la norma è la radice del prodotto scalare otteniamo

$$\|w\|^2 = \|z\|^2 + \frac{\langle v, w\rangle^2}{\|v\|^2} \geq \frac{\langle v, w\rangle^2}{\|v\|^2}$$

Da cui segue la tesi.

$$\|w\|^2\|v\|^2 \geq \langle v, w\rangle^2$$

□

10.2 Teorema del valore medio per curve a valori in $(V, \|\cdot\|)$

Derivabilità Una curva $\gamma : [a, b] \longrightarrow (V_{/\mathbb{R}}, \|\cdot\|)$ a valori in uno spazio normato reale si dice derivabile in $x = x_0$ se in un intorno di x_0 ammette uno "sviluppo polinomiale" di ordine 1, ossia esiste un vettore $\gamma'(x_0) \in V$, detto vettore tangente, tale che

$$\gamma(x) = \gamma(x_0) + (x - x_0)\gamma'(x_0) + o(x - x_0)$$

Teorema del valore medio Sia $\gamma :\ [a, b] \longrightarrow (V, \|\cdot\|)$ una curva derivabile in (a, b). Allora vale che

$$\|\gamma(b) - \gamma(a)\| \leq \sup_{a<t<b} \|\gamma'(t)\|(b - a)$$

Si osservi che questa riformulazione del teorema di Lagrange è valida anche nel caso reale, ma è molto più debole perché fornisce solo una disuguaglianza, la quale spesso è anche stretta.[12]

Dimostrazione Escludiamo il caso ovvio in cui il sup è infinito. Sia allora $\sup_{a<t<b} \|\gamma'(t)\| < M$. Mostriamo che vale la disuguaglianza per M, la cui generalità permetterà poi di affermare che la disuguaglianza debole vale anche per $\sup_{a<t<b} \|\gamma'(t)\|$. Definiamo la funzione:

$$g(x) = \|\gamma(x) - \gamma(a)\| - Mx$$

La funzione è continua perché somma di funzioni continue, e pertanto deve ammettere un minimo essendo definita su un intervallo chiuso e limitato. Scegliamo ora un punto $a \leq x_0 < b$ diverso da b, in modo che esista un $x_0 < x < b$. Mostriamo che la differenza $g(x) - g(x_0)$ è negativa, e dunque x_0 non può essere il minimo.

$$g(x) - g(x_0) = \|\gamma(x) - \gamma(a)\| - \|\gamma(x_0) - \gamma(a)\| - M(x - x_0)$$

Per disuguaglianza triangolare vale che

$$g(x) - g(x_0) \leq \|\gamma(x) - \gamma(x_0)\| - M(x - x_0)$$

Essendo γ derivabile in x_0 abbiamo che

$$g(x) - g(x_0) \leq \|\gamma'(x_0)(x - x_0) + o(x - x_0)\| - M(x - x_0) = (x - x_0)(\|\gamma'(x_0) + o(1)\| - M) \leq 0$$

L'ultima disuguaglianza è data dal fatto che abbiamo supposto $\sup_{a<t<b} \|\gamma'(t)\| < M$. Ora, dato che nessun $a \leq x_0 < b$ può essere il minimo, questo deve essere b. In particolare

$$g(b) \leq g(a) \quad \Longrightarrow \quad \|\gamma(b) - \gamma(a)\| \leq M(b - a)$$

Da cui la tesi.

□

[12] Ad esempio, la curva $\gamma(t) : \mathbb{R} \longrightarrow \mathbb{C}$ definita come $\gamma(t) = e^{it}$ assume lo stesso valore in $t = 0$ e $t = 2\pi$, ma $\sup_{0<t<2\pi} \left\| \frac{d}{dt} e^{it} \right\| = \sup_{0<t<2\pi} \|ie^{it}\| = 1$.

10.3 Variante integrale del TVM

Si dimostra che vale la seguente disuguaglianza:

$$\|\gamma(b) - \gamma(a)\| \leq {}^{*}\!\!\int_a^b \|\gamma'(t)\| \, dt$$

Quanto si sta affermando è leggermente più preciso del teorema precedente, in quanto, per definizione di integrale superiore,

$${}^{*}\!\!\int_a^b \|\gamma'(t)\| \, dt \leq \sup_{a<t<b} \|\gamma'(t)\|(b-a)$$

Dimostrazione Siano $q \subset p$ due partizioni di $[a, b]$. Allora

$$\|\gamma(b) - \gamma(a)\| \leq L(\gamma, q) \leq L(\gamma, p) = \sum_{k=1}^{n} \|\gamma(x_k) - \gamma(x_{k-1})\|$$

Ora, applichiamo il teorema del valor medio a ciascun intervallo:

$$\|\gamma(b) - \gamma(a)\| \leq L(\gamma, q) \leq \sum_{k=1}^{n} |I_k| \sup_{t \in (x_{k-1}, x_k)} \|\gamma'(t)\| = S(\|\gamma'\|, p)$$

Passando all'inf su p si ottiene la tesi:

$$\|\gamma(b) - \gamma(a)\| \leq {}^{*}\!\!\int_a^b \|\gamma'(t)\| \, dt$$

□

10.4 Calcolo della lunghezza tramite un integrale

Se al termine della precedente dimostrazione passiamo al sup su q otteniamo anche la seguente disuguaglianza:

$$L(\gamma, [a, b]) \leq {}^{*}\!\!\int_a^b \|\gamma'(t)\| \, dt$$

Si può mostrare che, se la curva è derivabile e la derivata è integrabile[13], allora la lunghezza della curva è l'integrale della sua derivata, ossia:

$$L(\gamma, [a, b]) = \int_a^b \|\gamma'(t)\| \, dt$$

Dimostrazione Mostriamo che la differenza in modulo tra i due termini dell'uguaglianza è infinitesima. Data una partizione p, per definizione dei due oggetti abbiamo

$$\left| \int_a^b \|\gamma'(t)\| \, dt - L(\gamma, [a, b]) \right| = \left| \inf_{p \in P} S(\|\gamma'\|, p) - \sup_{p \in P} L(\gamma, p) \right| \leq |S(\|\gamma'\|, p) - L(\gamma, p)| =$$

$$= \left| \sum_{k=1}^{n} |I_k| \sup_{t \in I_k} \|\gamma'(t)\| - \|\gamma(x_k) - \gamma(x_{k-1})\| \right| \leq \sum_{k=1}^{n} \sup_{t \in I_k} \big| |I_k| \|\gamma'(t)\| - \|\gamma(x_k) - \gamma(x_{k-1})\| \big|$$

Si è portato fuori il sup dato che l'unica quantità dipendente da t è $\|\gamma'(t)\|$. Ora, per disuguaglianza triangolare della norma, che vale anche in modulo per simmetria[14], abbiamo che

$$\left| \int_a^b \|\gamma'(t)\| \, dt - L(\gamma, [a, b]) \right| \leq \sum_{k=1}^{n} \sup_{t \in I_k} \big| |I_k| \|\gamma'(t)\| - \|\gamma(x_k) - \gamma(x_{k-1})\| \big| \leq$$

[13] Le ipotesi del teorema valgono in particolare per funzioni di classe C^1, derivabili e con derivata continua, che saranno quelle che studieremo principalmente.

[14] Dato che $\|u\| - \|v\| \leq \|u - v\|$ e viceversa, vale che $| \, \|u\| - \|v\| \, | \leq \|u - v\|$

$$\leq \sum_{k=1}^{n} \sup_{t\in I_k} \| |I_k|\gamma'(t) - (\gamma(x_k) - \gamma(x_{k-1}))\|$$

Ora, $|I_k| = x_k - x_{k-1}$. L'espressione dentro la norma è allora una funzione calcolata tra gli estremi x_{k-1} e x_k.

$$\left| \int_a^b \|\gamma'(t)\|\, dt - L(\gamma, [a,b]) \right| \leq \sum_{k=1}^{n} \sup_{t\in I_k} \|x\gamma'(t) - \gamma(x)\|_{x=x_{k-1}}^{x=x_k}$$

Applicando il teorema del valore intermedio, derivando rispetto a x, otteniamo

$$\left| \int_a^b \|\gamma'(t)\|\, dt - L(\gamma, [a,b]) \right| \leq \sum_{k=1}^{n} \sup_{t\in I_k} \sup_{x\in I_k} \|\gamma'(t) - \gamma'(x)\|\, |I_k| = \sum_{k=1}^{n} \rho(\gamma', p)|I_k|$$

Ora, se la derivata è integrabile, la sua variazione è infinitesima. Segue allora l'uguaglianza

$$L(\gamma, [a,b]) = \int_a^b \|\gamma'(t)\|\, dt$$

□

10.5 Polinomio di Taylor per curve

Sia $\gamma : [a,b] \longrightarrow (V, \|\cdot\|)$ una curva derivabile n volte in un punto x_0. Allora esiste ed è unico il suo sviluppo polinomiale di ordine n, che si identifica con il suo polinomio di Taylor:

$$T_n(\gamma, x, x_0) = \sum_{k=0}^{n} \gamma^{(k)} \frac{(x-x_0)^k}{k!}$$

Con leggeri accorgimenti, valgono tutte le proprietà già viste per il caso $V = \mathbb{R}$. In particolare, valgono le seguenti formule del resto, da considerare in norma.

Resto di Peano $\|\gamma(x) - T_n(\gamma, x, x_0)\| = o(x-x_0)^n$. La formula si ricava reiterando la costruzione già vista, che si basa sul fatto che un polinomio è un infinitesimo di ordine n se ha tutte le derivate nulle fino alla n-esima.

Resto di Lagrange Osservando che non si può trasportare al caso generale l'uguaglianza vista per $\mathbb{R}$, che si fondava sul teorema di Lagrange nel caso reale, si ha la seguente formula:

$$\|\gamma(x) - T_n(\gamma, x, x_0)\| \leq \sup t \in a, b \left\|\gamma^{(n+1)}(t)\right\| \frac{|x-t|^{n+1}}{(n+1)!}$$

Per dimostrarlo, ricordiamo innanzitutto che la derivata del polinomio di Taylor rispetto al centro dello sviluppo vale

$$\|\frac{d}{dt} T_n(\gamma, x, t)\| = \left\|\gamma^{(n+1)}(t)\right\| \frac{|x-t|^n}{n!}$$

Dunque, applicando il teorema del valor medio nella sua formulazione integrale, derivando rispetto a t, otteniamo:

$$\|\gamma(x) - T_n(\gamma, x, x_0)\| \leq^* \int_{\min(x,x_0)}^{\max(x,x_0)} \left\|\gamma^{(n+1)}(t)\right\| \frac{|x-t|^n}{(n)!}\, dt$$

Maggiorando la norma della derivata con il suo sup possiamo portarla fuori dall'integrale e ottenere quanto voluto:

$$\|\gamma(x) - T_n(\gamma, x, x_0)\| \leq \sup_{t\in(a,b)} \left\|\gamma^{(n+1)}(t)\right\| \int_{\min(x,x_0)}^{\max(x,x_0)} \frac{|x-t|^n}{(n)!}\, dt = \sup_{t\in(a,b)} \left\|\gamma^{(n+1)}(t)\right\| \frac{|x-x_0|^{n+1}}{(n+1)!}$$

□

Resto integrale Sì, dai, il lettore la fa facile. In realtà servono un paio di strutture preliminari che non sono state ancora definite rigorosamente, tra cui:

- Una teoria dell'integrazione secondo Riemann di curve a valori in spazi normati. In dimensione finita e in uno spazio vettoriale su $\mathbb{R}$ si può dire che una curva è integrabile se e solo se lo sono le sue componenti. Più in generale, si può adattare la definizione di integrabilità originale data da Riemann: una curva è integrabile se e solo se partizioni abbastanza piccole rendono infinitesima l'oscillazione della funzione. In altri termini

$$\forall \varepsilon\, \exists \delta \text{ t.c. } \forall p \in P,\ |p| < \delta \implies \rho(\gamma, p) = \sum_{k=1}^{n} |I_k| \sup_{a,b \in I_k} \|\gamma(b) - \gamma(a) < \varepsilon$$

- Lo spazio vettoriale normato deve essere completo rispetto alla distanza indotta dalla norma[15]. Si può dimostrare che questa condizione è equivalente al cosiddetto criterio di convergenza totale, secondo il quale se una sommatoria di vettori converge in norma allora converge anche come vettore. In simboli, lo spazio è completo, se e solo se

$$\sum_{k=0}^{+\infty} \|v_k\| < +\infty \quad \implies \quad w = \sum_{k=0}^{+\infty} v_k \in V$$

  Con una leggera improprietà linguistica, si può dire che il criterio di convergenza totale serve a assicurare l'esistenza di un "integrale superiore" e di uno inferiore, che in $\mathbb{R}$ era garantita dalla completezza.

- Serve una formulazione del teorema fondamentale del calcolo integrale, che si può riformulare come segue:
  Data una funzione γ derivabile e con derivata integrabile, vale che

$$\int_a^b \gamma'(x)\, dx = \gamma(b) - \gamma(a)$$

  La dimostrazione dovrebbe essere pressoché analoga a quella già vista.

10.6 Parametrizzazione in lunghezza d'arco

Curva regolare Una curva $\gamma : [a, b] \longrightarrow (V, \|\cdot\|)$ si dice regolare se è di classe C^1 e la norma della sua derivata non è mai nulla.

Parametrizzazione in lunghezza d'arco Una curva regolare può essere riparametrizzata in modo che $\|(\gamma \circ \sigma)'(t)\| = 1$, ossia in modo da essere "percorsa" a velocità costante.

Dimostrazione Sia $\sigma(x) = L(\gamma, [a, x])$. Per il teorema visto in 10.4 la lunghezza si può esprimere come integrale della derivata, dato che la curva è di classe C^1:

$$\sigma(x) = L(\gamma, [a, x]) = \int_a^x \|\gamma'(t)\|\, dt$$

Ora, σ è crescente in senso stretto, e dunque bigettiva, e di classe C^1, in quanto funzione integrale di una funzione continua. Allora anche la derivata è continua, e dunque è un omeomorfismo. Possiamo allora scrivere

$$\gamma = (\gamma \circ \sigma^{-1}) \circ \sigma = u \circ \sigma$$

Ora, derivando entrambi i membri per mezzo della "chain rule" otteniamo

$$\gamma'(t) = \frac{du}{d\sigma(t)} \sigma'(t) = \frac{du}{d\sigma(t)} \|\gamma'(t)\|$$

[15] Si dice che lo spazio è di Banach.

Valutando in norma le due scritture, è possibile semplificare $\|\gamma'(t)\|$, che per ipotesi di regolarità non è mai nullo. Allora otteniamo

$$\left\| \frac{du}{d\sigma(t)} \right\| = 1$$

Da cui la tesi.

$\square$

11 Funzioni e continuità

11.1 Definizione topologica di continuità

Intorno Un soprainsieme di un aperto contenente un punto x si dice intorno di x. In altri termini, è un intorno di x qualsiasi insieme che contenga un aperto contenente x.

Continuità Una funzione $f : X \longrightarrow Y$ tra spazi topologici si dice continua in un punto x_0 se ogni intorno di $f(x_0)$ contiene l'immagine di un intorno di x_0. In simboli

$$\forall \text{ V intorno di } f(x_0) \quad \exists \text{ U intorno di } x_0 \text{ t.c. } f(U) \subseteq V$$

(*) É equivalente dire che una funzione è continua $\iff$ la controimmagine di un aperto è un aperto. La dimostrazione è lasciata al lettore per esercizio.

Punto di accumulazione Un punto $x \in (X, \tau)$ si dice di accumulazione (per X) se in ogni intorno di x è presente almeno un altro punto dell'insieme. In simboli

$$\forall \text{ U intorno di } x \quad U \setminus \{x\} \cap X \neq \emptyset$$

Limite Data una funzione $f : X \longrightarrow Y$ tra spazi topologici, un punto $x_0 \in X$ di accumulazione per X e un punto $y_0 \in Y$, si dice che il limite di $f(x)$ per $x \to x_0$ tende a y_0 se la funzione

$$g(x) = \begin{cases} f(x) & \text{se } x \neq x_0 \\ y_0 & \text{se } x = x_0 \end{cases}$$

è continua in x_0.
Era altresì possibile definire prima il limite e poi la continuità in un punto, con le definizioni usuali.

11.2 Connessione

Un sottoinsieme di uno spazio topologico $E \subseteq (X, \tau)$ si dice connesso se non esistono due aperti $A_1, A_2 \in \tau$ disgiunti e non vuoti che ricoprono E, vale a dire $E \subseteq A_1 \cup A_2$.
Le seguenti proprietà sono equivalenti:

1. X è connesso;
2. Ogni funzione continua $F : X \longrightarrow \{0, 1\}$, dove l'insieme di arrivo ha la topologia discreta $\{\emptyset; \{0\}; \{1\}; \{0, 1\}; \}$, è costante;
3. Gli unici insiemi contemporaneamente aperti e chiusi sono $\emptyset$ ed X.

Dimostrazione L'equivalenza ($1 \iff 3$) è ovvia: se non fosse connesso, ci sarebbe un aperto di X il cui complementare è aperto, e viceversa. Per quanto riguarda ($1 \iff 2$), osserviamo che nella topologia discreta i singoletti $\{0\}$ e $\{1\}$ sono aperti, dunque una funzione continua è tale che le controimmagini dei singoletti sono aperti disgiunti in X. Se per assurdo X non fosse connesso, basterebbe definire una funzione che valesse 0 su uno dei due aperti e 1 sull'altro, e questa sarebbe continua ma non costante. Viceversa, una funzione continua e non costante avrebbe come controimmagini due aperti disgiunti che ricoprirebbero X.

□

Corollario In $\mathbb{R}$, gli insiemi connessi sono gli intervalli.
Infatti, se un insieme connesso E non fosse un intervallo, esisterebbero tre punti $x < y < z$, con $x, z \in E$ e $y \notin E$. Ma allora E potrebbe essere ricoperto con gli aperti $(-\infty, y)$ e $(y, +\infty)$, evidentemente disgiunti e non vuoti.
Viceversa, per assurdo un intervallo non sia connesso. WLOG, supponiamo che l'intervallo sia della forma $I = (a, b)$, dato che il caso di intervalli illimitati è analogo. Siano A_1 e A_2 gli aperti che

ricoprono I. WLOG, sia $x_0 \in A_1 \cap I$ e supponiamo che esistano degli elementi di $A_2 \cap I$ maggiori di x_0. Sia $x_1 = \sup\{x : [x_0, x] \subseteq A_1 \cap I\}$, che esiste finito e non coincide con b. Dunque $x_1 \in I$. Tuttavia, per massimalità non può appartenere a $A_1 \cap I$, altrimenti ci sarebbe una palla di centro x_1 contenuta in $A_1 \cap I$. D'altronde non può nemmeno appartenere a $A_2 \cap I$, o non sarebbe aderente a $A_1 \cap I$.

□

11.3 Continuità per successioni

Negli spazi metrici si fa tutto per successioni

Vecchio proverbio cinese

Definizione Una funzione $f : X \longrightarrow Y$ tra spazi topologici si dice sequenzialmente continua, o continua per successioni, in un punto $\bar{x} \in X$ se, comunque si scelga una successione x_n convergente a $\bar{x}$, le immagini dei termini della successione convergono a $f(\bar{x})$. In simboli

$$\forall (x_n)_{n\in\mathbb{N}} \subseteq X \text{ t.c. } \lim_{n\to+\infty} x_n = \bar{x} \quad \text{vale che} \quad \lim_{n\to+\infty} f(x_n) = f(\bar{x})$$

Una funzione continua è sequenzialmente continua Basta applicare la definizione di continuità.

In $\mathbb{R}$, una funzione sequenzialmente continua è continua Per assurdo la funzione non sia continua in $\bar{x}$. Allora esiste un intorno di $f(\bar{x})$ che non contiene completamente l'immagine di nessun intorno di $\bar{x}$. In simboli

$$\exists \text{ V intorno di } f(\bar{x}) \text{ t.c. } \forall \text{ U intorno di } \bar{x} \text{ vale che } f(U) \nsubseteq V$$

Consideriamo gli intorni U_n di raggio $1/n$ e sia y_n un punto di $f(U)\setminus V$. Sia x_n una controimmagine di y_n. Poiché gli x_n tendono a $\bar{x}$ per costruzione e la funzione è sequenzialmente continua, si deve avere che

$$\lim_{n\to+\infty} y_n = \lim_{n\to+\infty} f(x_n) = f(\bar{x})$$

Ma ciò è impossibile, perché allora la successione y_n dovrebbe essere contenuta definitivamente in ogni intorno di $f(\bar{x})$, e quindi anche in V.

□

Se si rilegge la dimostrazione, si osserva che è possibile indebolire l'ipotesi di continuità sequenziale: è sufficiente che ogni successione $x_n \to \bar{x}$ ammetta una sotto-successione $x_{n(k)}$ le cui immagini convergano a $f(\bar{x})$. In simboli, f è continua in $\bar{x}$ sse:

$$\forall (x_n)_{n\in\mathbb{N}} \subseteq X \text{ t.c. } \lim_{n\to+\infty} x_n = \bar{x} \quad \exists (x_{n(k)})_{k\in\mathbb{N}} \text{ t.c. } \lim_{k\to+\infty} f(x_{n(k)}) = f(\bar{x})$$

(*) Il teorema non è vero in ogni spazio topologico. In generale, le proprietà che riguardano le successioni, tra cui l'equivalenza tra continuità sequenziale e continuità topologica, valgono in spazi topologici "primo-numerabili", nei quali ciascun punto ammette un sistema fondamentale di intorni numerabile. Un sistema fondamentale di intorni è un insieme I di intorni cofinale per inclusione insiemistica, ossia tale che per ogi intorno U esiste un intorno $V \in I$ tale che $V \subseteq U$. Per fortuna, tutti gli spazi metrici sono primo-numerabili, e dunque negli spazi metrici è possibile caratterizzare tutto con le successioni.

11.4 Continuità uniforme

Una funzione è continua se, per ogni bullone, esiste una chiave inglese che lo avvita. Una funzione è uniformemente continua se esiste una chiave universale per tutti i tipi di bulloni.

Insegna davanti al "Paradiso della brugola"

Una funzione tra spazi metrici $f : (X, d_x) \longrightarrow (Y, d_y)$ si dice uniformemente continua se

$$\forall \varepsilon \in \mathbb{R}^+ \, \exists \delta_\varepsilon \in \mathbb{R}^+ \text{ t.c. } \forall a, b \in X \; d_x(a, b) < \delta_\varepsilon \implies d_y(f(a), f(b)) < \varepsilon$$

Se confrontiamo la definizione con quella della continuità usuale, ossia

$$\forall \varepsilon \in \mathbb{R}^+ \, \forall a \in X \, \exists \delta_{\varepsilon,\, a} \in \mathbb{R}^+ \text{ t.c. } \forall b \in X \; d_x(a, b) < \delta_{\varepsilon,\, a} \implies d_y(f(a), f(b)) < \varepsilon$$

Possiamo vedere che la differenza sta solo nell'inversione tra $\forall a \in X$ e $\exists \delta_\varepsilon \in \mathbb{R}^+$. In pratica, se per quanto riguarda le funzioni continue il δ dipende dal punto a scelto, per le funzioni uniformemente continue esso è una proprietà della curva, dipendente solo da ε.
É facile osservare che le funzioni lipschitziane sono uniformemente continue, e che le funzioni uniformemente continue sono continue. Non è sempre vero il viceversa: esistono funzioni continue ma non uniformemente (per esempio $f(x) = x^2$) e funzioni uniformemente continue ma non lipschitziane (per esempio, $g(x) = \sqrt{x}$ in $[0, 1]$).

Modulo di continuità Una funzione f è uniformemente continua se e solo se esiste una funzione $\omega : [0, +\infty] \longrightarrow [0, +\infty]$, detta modulo di continuità [16], con le seguenti proprietà:

1. ω è monotona crescente in senso debole;
2. $\omega(0) = 0$ ed è continua in $x = 0$
3. $d(x, y) < t \implies d(f(x), f(y)) \leq \omega(t)$.

Dimostrazione ($\Longleftarrow$) É evidente che la funzione è uniformememente continua, perché basta rileggere la proprietà 3 assumendo $t = \delta$ tale che $\omega(\delta) \leq \varepsilon$. L'esistenza di un δ fatto così è garantita dalla continuità in zero, in quanto $\forall \varepsilon \, \exists \delta$ tale che $\omega(\delta) \leq \varepsilon$.
($\Longrightarrow$) Mostriamo che è possibile costruire un modulo di continuità data una funzione uniformemente continua. Definiamo

$$\omega(t) = \sup\{d(f(x), f(y)) \text{ t.c. } d(x, y) \leq t\}$$

É ovvio che la funzione costruita è crescente, in quanto si fa il sup di un insieme sempre più grande, e che assume sempre valori positivi o nulli perché nell'insieme di cui si fa il sup c'è sempre $0 = d(f(x), f(x))$. Anche la proprietà 3 è evidente per costruzione. Rimane da mostrare che è continua in $x = 0$, ma questo segue dal fatto che, essendo la funzione uniformemente continua, $\forall \varepsilon \, \exists \delta$ tale che $d(a, b) < \delta \implies d(f(a), f(b)) < \varepsilon$ Allora passando al sup si ottiene che $\omega(\delta) \leq \varepsilon$.

□

Si osserva che le funzioni Lipschitziane sono uniformemente continue anche perché hanno un modulo di continuità, del tipo $\omega(t) = Lt$ dove L è la costante di Lipschitz. Similmente si può definire la classe delle funzioni H́olderiane, uniformemente continue di modulo $\omega(t) = Lt^\alpha$, con $\alpha \in (0, 1]$. Dalla definizione segue subito che le funzioni Lipschitziane sono H́olderiane di costante $\alpha = 1$.

[16] Si rimanda alla pagina https://en.wikipedia.org/wiki/Modulus_of_continuity. L'autore è un certo "Pmajer". Sembra che i riferimenti bibliografici siano fittizi.

11.5 Discontinuità di una funzione monotona

Una funzione monotona ammette limite da destra $f^+(x)$ e da sinistra $f^-(x)$ in ogni punto x interno al dominio. Ciò segue facilmente dal fatto che una successione monotona ammette limite. Inoltre, supponendo WLOG che la funzione sia crescente, il limite da sinistra è minore o uguale del limite da destra, perché le disuguaglianze passano al limite. Da queste osservazioni seguono due importanti corollari:

Corollario Una funzione monotona non ammette punti di discontinuità di seconda specie nei punti interni al dominio. Questo perché i limiti da destra e da sinistra esistono e devono essere finiti, dato che devono essere l'uno maggiore dell'altro. Inoltre, non ammette nemmeno punti di discontinuità di terza specie, o non sarebbe rispettata la monotonia.

Corollario del corollario Una funzione monotona ammette solo discontinuità di prima specie. L'insieme dei punti di discontinuità è al più numerabile.

Dimostrazione Wlog, sia $f(x)$ crescente. Un punto di discontinuità è caratterizzato dal fatto che $f^-(x) < f^+(x)$. Per ogni punti di discontinuità si sceglie un razionale p_x tale che $f^-(x) < p_x < f^+(x)$. Si osserva che ogni razionale è associato univocamente al suo punto di discontinuità, perché la funzione è crescente. Infatti, presi comunque $x < y$ due punti di discontinuità, si ha che $f^-(x) < p_x < f^+(x) \leq f^-(y) < p_y < f^+(y)$. Dunque, l'insieme dei punti di discontinuità della funzione è in bigezione con un sottoinsieme di $\mathbb{Q}$, che deve essere al più numerabile.

$\square$

(*) In generale, è possibile mostrare che per qualsiasi funzione i punti di discontinuità di prima e terza specie sono al più numerabili, procedendo in modo simile ma associando a un punto di discontinuità una terna di razionali.

11.6 Funzioni convesse

Combinazione convessa Dato uno spazio vettoriale V, una combinazione lineare di due vettori $\mathbf{x}$ e $\mathbf{y}$ si dice combinazione convessa se la somma dei coefficienti vale 1. In simboli

$$t\mathbf{x} + (1-t)\mathbf{y}, \text{ con } t \in [0,1]$$

Si osservi che è equivalente scrivere $\mathbf{y} + t(\mathbf{x} - \mathbf{y})$. Da questa formula, si vede che al variare di t le combinazioni convesse descrivono graficamente il segmento tra i due vettori. In particolare, se $\mathbf{x}$ e $\mathbf{y}$ sono numeri reali, si ha che una loro combinazione convessa appartiene all'intervallo $[x, y]$.

Funzione convessa Una funzione $f : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ si dice convessa se, dati due vettori $\mathbf{x}$ e $\mathbf{y}$, l'immagine di una loro combinazione convessa è minore o uguale della combinazione convessa delle immagini. In simboli

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}), \text{ con } t \in [0,1]$$

Graficamente, se lo spazio di partenza è $\mathbb{R}$, la funzione sta sotto la corda tra i punti $(x, f(x))$ e $(y, f(y))$. Infatti, si verifica facilmente che il termine alla RHS è il valore assunto dalla retta $y - \mathbf{y} = \frac{f(\mathbf{y})-f(\mathbf{x})}{y-x}(x - \mathbf{x})$ nel punto $x = \mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y}$.

11.7 Proprietà delle funzioni convesse di variabile reale

Dominio di finitezza Si dice dominio di finitezza di una funzione convessa l'insieme dei punti in cui essa assume valore finito. Si osserva facilmente che, se lo spazio di partenza è $\mathbb{R}$, il dominio di finitezza è un intervallo, ossia se due punti x, y hanno immagine finita anche ogni loro combinazione convessa ha immagine finita, per la convessità della funzione.

Teorema Una funzione reale di variabile reale è (strettamente) convessa $\iff$ i rapporti incrementali nel suo dominio di finitezza sono (strettamente) crescenti. Ossia

$$x < y < z \quad \implies \quad \frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y}$$

Si osservi che si sta chiedendo che il rapporto incrementale, come funzione in due variabili $R(x, y)$, sia crescente in una delle due se si fissa l'altra.

Dimostrazione Dimostriamo una sola delle due disuguaglianze, l'altra è analoga. Inoltre, dimostriamo solo l'implicazione ($\implies$), dato che l'altra segue ripercorrendo al contrario i passaggi della dimostrazione.
Posto $y = tx + (1 - t)z$, per definizione di funzione convessa si ha che

$$f(y) \leq tf(x) + (1 - t)f(z)$$

Sottraiamo da entrambe le parti $f(x)$ e dividiamo per $y - x = (t - 1)x + (1 - t)z$, che è positivo e non cambia il segno della disuguaglianza.

$$\frac{f(y) - f(x)}{y - x} \leq\leq \frac{(t - 1)f(x) + (1 - t)f(z)}{(t - 1)x + (1 - t)z} = \frac{f(z) - f(x)}{z - x}$$

□

Una funzione convessa è localmente lipschitziana. Infatti, dati due punti $x < y$ interni all'intervallo di finitezza (a, b), esistono $y < \beta < \gamma < b$ tali che

$$f(y) - f(x) \leq (y - x)\frac{f(\gamma) - f(\beta)}{\gamma - \beta}$$

Ne segue che una funzione convessa è localmente uniformemente continua, eccetto al più agli estremi del dominio di finitezza. Ad esempio, la funzione

$$f(x) = \begin{cases} 1 \text{ se } |x| = 1 \\ 0 \text{ se } |x| < 1 \end{cases}$$

è convessa ma discontinua agli estremi.

11.8 Derivabilità di una funzione convessa

Esistenza di derivata destra e sinistra Il fatto che i rapporti incrementali siano debolmente crescenti fa sì che in ogni punto del dominio di finitezza, eccetto al più agli estremi, siano definite la derivata destra e sinistra della funzione, e che queste siano monotone crescenti (debolmente). Tuttavia non è detto che coincidano: la funzione $f(x) = |x|$ è convessa ma non derivabile in $x = 0$.

Derivabilità Una funzione convessa è derivabile se e solo se la derivata destra è continua.

Dimostrazione

- ($\implies$) Se la funzione è derivabile, la sua derivata coincide con la derivata destra. Inoltre, dato che la derivata destra di una funzione convessa è monotona, essa ammette solo discontinuità di prima specie. Tuttavia, se ci fossero dei salti la funzione non sarebbe derivabile in quei punti, e quindi la derivata destra è continua.
- ($\impliedby$) Per assurdo, la funzione non sia derivabile. Allora esiste un punto $x_0 \in (a, b)$ in cui la derivata destra e sinistra esistono reali e distinte. Sia $f'_+(x_0) - f'_-(x_0) = l$. Per continuità di f'_+, esiste un intorno sinistro di x_0 in cui

$$f'_-(x_0) < f'_+(x) < f'_+(x_0)$$

Tuttavia, essendo i rapporti incrementali crescenti, abbiamo che per $x < x_0$, $f'_+(x) \leq f'_-(x_0)$. Da ciò segue l'assurdo.

□

12 Successioni

12.1 Convergenza di sotto-successioni

Fatto Una successione x_n a valori in uno spazio topologico converge a $\bar{x} \iff$ sia la successione dei termini di indice pari x_{2n} sia quella dei dispari x_{2n+1} convergono a $\bar{x}$.
La dimostrazione di questo fatto è immediata. É invece interessante provare a pensare sotto quali condizioni la convergenza a un certo $\bar{x}$ di un numero finito k di sotto-successioni del tipo x_{nm_j}, ossia dei multipli di certi m_j, $j = 1, \dots, k$, implica la convergenza a $\bar{x}$ di tutta la successione.

Proprietà di Uryshon Una successione x_n converge a $\bar{x} \iff$ ogni sotto-successione $x_{n(k)}$ ammette una sotto-sotto-successione $x_{n(k(h))}$ convergente a $\bar{x}$.

Dimostrazione La freccia $\Longrightarrow$ è ovvia, dato che ogni sotto-successione converge allo stesso limite della successione madre. Viceversa, per assurdo la successione non converga a $\bar{x}$. Ciò significa che esiste un intorno U di $\bar{x}$ in cui la successione non sta frequentemente. Quindi esiste una sotto-successione $x_{n(k)}$ di termini che non stanno in U. Da questa sotto-successione non si può estrarre alcuna sotto-sotto-successione convergente a $\bar{x}$, dato che non può mai stare in un certo intorno di $\bar{x}$.

$\square$

Fatto Se una successione x_n in uno spazio metrico converge a un certo $\bar{x}$, allora è possibile estrarre una sotto-successione che converge a $\bar{x}$ con velocità arbitrariamente grande. In altri termini, data una successione ε_n decrescente e infinitesima qualsiasi, è possibile estrarre una sotto-successione $x_{n(k)}$ tale che $d(x_{n(k)}, \bar{x}) \leq \varepsilon_k \ \forall\, k \in \mathbb{N}$.
La dimostrazione è ovvia e si basa sul fatto che la distanza dal limite è infinitesima, cioè piccola a piacere.

12.2 Ogni successione in $\mathbb{R}$ ammette una sotto-successione monotona (debolmente)

Più in generale, la tesi vale per qualsiasi spazio $(X, \leq)$ totalmente ordinato.

Dimostrazione Premettiamo la seguente definizione, che ci tornerà utile nella dimostrazione. Si dice coda n-esima di una successione la sotto-successione dei termini dall'n-esimo in poi.
Distinguiamo ora due casi.

1. Supponiamo che esiste una coda, diciamo la coda n-esima, priva di massimo. Allora ogni coda successiva sarà ancora priva di massimo, perché si toglieranno al più un numero finito di termini. Si definisce la seguente sotto-successione di indici $n(k)$:

$$\begin{cases} n(0) = n; \\ n(k+1) = \min\{m > n(k) \text{ t.c. } x_m > x_{n(k)}\} \end{cases}$$

   La definizione è ben posta, perché dato che non esiste un massimo è sempre possibile trovare un indice successivo il cui termine è maggiore dei precedenti. Si osserva inoltre che per costruzione gli $n(k)$ sono strettamente crescenti, e dunque la sotto-successione $x_{n(k)}$ è ben definita. Infine, per costruzione essa è monotona crescente in senso stretto.

2. Supponiamo invece che ogni coda ammetta un massimo. Allora consideriamo la sotto-successione degli $x_{n(k)}$ che sono uguali al massimo della loro coda. Dato che ogni coda ha massimo, la sotto-successione è infinita, e dunque ben definita. É evidente inoltre che il massimo di una coda non potrà che essere maggiore o uguale dei massimi delle code successive, e cioè che la sotto-successione $x_{n(k)}$ è debolmente decrescente.

$\square$

La condizione 1 rispecchia il mito moderno del progresso, nonché l'unico mito attuale: è sempre possibile superare i picchi del passato. Viceversa, la condizione 2 rappresenta il mito dell'età dell'oro, la mamma di tutti i miti: c'è stato un massimo in passato, ora si può solo decrescere.

A lezione di filosofia della Storia con Pietro Majer

12.3 Teorema di Bolzano-Weierstrass

Ogni successione limitata di reali ammette una sotto-successione convergente.
L'enunciato è un corollario del teorema precedente (se la successione monotona è limitata il suo limite esiste finito), ma ha una sua pregnanza notevole. Generalizzando, è possibile dire che ogni successione limitata in $\mathbb{R}^n$, con $n \in \mathbb{N}$, ammette una sotto-successione convergente. Per dimostrarlo, basta "proiettare" la successione sulla componente x ed estrarre una sotto-successione convergente. Dopodiché, si proietta la sotto-successione sull'asse y e si estrae una sotto-sotto-successione le cui componenti x e y convergono contemporaneamente. Reiterano di procedimento un numero finito di volte segue la tesi.
Si osservi che è equivalente dire che un sottoinsieme chiuso e limitato di $\mathbb{R}^n$ è compatto[17]. La limitatezza equivale a richiedere che i limiti non siano infiniti; la chiusura a chiedere che i limiti stiano ancora nello sottoinsieme.

12.4 (*)Teorema delle contrazioni

Questo teorema è sempre stato implicitamente assunto a lezione, ma non lo si è mai formalizzato con le giuste ipotesi. Se ne fornisce qui l'enunciato con la relativa dimostrazione.

Definizione Una funzione continua è una contrazione se è lipschitziana di costante strettamente minore di 1, ossia se

$$|f(x) - f(y)| \leq L\,|x - y| \quad \text{con } L < 1$$

Teorema delle contrazioni Una contrazione a valori in uno spazio metrico completo ammette uno e un solo punto fisso, il quale è il limite della successione

$$\begin{cases} x_0 = \alpha, \quad \alpha \in I \\ x_{n+1} = f(x_n) \end{cases}$$

(*)Dimostrazione L'unicità segue banalmente dalla lipschitzianità di costante $L < 1$. Infatti, se ci fossero due punti fissi x e y, si avrebbe che $|x-y| = |f(x) - f(y)| < |x-y|$, il che è assurdo. Per l'esistenza, consideriamo la successione x_n definita nel testo. La distanza tra due punti successivi vale

$$|x_n - x_{n+1}| = |f(x_{n-1}) - f(x_n)| \leq L\,|x_{n-1} - x_n| \leq \ldots \leq L^n\,|x_0 - x_1|$$

Applicando la disuguaglianza triangolare e la proprietà appena mostrata, si ottiene che la distanza tra due punti x_n e x_m, con $n > m$, vale

$$|x_n - x_m| \leq |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \ldots + |x_{m+1} - x_m|$$

[17]Questo risultato è noto anche come teorema di Heine-Borel, e vale soltanto in $\mathbb{R}^n$. In generale, non è vero che un sottoinsieme chiuso e limitato di uno spazio metrico è compatto, ma vale sempre il viceversa, ossia il fatto che un insieme compatto è chiuso e limitato. Si veda a proposito 7.8.

$$|x_n - x_m| \leq |x_1 - x_0| \left(L^n + \ldots + L^m\right) \leq |x_1 - x_0| \sum_{k=m}^{+\infty} L^k$$

La somma alla RHS è la coda della serie geometrica, che è infinitesima poiché la serie converge. Dunque la distanza tra due punti della successione è infinitesima, e cioè la successione è di Cauchy. Ciò implica l'esistenza del limite, poiché lo spazio è completo e cioè ogni successione di Cauchy converge. Sia $\bar{x}$ il limite. Ora, poiché le funzioni lipschitziane sono continue, e dunque sequenzialmente continue, si ha che

$$f(\bar{x}) = \lim_{n \to +\infty} f(x_n) = \lim_{n \to +\infty} x_{n+1} = \bar{x}$$

Ossia il limite della successione è un punto fisso della funzione.

□

12.5 (*)Cenni al Teorema di Stolz-Cesaro

Siano a_n, b_n due successioni a valori reali. Se

1. b_n è strettamente monotona;
2. Vale una delle due condizioni seguenti:
   - $\lim_{n \to +\infty} a_n = \lim_{n \to +\infty} b_n = 0$
   - $\lim_{n \to +\infty} b_n = \pm\infty$
3. Esiste $\lim_{n \to +\infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n} = L \in [-\infty, +\infty]$

Allora esiste il limite del rapporto $\frac{a_n}{b_n}$ e vale L.
Il teorema può essere visto come la controparte discreta del teorema di de l'Hospital. La dimostrazione sono lunghe e noiose verifiche in ciascuno dei quattro casi possibili (L finito o infinito, $\lim b_n = 0$ oppure ∞).

13 Funzione esponenziale complessa

13.1 La successione $(1+x/n)^n$

Consideriamo la successione $(1+x/n)^n$.

- La successione è crescente per $n > x^-$, dove x^- è la parte negativa di x. Infatti, dalla disuguaglianza delle medie segue che

$$\sqrt[n+1]{\left(1+\frac{x}{n}\right)^n} < \frac{1+n\left(1+\frac{x}{n}\right)}{n+1} = 1+\frac{x}{n+1}$$

  Da cui segue la tesi elevando alla $n+1$ ambo i membri. La condizione $n > x^-$ è necessaria perché sia ben definita la media geometrica, che richiede argomento positivo.

- Si definisce la funzione "esponenziale di x" come

$$\exp(x) = \lim_{n\to+\infty}\left(1+\frac{x}{n}\right)^n \tag{9}$$

  Tale quantità esiste, essendo la successione strettamente crescente, ed è sempre positiva perché definitivamente $(1+x/n) > 0$. Particolare rilevanza ha il valore che la funzione assume in $x=1$, detto numero di Nepero e indicato con e.

- Vale che $\exp(x) \geq 1+x$.
  La disuguaglianza è banalmente vera per $x < -1$. Negli altri casi si sfrutta la monotonia della successione per dire che

$$1+x = \left(1+\frac{x}{1}\right)^1 \leq \left(1+\frac{x}{n}\right)^n \leq \exp(x)$$

- $\exp(x)$ è strettamente crescente, e dunque assume sempre valori reali minori di $+\infty$. Infatti, sia $y > x$. Eseguendo la differenza tra i termini n-esimi delle relative successioni e sviluppando la differenza tra potenze n-esime si ottiene

$$\left(1+\frac{y}{n}\right)^n - \left(1+\frac{x}{n}\right)^n = \left(\frac{y-x}{n}\right)\sum_{k=0}^{n-1}\left(1+\frac{y}{n}\right)^k\left(1+\frac{x}{n}\right)^{n-k-1}$$

  Sostituendo tutti i fattori con il massimo o con il minimo si ottiene che

$$(y-x)\left(1+\frac{x}{n}\right)^{n-1} \leq \left(1+\frac{y}{n}\right)^n - \left(1+\frac{x}{n}\right)^n \leq (y-x)\left(1+\frac{y}{n}\right)^{n-1}$$

  Consideriamo ora il limite per $n\to+\infty$. Si osserva che il limite dei termini laterali vale

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

  dove l'ultimo passaggio segue dal fatto che $\lim_{n\to+\infty}\left(1+\frac{y}{n}\right)^{-1} = 1$. Si conclude dunque che

$$0 < (y-x)\exp(x) \leq \exp(y)-\exp(x) \leq (y-x)\exp(y)$$

  La disuguaglianza a sinistra dice che $\exp(x)$ è strettamente crescente, essendo positiva la differenza tra le immagini. La parte a destra dice che, nell'intervallo $]-\infty;y]$ la funzione è lipschitziana di costante $\exp(y)$, ossia che la funzione è localmente lipschitziana. Da ciò consegue che è continua, e in particolare localmente uniformemente continua.

- Se una successione x_n tende a $x \in \mathbb{R}$, allora $\lim_{n\to+\infty}\left(1+\frac{x_n}{n}\right)^n = \exp(x)$. Infatti, la successione x_n è convergente e dunque limitata superiormente da un qualche C applicando la disuguaglianza precedente a x_n e x, si ha

$$0 \leq \left|\left(1+\frac{x}{n}\right)^n - \left(1+\frac{x_n}{n}\right)^n\right| \leq |x_n - x|\exp(C)$$

  Poiché la RHS tende a zero, per teorema della Pula si conclude quanto voluto.

13.2 $\exp(x)$ come isomorfismo di gruppi

Si mostra ora che $\exp(x)$ è un isomorfismo di gruppi ordinati tra $(\mathbb{R},+)$ e $(\mathbb{R}^+,*)$.

- Si è già mostrato che $\exp(x)$ è strettamente monotona.
- Mostriamo che è lineare, ossia che $\exp(a+b)=\exp(a)\exp(b)$. Si ha che

$$\left(1+\frac{a}{n}\right)^n\left(1+\frac{b}{n}\right)^n=\left(1+\frac{a+b+\frac{ab}{n}}{n}\right)^n$$

  Poiché il numeratore della frazione tende a $a+b$, per lemma precedente il limite del prodotto è $\exp(a+b)$, come voluto.

- Manca solo mostrare che è suriettiva. Sia allora $c\in\mathbb{R}^+$ e mostriamo che esiste un x tale che $\exp(x)=c$. WLOG, assumiamo che $c\geq 1$, dato che il caso complementare si ottiene passando ai reciproci. Definiamo la successione $x_n=n\,(\sqrt[n]{c}-1)$.Si osserva che invertendo la definizione si ottiene $c=\left(1+\frac{x_n}{n}\right)^n\ \forall n$.
  Ora, la successione è sempre non negativa essendo $c\geq 1$. Inoltre, è crescente, in quanto

$$\left(1+\frac{x_{n+1}}{n+1}\right)^{n+1}=c=\left(1+\frac{x_n}{n}\right)^n\leq\left(1+\frac{x_n}{n+1}\right)^{n+1}$$

  L'ultima disuguaglianza segue dal fatto che la successione $\left(1+\frac{x_n}{n}\right)^n$, fissato x_n, è crescente in n. Dalla disuguaglianza segue $x_n\geq x_{n+1}$. Dunque la successione ammette limite reale $\bar{x}$. Vale allora che, applicando il lemma dell'esponenziazione di successioni,

$$c=\lim_{n\to+\infty}c=\lim_{n\to+\infty}\left(1+\frac{x_n}{n}\right)^n=\lim_{n\to+\infty}\left(1+\frac{\bar{x}}{n}\right)^n=\exp(\bar{x})$$

  Si è dunque mostrata una controimmagine di c.

□

13.3 Serie esponenziale

La convergenza della successione $\left(1+\frac{x_n}{n}\right)^n$ è discretamente lenta, perciò è utile definire l'esponenziale come serie di potenze, che converge molto più rapidamente. Sviluppiamo il termine n-esimo della successione come potenza di un binomio.

$$\exp(x)=\lim_{n\to+\infty}\left(1+\frac{x_n}{n}\right)^n=\lim_{n\to+\infty}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{x}{n}\right)^k=\lim_{n\to+\infty}\sum_{k=0}^{+\infty}\binom{n}{k}\left(\frac{x}{n}\right)^k$$

Sviluppiamo il binomiale col teorema dei tre fattoriali:

$$\exp(x)=\lim_{n\to+\infty}\sum_{k=0}^{+\infty}\frac{n!}{k!(n-k)!}\left(\frac{x}{n}\right)^k=\lim_{n\to+\infty}\sum_{k=0}^{+\infty}\frac{n(n-1)\dots(n-k+1)}{n^k}\frac{x^k}{k!}$$

Si osserva che il termine generale della sommatoria, che è una successione in n, è maggiorato dalla successione $b_k(n)=\frac{x^k}{k!}=\text{cost}$. Inoltre, la sommatoria su k dei b_k è finita. Infatti, definitivamente vale che

$$b_k=\frac{x^k}{k!}\geq 2\frac{x^{k+1}}{(k+1)!}=b_{k+1}$$

cioè definitivamente ogni termine è minore della metà del precedente. Questo significa che la serie converge, perché definitivamente cresce più lentamente della serie geometrica di ragione $1/2$. (*)Più formalmente, la tesi segue per criterio del rapporto generalizzato.

Valgono dunque le ipotesi del teorema di convergenza dominata, e dunque il limite della sommatoria è la sommatoria dei limiti delle singole successioni. Vale dunque che

$$\exp(x) = \lim_{n\to+\infty} \sum_{k=0}^{+\infty} \frac{n(n-1)\dots(n-k+1)}{n^k}\frac{x^k}{k!} = \sum_{k=0}^{+\infty} \lim_{n\to+\infty} \frac{n(n-1)\dots(n-k+1)}{n^k}\frac{x^k}{k!}$$

Si giunge dunque alla seguente espressione

$$\exp(x) = \sum_{k=0}^{+\infty} \frac{x^k}{k!} \tag{10}$$

13.4 Stima del resto

Cerchiamo di quantificare la velocità con cui la successione $\left(1+\frac{1}{n}\right)^n$ e la corrispondente serie convergono a e, tramite la stima della differenza tra il termine n-esimo e il limite.

- Studiamo prima la convergenza della serie.

  $$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| = \left| \sum_{k=n+1}^{+\infty} \frac{1}{k!} \right| \leq \sum_{k=n+1}^{+\infty} \frac{1}{k!}$$

  Dove l'ultimo passaggio segue per disuguaglianza triangolare. Posto $j = k - n - 1$, si ha

  $$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| \leq \frac{1}{(n+1)!} \sum_{j=0}^{+\infty} \frac{1}{(n+1)^j}$$

  Dove l'ultima disuguaglianza segue stimando il fattoriale con il primo termine ripetuto j volte. Sviluppando come serie geometrica si ottiene

  $$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| \leq \frac{1}{(n+1)!} \frac{1}{1-\frac{1}{n+1}} = \frac{n+1}{(n+1)!(n+1-1)} = \frac{1}{n!n}$$

  Si osserva che la serie converge molto rapidamente, essendo altissimo il tasso di divergenza del fattoriale.

- Consideriamo adesso la successione $\left(1+\frac{1}{n}\right)^n$. Si omette il modulo perché si è già mostrato che la successione converge ad e dal basso.

  $$e - \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{+\infty} \frac{1}{k!} - \left(1+\frac{1}{n}\right)^n$$

  Sviluppando la potenza del binomio e raccogliendo i termini dello stesso grado, si ottiene

  $$e - \left(1+\frac{1}{n}\right)^n = \sum_{k=n+1}^{+\infty} \frac{1}{k!} + \left(1 - \binom{n}{0}\right) + \left(1 - \frac{1}{n}\binom{n}{1}\right) + \left(\frac{1}{2} - \frac{1}{n^2}\binom{n}{2}\right)\dots$$

  Tutti i termini sono positivi, in quanto ciascuna successione $\frac{1}{n^k}\binom{n}{k}$ converge dal basso a $\frac{1}{k!}$. Possiamo pertanto sottostimare il resto con il primo termine non nullo:

  $$e - \left(1+\frac{1}{n}\right)^n \geq \frac{1}{2} - \frac{1}{n^2}\binom{n}{2} = \frac{1}{2} - \frac{n-1}{2n} = \frac{1}{2n}$$

  Il resto ottenuto è enormemente maggiore del resto della serie: dopo 500 termini il la serie approssima e con una precisione di circa 1500 cifre, mentre la successione è precisa solo per le prime 3.

13.5 Irrazionalità di e

Una caratteristica dei numeri razionali è che, prima o poi, il loro prodotto con il fattoriale è un intero. Mostriamo che ciò non accade per e. Infatti

$$n!e = n!\sum_{k=0}^{+\infty}\frac{1}{k!} = n!\sum_{k=0}^{n}\frac{1}{k!} + n!\sum_{k=n+1}^{+\infty}\frac{1}{k!}$$

Il primo termine è un intero. Il secondo può essere stimato con il resto n-esimo della serie:

$$0 < n!\sum_{k=n+1}^{+\infty}\frac{1}{k!} \leq n!\frac{e}{(n+1)!} = \frac{e}{n+1} < 1$$

Dove l'ultima disuguaglianza vale per $n \geq 2$. Dunque, la somma dei due membri non può mai essere un intero. Ne segue che e è irrazionale. É possibile, ma non è nelle mie facoltà, mostrare che e è anche trascendente.

□

13.6 Esponenziale complessa

É possibile definire la funzione esponenziale anche in $\mathbb{C}$, praticamente ripetendo quanto appena detto. L'unica differenza è che non è iniettiva, e dunque non ammette inversa, come omomorfismo da $(\mathbb{C}, +)$ in $(\mathbb{C}\setminus\{0\}, *)$. Ad esempio, 1 ha come controimmagine l'insieme infinito $\left\{e^{2\pi ki}, \quad k \in \mathbb{Z}\right\}$. Consideriamo ora il modulo di $\exp(z)$, con $z \in \mathbb{C}$. Per definizione

$$|e^z| = \sqrt{e^z e^{\bar{z}}} = e^{(z+\bar{z})/2} = e^{Re(z)}$$

Cioè il modulo dipende solo dalla parte reale dell'esponente. La parte immaginaria definisce l'argomento, ossia l'angolo che il vettore e^z descrive nel piano di Argand-Gauss.

13.7 $|e^{ix} - e^{iy}| \leq |x - y|$

In pratica, mostreremo che il modulo della differenza tra due vettori unitari è minore dell'arco compreso tra i due. In altri termini, si può anche dire che la funzione che associa a un reale x il vettore di modulo 1 e angolo x è lipschitziana di costante 1.

Dimostrazione Riduciamoci al caso $y = 0$. Il caso generale segue moltiplicando da ambo i lati per $|e^{iy}| = 1$. Vogliamo allora mostrare che $\left|e^{ix} - 1\right| \leq |x|$.
Majer qui ha fatto un conto molto brutto. A mio avviso, col senno di poi basta considerare i due vettori che da $(0, 0)$ puntano a $(1, 0)$ e $(\cos x, \sin x)$. La loro differenza è la corda che congiunge questi due punti, la cui lunghezza è sicuramente minore o uguale dell'arco sotteso, lungo x.

□

14 Funzioni trigonometriche

14.1 Definizione di seno e coseno

Si definiscono come la parte reale e immaginaria di e^{ix}.

$$\begin{cases} \cos x = Re\left(e^{ix}\right) = \frac{e^{ix}+e^{-ix}}{2} = \sum_{k=0}^{+\infty}(-1)^k \frac{x^{2k}}{(2k)!} \\ \\ \sin x = Im\left(e^{ix}\right) = \frac{e^{ix}-e^{-ix}}{2i} = \sum_{k=0}^{+\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!} \end{cases}$$

Tutte le proprietà note delle funzioni goniometriche sono ricavabili da questa definizione, sebbene sia abbastanza difficile fare vedere che sono periodiche.

14.2 Definizione di π

Consideriamo la funzione $\cos x$ come serie di potenze. Si osserva che per $|x| \leq 2$ il termine generale della serie è decrescente in modulo da $k = 1$. Infatti

$$\frac{x^{2k}}{(2k)!} \geq \frac{x^{2k+2}}{(2k+2)!} \quad \longrightarrow \quad 4k^2 + 6k + 2 \geq x^2 \quad \longrightarrow \quad 4 \geq x^2$$

Dunque, per criterio di Leibniz per le serie a segni alterni, il limite della somma è compreso tra le somme pari e le somme dispari. In particolare, $1 - x^2/2 \leq \cos x \leq 1 - x^2/2 + x^4/24$. Essendo $|x| \leq 2$ il termine a destra è minore o uguale di $1 - x^2/2 + 4x^2/24 = 1 - x^2/3$, perciò si giunge a

$$1 - x^2/2 \leq \cos x \leq 1 - x^2/3$$

Per conseguenza del teorema di esistenza degli zeri, il più piccolo zero positivo del coseno si trova tra gli zeri delle due parabole ai lati. Si definisce π il doppio di tale numero. Vale allora che

$$\pi = 2\min\{x \in \mathbb{R}^+ \text{ t.c. } \cos x = 0\}, \quad \sqrt{2} \leq \pi/2 \leq \sqrt{3}$$

Dalla prima relazione fondamentale della trigonometria segue che $\sin \pi/2 = 1$, da cui

$$e^{i\pi/2} = i \sin \pi/2 + \cos \pi/2 = i$$

Da questa formula derivano due cose. La prima è che il minimo esponente positivo τ per cui $e^{i(x+\tau)} = e^{ix}$ è 2π, dato che $e^{2i\pi} = 1$. In altri termini, seno e coseno sono periodiche di periodo 2π. La seconda è la formula ritenuta la più bella di tutta la matematica:

$$e^{i\pi} + 1 = 0$$

14.3 Sviluppo alternativo del seno

Nella sezione 4.2 abbiamo mostrato la seguente formula

$$(1+z)^n - (1-z)^n = 2nz \prod_{0<k<n/2} \left(1 + \frac{z^2}{\tan^2 \frac{\pi k}{n}}\right)$$

Applichiamo tale sviluppo in $z = \frac{ix}{n}$, con $x \in \mathbb{R}$. Alla LHS si ottengono le successioni i cui limiti sono rispettivamente e^{ix} e e^{-ix}. Dunque

$$\left(1 + \frac{ix}{n}\right)^n - \left(1 - \frac{ix}{n}\right)^n = 2n\frac{ix}{n} \prod_{0<k<n/2} \left(1 - \frac{x^2}{n^2 \tan^2 \frac{\pi k}{n}}\right)$$

Dividiamo per $2i$ e passiamo al limite. Alla LHS otteniamo, dividendo per $2i$

$$\lim_{n\to+\infty} \frac{\left(1 + \frac{ix}{n}\right)^n - \left(1 - \frac{ix}{n}\right)^n}{2i} = \frac{e^{ix} - e^{-ix}}{2i} = \sin x$$

Dunque
$$\sin x = \lim_{n\to+\infty} x \prod_{0<k<n/2} \left(1 - \frac{x^2}{\left(n\tan\frac{\pi k}{n}\right)^2}\right)$$
Il limite del prodotto è il prodotto dei limiti per la variante per i prodotti del teorema di convergenza dominata. Per applicare il teorema dobbiamo mostrare che il termine generale del prodotto è definitivamente limitato da una costante, e che il prodotto delle costanti converge. Intanto, osserviamo che il termine generale è decrescente. Consideriamo solo l'andamento del termine $n\tan\frac{\pi k}{n}$ al variare di n. Assumiamo che sia una funzione di variabile reale y e calcoliamone la derivata.
$$\frac{d}{dy}\left(y\tan\frac{\pi k}{y}\right) = \tan\frac{\pi k}{y} + y\frac{1}{\cos^2\frac{\pi k}{y}}\frac{-\pi k}{y^2} = \tan\frac{\pi k}{y} - \frac{\pi k}{y\cos^2\frac{\pi k}{y}} = \frac{y\sin\frac{\pi k}{y}\cos\frac{\pi k}{y} - \pi k}{y\cos^2\frac{\pi k}{y}} = \frac{\frac{\sin\frac{2\pi k}{y}}{2/y} - \pi k}{y\cos^2\frac{\pi k}{y}}$$
Ora, per $y \to +\infty$ il rapporto $\frac{\sin\frac{2\pi k}{y}}{2/y}$ tende a πk dal basso. Dunque il numeratore è negativo, e dato che il denominatore deve essere positivo tutta la derivata è negativa.
Ora, se la successione $n\tan\frac{\pi k}{n}$ è decrescente, lo è anche il termine generale della produttoria $\left(1 - \frac{x^2}{n^2\tan^2\frac{\pi k}{n}}\right)$ Perciò, poiché definitivamente il denominatore è positivo e superiormente limitato, avremo che
$$\left(1 - \frac{x^2}{n^2\tan^2\frac{\pi k}{n}}\right) < 1$$
Ora, $\frac{x^2}{n^2\tan^2\frac{\pi k}{n}} \leq C\frac{1}{k^2} := b_k$ per una certa costante C indipendente da k e da n. Visto che $\sum_k b_k < +\infty$ possiamo applicare il teorema di convergenza dominata (vedasi 21.6), dunque possiamo scambiare il limite e la produttoria ottenendo
$$\sin x = \lim_{n\to+\infty} x \prod_{0<k<n/2} \left(1 - \frac{x^2}{\left(n\tan\frac{\pi k}{n}\right)^2}\right) = x\prod_{k\in\mathbb{N}} \lim_{n\to+\infty} \left(1 - \frac{x^2}{\left(n\tan\frac{\pi k}{n}\right)^2}\right) = x\prod_{k\in\mathbb{N}}\left(1 - \frac{x^2}{\pi^2 k^2}\right)$$
Ossia
$$\sin x = x\prod_{k\in\mathbb{N}}\left(1 - \frac{x^2}{\pi^2 k^2}\right) \tag{11}$$

14.4 Prodotto di Wallis

Calcolando la formula alternativa dello sviluppo del seno, mostrata nella sezione 14.3 in $\pi/2$ si giunge a un'approssimazione di π. Infatti
$$\sin\pi/2 = \pi/2\prod_{k\in\mathbb{N}}\left(1 - \frac{1}{4k^2}\right) = \pi/2\prod_{k\in\mathbb{N}}\frac{2k-1}{2k}\frac{2k+1}{2k}$$
Poiché, come è noto, $\sin\pi/2 = 1$, si giunge a
$$\pi = 2\prod_{k\in\mathbb{N}}\frac{2k}{2k-1}\frac{2k}{2k+1} = 2\left(\frac{2}{1}\frac{2}{3}\frac{4}{3}\frac{4}{5}\dots\right) \tag{12}$$
La formula è nota come prodotto di Wallis. Va detto che la convergenza di questa produttoria è discretamente lenta. Nella sezione 5.6 si mostra una stima più rapida di π:
$$\pi = 4 - \frac{4}{3} + \frac{4}{5} - \frac{4}{7}\dots \tag{13}$$

14.5 Problema di Basilea

Nel 1644[18] Pietro Mengoli (sì, proprio quello della serie) pose alla comunità scientifica il seguente problema: a cosa tende la serie dei reciproci dei quadrati dei naturali? Il fatto che convergesse era

[18] Funfact: Le prime quattro cifre del risultato coincidono con l'anno in cui è stato presentato il problema. Coincidenze? Io non credo.

noto da tempo, ma non si riusciva a trovare il valore del limite. La soluzione fu trovata da Carl Friedrich Gauss (sì, proprio quello di praticamente tutta la matematica classica) solo nel 1735. Mostreremo ora, sebbene non seguendo la dimostrazione del mitico Gauss, che la serie converge a

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \tag{14}$$

Dimostrazione Dalla formula alternativa dello sviluppo del seno sappiamo che

$$\sin x = x \prod_{k \in \mathbb{N}} \left(1 - \frac{x^2}{\pi^2 k^2}\right)$$

É possibile sviluppare la produttoria come somma dei prodotti parziali tra i termini, come mostrato in 16.7

$$\sin x = x \sum_{J \subseteq \mathbb{N} \text{ finito}} (-1)^{|J|} \prod_{k \in J} \frac{x^2}{\pi^2 k^2}$$

Raccogliamo i prodotti in modo da mettere insieme tutti gli insiemi J con la stessa cardinalità. Otteniamo che

$$\sin x = x \left(1 - \sum_{k \in \mathbb{N}} \frac{x^2}{\pi^2 k^2} + \sum_{k,\, h \in \mathbb{N}}^{k>h} \frac{x^4}{\pi^4 k^2 h^2} \dots\right)$$

Compariamo questo sviluppo con lo sviluppo in serie del seno, derivante dalla serie esponenziale:

$$\sin x = x - \sum_{k \in \mathbb{N}} \frac{x^3}{\pi^2 k^2} + \sum_{k,\, h \in \mathbb{N}}^{k>h} \frac{x^5}{\pi^4 k^2 h^2} + \mathcal{O}(x^7) = x - \frac{x^3}{6} + \frac{x^5}{120} + \mathcal{O}(x^7) \tag{15}$$

Dato che lo sviluppo polinomiale infinito deve essere unico, i termini dello stesso grado devono essere uguali. Dunque, in particolare, eguagliando i termini di grado 3, si ha

$$\sum_{k \in \mathbb{N}} \frac{1}{\pi^2 k^2} = \frac{1}{6}$$

Da cui, moltiplicando per π^2, segue la tesi.

$$\sum_{k \in \mathbb{N}} \frac{1}{k^2} = \frac{\pi^2}{6}$$

□

14.6 (*) Somma delle potenze quarte

Procedendo a comparare i termini dell'espressione 15, si ottiene la seguente formula

$$\sum_{k,\, h \in \mathbb{N}}^{k>h} \frac{1}{k^2 h^2} = \frac{\pi^4}{120}$$

che, in sé, è poco utile. Tuttavia, possiamo sfruttare il risultato del problema di Basilea. Infatti

$$\left(\sum_{k \in \mathbb{N}} \frac{1}{k^2}\right)^2 = \frac{\pi^4}{36}$$

Sviluppiamo la RHS:

$$\left(\sum_{k \in \mathbb{N}} \frac{1}{k^2}\right)^2 = \left(\sum_{k \in \mathbb{N}} \frac{1}{k^2}\right)\left(\sum_{h \in \mathbb{N}} \frac{1}{h^2}\right) = \sum_{k,\, h \in \mathbb{N}} \frac{1}{k^2 h^2} = 2 \sum_{k,\, h \in \mathbb{N}}^{k>h} \frac{1}{k^2 h^2} + \sum_{k,\, h \in \mathbb{N}}^{k=h} \frac{1}{k^4}$$

Sostituendo troviamo allora che

$$\frac{\pi^4}{36} = \left(\sum_{k\in\mathbb{N}} \frac{1}{k^2}\right)^2 = 2\sum_{k,\,h\in\mathbb{N}}^{k>h} \frac{1}{k^2h^2} + \sum_{k,\,h\in\mathbb{N}}^{k=h} \frac{1}{k^4} = \frac{\pi^4}{60} + \sum_{k,\,h\in\mathbb{N}}^{k=h} \frac{1}{k^4}$$

Da cui

$$\sum_{k,\,h\in\mathbb{N}}^{k=h} \frac{1}{k^4} = \pi^4\left(\frac{1}{36} - \frac{1}{60}\right) = \frac{\pi^4}{90}$$

$\square$

15 Sommatorie e serie

15.1 Somme di reali positivi

Data una famiglia di reali positivi, indicizzati su un insieme I, si definisce la loro somma come l'estremo superiore delle somme su sottoinsiemi di indici finiti.

$$\sum_{i\in I} a_i = \sup \left\{ \sum_{j\in J} a_j, \quad J \subseteq I, \quad , |J| < \aleph_0 \right\}$$

Valgono le seguenti proprietà:

1. MONOTONIA V1: Se $a_i \leq b_i\ \forall i \in I$ allora $\sum a_i \leq \sum b_i$;
2. MONOTONIA V2: Se $I' \subseteq I$ allora $\sum_{I'} a_i \leq \sum_I a_i$;
3. ASSOCIATIVITA': Sia $\{I_K\}_{K\in\Lambda}$ una partizione dell'insieme degli indici. Allora la somma su tutto l'insieme è la somma delle somme parziali sui sottoinsiemi.

$$\sum i \in I a_i = \sum_{K\in\Lambda} \left(\sum_{i\in I_K} a_i \right)$$

4. Il prodotto di due sommatorie positive è la somma dei prodotti.

$$\sum_{i\in I,\, j\in J} a_i b_j = \left(\sum_{i\in I} a_i \right) \left(\sum_{j\in J} b_j \right)$$

5. Per una somma indicizzata su $\mathbb{N}^2$ vale la seguente formula

$$\sum_{(i,j)\in\mathbb{N}^2} a_{i,j} = \sum_{n\in\mathbb{N}} \sum_{i+j=n} a_{i,j}$$

che converge sse convergono entrambe le somme.

Dimostrazione Le prime due proprietà sono banali, la quarta e la quinta seguono dalla terza. Si dimostra la doppia disuguaglianza per la proprietà associativa.

- Comunque si scelga un sottoinsieme $J \subseteq I$ finito di indici, esiste un insieme finito $H \subseteq \Lambda$ tale che $J \subseteq \bigcup_{K\in H} I_K$. Allora per monotonia

$$\sum_{i\in J} a_i \leq \sum_{K\in H} \sum_{i\in K} a_i \leq \sum_{K\in\Lambda} \sum_{i\in K} a_i$$

Poiché la disuguaglianza vale per ogni somma su un J finito, passando al sup si ha che

$$\sum_{i\in I} a_i \leq \sum_{K\in\Lambda} \sum_{i\in K} a_i$$

- Viceversa, dato comunque un sottoinsieme finito di indici $H \subseteq \Lambda$, si ha che, per monotonia

$$\sum_{K\in H\wedge i\in I_K} a_i \leq \sum_{i\in I} a_i$$

La disuguaglianza si conserva passando al sup a sinistra, perciò

$$\sum_{K\in\Lambda} \sum_{i\in K} a_i \leq \sum_{i\in I} a_i$$

□

15.2 Termini non nulli di una sommatoria convergente

Se una sommatoria di reali positivi converge, ossia assume valore reale non infinito, allora l'insieme di indici per cui $a_i \neq 0$ è al più numerabile.

Dimostrazione Consideriamo gli insiemi $A_n = \left\{ i \in I \quad | \quad a_i > \frac{1}{n} \right\}$. Ciascuno di questi insiemi deve essere finito, altrimenti la somma non sarebbe finita, in quanto l'insieme dei multipli di $\frac{1}{n}$ è illimitato in $\mathbb{R}$ per proprietà di Archimede. Ora, l'insieme degli indici per cui $a_i \neq 0$ è l'unione degli A_n, indicizzata su $\mathbb{N}$. Si conclude osservando che l'unione numerabile di insiemi al più numerabili è numerabile, cosa che segue facilmente dal fatto che $\mathbb{N}$ è in bigezione con $\mathbb{N}^2$.

□

15.3 Somme di famiglie generiche di reali

Data una famiglia di reali $(a_i)_{i \in I}$, se vale che

- $\forall i \in I \ \exists p_i \geq 0, \quad n_i \geq 0$ reali tali che $a_i = p_i - n_i$;
- $\sum_{i \in I} p_i < +\infty$;
- $\sum_{i \in I} n_i < +\infty$;

Allora è possibile definire la somma degli a_i come

$$\sum_{i \in I} a_i = \sum_{i \in I} p_i - \sum_{i \in I} n_i$$

Dove le due somme alla RHS sono somme di reali positivi, definite come estremo superiore delle somme parziali.
Si osserva che la definizione è ben posta, poiché prescinde dalla scelta dei p_i e degli n_i. Infatti, sia $a_i = p_i - n_i = p'_i - n'_i$. Allora separando i termini e sommando si ottiene

$$\sum_{i \in I} p_i + \sum_{i \in I} n'_i = \sum_{i \in I} n_i + \sum_{i \in I} p'_i$$

Poiché nessuna delle somme considerate è illimitata, è possibile separare i termini e ottenere che

$$\sum_{i \in I} a_i = \sum_{i \in I} p_i - \sum_{i \in I} n_i = \sum_{i \in I} p'_i - \sum_{i \in I} n'_i$$

Risulta allora comodo scegliere p_i come la parte positiva di a_i, e analogamente n_i come la sua parte negativa. Tali quantità si definiscono come

$$a^+ = \max(a, 0) \quad a^- = \max(-a, 0) \quad a^+ + a^- = |a|$$

Si osserva che la somma tra la parte positiva e la parte negativa è il modulo di a. Questo significa che, se valgono le condizioni dette sopra, si può scrivere

$$\sum_{i \in I} |a_i| = \sum_{i \in I} a^+{}_i + \sum_{i \in I} a^-{}_i$$

Se le somme alla RHS sono finite, anche la somma dei moduli è finita, e viceversa se la somma dei moduli è finita esiste sempre una decomposizione in parte positiva e negativa.
Se vale che $\sum_{i \in I} |a_i| < +\infty$, la famiglia $(a_i)_{i \in I}$ si dice assolutamente sommabile.
É possibile dimostrare che valgono le stesse proprietà viste in 15.1 che valgono per le somme sui reali positivi, applicando la decomposizione in parte positiva e negativa. In particolare, nella proprietà sul prodotto tra somme, se le singole famiglie sono assolutamente sommabili lo è anche la famiglia dei prodotti.

Osservazione Il concetto di somma per famiglie assolutamente sommabili è intrinsecamente commutativo, ossia prescinde dall'ordine in cui vengono sommati i termini. Più formalmente, comunque si scelga una permutazione σ dell'insieme I degli indici, vale che

$$\sum_{i\in I} |a_i| < +\infty \quad \Longrightarrow \quad \sum_{i\in I} a_i = \sum_{i\in I} a_{\sigma(i)}$$

Invece, se una serie converge ma non assolutamente, l'ordine con cui si sommano gli addendi è fondamentale. Si dimostra che, comunque si scelga un numero reale L, esiste una permutazione degli indici tale che la serie permutata tende a L. Si dice che la convergenza assoluta è incondizionata, mentre non lo è la convergenza semplice.

15.4 Prodotti di famiglie di reali positivi

Sfruttando l'isomorfismo $\exp(x)$ tra $(\mathbb{R}, +)$ e $(\mathbb{R}^+, *)$, è possibile definire il prodotto infinito di una famiglia di numeri reali positivi come

$$\prod_{i\in I} a_i = \exp \sum_{i\in I} \log(a_i)$$

Tale definizione permette di traslare tutte le proprietà viste per le somme sui prodotti. In particolare, la definizione di somma per reali positivi si traduce in una definizione di prodotto di reali positivi maggiori di 1:

$$\forall i \in I (b_i \geq 1) \quad \Longrightarrow \quad \prod_{i\in I} b_i = \sup \left\{ \prod_{i\in J} b_i, \quad J \subseteq I \text{ finito} \right\}$$

Il prodotto converge se tende a un limite diverso da zero (che corrisponde a $e^{-\infty}$).

16 Calcolo di serie e produttorie

16.1 Criterio di condensazione di Cauchy

Sia a_n una successione debolmente decrescente di reali positivi. Allora vale che

$$\sum_{i\in\mathbb{N}} a_i < +\infty \iff \sum_{k\in\mathbb{N}} 2^k a_{2^k} < +\infty$$

Dimostrazione

- ($\Rightarrow$) Sia $J_k = \left\{n \in \mathbb{N} \text{ t.c. } 2^{k-1} < n \leq 2^k\right\}$. Si osserva che i J_k descrivono una partizione di $\mathbb{N}$, e che ognuno di essi ha cardinalità $|J_k| = 2^{k-1}$. Ora, essendo a_n debolmente decrescente, vale che
$$2^{k-1}a_{2^k} \leq \sum_{n\in J_k} a_n$$
Sommando su $k \in \mathbb{N}$ si ha che
$$\sum_{k\in\mathbb{N}} 2^{k-1}a_{2^k} \leq \sum_{n\in\mathbb{N}} a_n < +\infty$$
Segue allora la tesi, dal momento che
$$\sum_{k\in\mathbb{N}} 2^k a_{2^k} = a_1 + 2\sum_{k\in\mathbb{N}} 2^{k-1}a_{2^k} < +\infty$$
- ($\Leftarrow$) Sia ora $I_k = \left\{n \in \mathbb{N} \text{ t.c. } 2^k \leq n < 2^{k+1}\right\}$. Come prima, gli I_k partizionano $\mathbb{N}$. Sommando su ciascun I_k vale che
$$\sum_{n\in I_k} a_n \leq 2^k a_{2^k}$$
Dunque, sommando su k, segue la tesi.
$$\sum_{n\in\mathbb{N}} a_n \leq \sum_{k\in\mathbb{N}} 2^k a_{2^k} < +\infty$$

□

16.2 Criterio di Abel-Dirichlet

Sommazione per parti di Abel Date due successioni a_n, b_n, sia $B(n)$ la somma parziale della seconda. Allora per ogni n vale la seguente identità:

$$\sum_{k=0}^{n} a_k b_k = B(n)a_{n+1} - \sum_{k=0}^{n} B(k)(a_{k+1} - a_k)$$

La dimostrazione è una banale verifica. Se si sottrae la somma n-esima dalla somma m-esima si ottiene la seguente formula più generale:

$$\sum_{k=n}^{m} a_k b_k = B(m)a_{m+1} - B(n)a_{n+1} - \sum_{k=n+1}^{m} B(k)(a_{k+1} - a_k)$$

Si osservi che la formula è l'equivalente discreto della più nota integrazione per parti.

Criterio di Abel-Dirichlet Sia a_n decrescente e infinitesima e sia $b(n)$ a somme parziali limitate, ossia $|B(n)| < M$ per ogni n. Allora la serie $\sum_{n=0}^{+\infty} a_n b_n$ converge.

Dimostrazione Mostriamo che la successione delle somme parziali è di Cauchy. Fissiamo un ε e procediamo sommando per parti:

$$\left|\sum_{k=n}^{m} a_k b_k\right| = \left|B(m)a_{m+1} - B(n)a_{n+1} - \sum_{k=n+1}^{m} B(k)(a_{k+1} - a_k)\right| \leq$$

$$\leq |B(m)a_{m+1}| + |B(n)a_{n+1}| + \left|\sum_{k=n+1}^{m} B(k)(a_{k+1} - a_k)\right|$$

Ogni $B(n)$ si maggiora in modulo con la costante M. Otteniamo dunque

$$\left|\sum_{k=n}^{m} a_k b_k\right| \leq M\left(|a_{m+1}| + |a_{n+1}| + \left|\sum_{k=n+1}^{m} a_{k+1} - a_k\right|\right) = M\left(|a_{m+1}| + |a_{n+1}| + |a_{m+1} - a_{n+1}|\right)$$

$$\left|\sum_{k=n}^{m} a_k b_k\right| \leq 2M(|a_{m+1}| + |a_{n+1}|)$$

Dato che la successione a_n è infinitesima, definitivamente è minore di $\varepsilon/4M$. Dunque otteniamo che, definitivamente,

$$\left|\sum_{k=n}^{m} a_k b_k\right| \leq \varepsilon$$

Da cui la tesi.

□

16.3 (*)Serie binomiale

Consideriamo la serie a valori z complessi

$$\sum_{k=0}^{+\infty} \binom{a}{k} z^k$$

Dove $a \in \mathbb{R}$ e dove per definizione

$$\binom{a}{k} = \frac{a(a-1)\dots(a-k+1)}{k!}$$

Ci chiediamo per quali valori di z e di a converga. Osserviamo innanzitutto che, per a fissato, la serie è una serie di potenze di z. Questo tipo di serie ammette un raggio di convergenza, ossia esiste $R \in [0, +\infty]$ tale che

- se $|z| < R$, la serie converge assolutamente;
- se $|z| > R$, la serie diverge;
- se $|z| = R$, in generale non si può dire nulla.

In altri termini, l'insieme di convergenza della serie, nel piano complesso, è un cerchio di raggio R, con la convenzione che per $R = 0$ il cerchio coincide con l'origine, mentre per $R = +\infty$ il cerchio è tutto $\mathbb{C}$.
Per determinare il raggio di convergenza, solitamente si applica il criterio della radice alla serie. Per comodità, in questo frangente useremo il criterio del rapporto. Prendiamo allora il limite del modulo del rapporto tra i termini successivi:

$$\lim_{k\to+\infty}\left|\frac{\binom{a}{k+1} z^{k+1}}{\binom{a}{k} z^k}\right| = \lim_{k\to+\infty}\left|\frac{a-k}{k+1}\right| |z| = |z|$$

Dunque otteniamo innanzitutto che se $|z| < 1$ la serie converge assolutamente, mentre per $|z| > 1$ la serie diverge. Rimane ora da studiare il caso $|z| = 1$, al variare di $a \in \mathbb{R}$.
Osserviamo innanzitutto che se $a \in \mathbb{N}$ la somma è finita, dato che essa rappresenta lo sviluppo del binomio di Newton di $(1+z)^a$. Limitiamoci allora ai casi restanti, cercando dapprima i valori di a per cui la serie converge assolutamente. Consideriamo il modulo del termine generale:

$$\left|\binom{a}{k}\right| = \left|\frac{a(a-1)\dots(a-k+1)}{k!}\right| = \left|(-1)^k \frac{(k-a-1)\dots a}{k!}\right| = \left|\frac{(k-a-1)!}{k!(-a-1)!}\right|$$

Abbiamo potuto scrivere $(-a-1)!$ dato che l'argomento non è un numero intero negativo, in quanto si è escluso il caso $a \in \mathbb{N}$. Ora, per sviluppo asintotico del fattoriale, di cui a 20.7, possiamo scrivere

$$\left|\frac{(k-a-1)!}{k!(-a-1)!}\right| = \left|\frac{k!k^{-a-1}}{k!(-a-1)!}(1+o(1))\right| = \frac{1+o(1)}{|(-a-1)!|\, k^{a+1}}$$

Dunque, per confronto con l'armonica generalizzata, la serie converge assolutamente se e solo se $a > 0$.
Chiediamoci ora se ci siano valori per cui la serie converge ma non assolutamente. Partiamo escludendo i casi in cui il termine generale non è infinitesimo. Per quanto già visto, possiamo scrivere

$$\left|\binom{a}{k}\right| = \frac{1+o(1)}{|(-a-1)!|\, k^{a+1}}$$

Se $a \leq -1$ il termine generale non è infinitesimo, dunque sicuramente la serie non converge. Ci siamo ristretti dunque a studiare cosa succeda nel caso $-1 < a \leq 0$. Per cercare di ricondurci al caso $a > 0$ scriviamo la formula di Stiefel per i binomiali:

$$\binom{a+1}{k} = \binom{a}{k} + \binom{a}{k-1}$$

Sommiamo da entrambe le parti per k da 1 a n, moltiplicando ogni termine per z^k. Per completare le somme, aggiungiamo il fatto che $\binom{a+1}{0} = \binom{a}{0} = 1$.

$$\sum_{k=0}^{n} \binom{a+1}{k} z^k = \sum_{k=0}^{n} \binom{a}{k} z^k + \sum_{k=1}^{n} \binom{a}{k-1} z^k =$$

$$= \sum_{k=0}^{n} \binom{a}{k} z^k + z \sum_{k=0}^{n-1} \binom{a}{k} z^k = (1+z) \sum_{k=0}^{n-1} \binom{a}{k} z^k + \binom{a}{n} z^n$$

Da cui, se $z \neq -1$, possiamo esplicitare la somma parziale $n-1$-esima come

$$\sum_{k=0}^{n-1} \binom{a}{k} z^k = \frac{1}{1+z}\left(\sum_{k=0}^{n} \binom{a+1}{k} z^k - \binom{a}{n} z^n\right)$$

Ora, facciamo tendere $n \to +\infty$. Se $a > -1$, la serie di destra converge assolutamente; inoltre si è già osservato che il secondo termine è infinitesimo. Dunque se $-1 < a \leq 0$ anche la somma di sinistra converge, sebbene non assolutamente.
L'unico caso rimasto da considerare è il caso $-1 < a \leq 0$, con $z = -1$. La serie diventa una serie a segni alterni del tipo $\sum_{k=0}^{+\infty} \binom{a}{k}(-1)^k$, che converge per riterio di Leibniz dato che abbiamo già osservato che il coefficiente binomiale in modulo è decrescente e infinitesimo per $k \to +\infty$.
Ricapitolando:

- Se $a \in \mathbb{N}$ la serie è una somma finita;
- Se $|z| > 1$ la serie non converge;
- Se $|z| < 1$ la serie converge assolutamente;
- Se $|z| = 1$ e $a > 0$ la serie converge assolutamente;

- Se $|z| = 1$ e $-1 < a \leq 0$ la serie converge ma non assolutamente;
- Se $|z| = 1$ e $a < -1$ la serie non converge.

Quanto visto finora acquista senso se si considera che la serie è lo sviluppo di Taylor della funzione (analitica)

$$f(z) = (1+z)^{\alpha} = \sum_{k=0}^{+\infty} \binom{a}{k} z^k$$

Come si verifica facilmente per induzione.

16.4 (*)$\sum_{k\in\mathbb{N}} kc^k = \frac{c}{(c-1)^2}$

Dimostrazione Si assume ovviamente che $|c| < 1$, altrimenti la serie divergerebbe. Consideriamo la serie

$$\left(\sum_{k=0}^{+\infty} c^k\right)^2 = \left(\sum_{k=0}^{+\infty} c^k\right)\left(\sum_{h=0}^{+\infty} c^h\right)$$

Sviluppando come prodotto di Cauchy e raccogliendo i termini con lo stesso esponente si ha che

$$\left(\sum_{k=0}^{+\infty} c^k\right)^2 = \sum_{n\in\mathbb{N}}\left(\sum_{k+h=n} c^k c^h\right) = \sum_{n=0}^{+\infty}(n+1)c^n$$

Ora, si scinde la RHS in due parti

$$\left(\sum_{k=0}^{+\infty} c^k\right)^2 = \sum_{n=0}^{+\infty} nc^n + \sum_{n=0}^{+\infty} c^n$$

Portando a sinistra le due serie geometriche si ha la tesi

$$\sum_{n=0}^{+\infty} nc^n = \left(\sum_{n=0}^{+\infty} c^n\right)^2 - \sum_{n=0}^{+\infty} c^n = \frac{1}{(1+c)^2} - \frac{1}{1-c} = \frac{c}{(1-c)^2}$$

□

16.5 (*)Generalizzazione del caso precedente

Si vuole trovare uno sviluppo in serie per $1/(1+c)^n$, con $|c| < 1$ e $n \in \mathbb{N}$. Sviluppando come prodotto di serie geometriche e applicando la proprietà associativa, si ha

$$\frac{1}{(1-c)^n} = \left(\sum_{k=0}^{+\infty} c^k\right)^n = \sum_{(k_1,k_2\ldots k_n)\in\mathbb{N}^n} c^{k_1+k_2+\ldots+k_n}$$

Si vogliono raccogliere i termini con lo stesso esponente, ossia ci si chiede in quanti modi sia possibile scrivere un intero M come somma di n interi. La risposta sono le combinazioni con ripetizione di $n+k-1$ oggetti in k posti (vedi 2.2). Dunque

$$\frac{1}{(1-c)^n} = \sum_{k\in\mathbb{N}} \binom{n+k-1}{k} c^k$$

□

16.6 (*)Serie di Mengoli generalizzata

$$\sum_{k=1}^{+\infty} \frac{1}{\prod_{i=1}^{r}(k+i)} = \frac{1}{r\prod_{k=1}^{r} k} \tag{16}$$

Dimostrazione Applicando la partial fraction decomposition, il termine generale diventa

$$\frac{1}{\prod_{i=1}^{r}(k+i)} = \frac{A}{k(k+1)\dots(k+r-1)} + \frac{B}{(k+1)\dots(k+r)} = \frac{A(k+r)+Bk}{\prod_{i=1}^{r}(k+i)}$$

Da ciò segue $A = 1/r$ e $B = -1/r$. Si riconosce allora una serie telescopica, perché ciascun termine si semplifica col successivo e resta solo il termine iniziale. Dunque

$$\sum_{k=1}^{+\infty} \frac{1}{\prod_{i=1}^{r}(k+i)} = \frac{1}{r}\left(\sum_{k=1}^{+\infty} \frac{1}{k\dots(k+r-1)} - \frac{1}{(k+1)\dots(k+r)}\right) = \frac{1}{r\prod_{i=1}^{r} k}$$

□

16.7 Prodotto di $(1 + a_i)$

Data una famiglia di reali positivi $(a_i)_{i\in I}$, vale la seguente formula

$$\prod_{i\in I}(1+a_i) = \sum_{J\in P_F(I)} \prod_{i\in J} a_i \tag{17}$$

Dove $P_F(I)$ indica le parti finite.

Dimostrazione Per insiemi finiti di indici, la tesi si dimostra banalmente per induzione. Passiamo al caso in cui I è infinito.
Per definizione di prodotto di numeri maggiori di 1, si ha che

$$\prod_{i\in I}(1+a_i) = \sup\left\{\prod_{i\in J}(1+a_i), \quad J\subseteq I \text{ finito}\right\} = \sup A$$

Analogamente, per definizione di somma di quantità positive, si ha che l'altro termine dell'uguaglianza vale

$$\sum_{J\in P_F(I)} \prod_{i\in J} a_i = \sup\left\{\sum_{k=1}^{n}\prod_{i\in J_k} a_i, \quad J_k\subseteq I \text{ finito}, \quad n\in\mathbb{N}\right\} = \sup B$$

Mostreremo ora che i due sup sono uguali.

- Mostriamo che $A \subseteq B$. Infatti, dato un insieme finito di indici J, esistono un numero finito di insiemi J_1, $J_2 \dots J_n$ finiti per cui vale

  $$\prod_{i\in J}(1+a_i) = \sum_{k=1}^{n}\prod_{i\in J_k} a_i$$

  Dunque ogni elemento di A è anche un elemento di B. Ciò implica che $\sup A \leq \sup B$.

- Prendiamo ora un elemento di B e osserviamo che esso è maggiorato da $\sup A$, in quanto

  $$\sum_{k=1}^{n}\prod_{i\in J_k} a_i \leq \prod_{i\in\bigcup_{k=1}^{n} J_k}(1+a_i) \leq \prod_{i\in I}(1+a_i) = \sup A$$

  Dunque, passando al sup si ha che $\sup A \geq \sup B$.

- Per doppia disuguaglianza segue che $\sup A = \sup B$, come voluto.

□

16.8 (*) Stima del resto dell'armonica generalizzata

Vogliamo stimare la coda n-esima di $\sum_{k=0}^{+\infty} \frac{1}{k^{1+\alpha}}$, con $\alpha > 0$ in modo che la serie converga. Stimiamo la serie con l'integrale, sperando che nel passaggio dal discreto al continuo i due valori non si discostino di molto (cosa che andrebbe verificata, ma abbonatemi almeno questa, *n.d.A.*)

$$\left|\sum_{k=n}^{+\infty} \frac{1}{k^{1+\alpha}}\right| \approx \int_n^{+\infty} \frac{dx}{x^{1+\alpha}} = \frac{1}{\alpha n^\alpha} \gg \frac{1}{n^{\alpha+1}}$$

Ciò mostra, in particolare, che l'armonica generalizzata non converge in modo geometrico, in quanto la serie geometrica ha un resto dello stesso ordine dell'ultimo termine.

17 Serie generatrici e applicazioni in Combinatoria

17.1 Serie generatrice di una successione per ricorrenza

Le seguenti condizioni sono equivalenti:

1. La successione $F(n)$ risolve un'equazione alle differenze finite lineare, omogenea e a coefficienti costanti, cioé
$$F(n+m) = \sum_{k=0}^{m-1} a_k F(n+k)$$
2. $F(n)$ sono i coefficienti dello sviluppo in serie di una funzione razionale propria $p(x)/q(x)$ in $x = 0$, in un intorno sufficientemente piccolo;
3. $F(n)$ è una combinazione lineare di successioni del tipo $n^j \alpha^n$, con α complesso e j naturale (eventualmente nullo).

$1 \iff 3$ La coimplicazione coincide con quanto mostrato nella sezione 22.6, dunque si omettono ulteriori precisazioni.

$1, 3 \implies 2$ Consideriamo la funzione

$$\phi(x) = \sum_{k=0}^{+\infty} F(k)x^k = \sum_{k=0}^{+\infty} x^k \left(\sum k^j \alpha^k\right)$$

La funzione è una serie di potenze con raggio di convergenza non nullo, come si può verificare con il criterio della radice. Inoltre, per costruzione, gli $F(k)$ coincidono con i coefficienti dello sviluppo di Taylor di ϕ in $x = 0$.
Ora, dato che i coefficienti seguono una regola ricorsiva, anche la funzione si comporta allo stesso modo, ossia $\phi(x)$ è identificata dalla seguente equazione funzionale:

$$\phi(x) = \sum_{k=0}^{+\infty} F(k)x^k = F(0) + xF(1) + \ldots + x^m F(m) + \sum_{n=0}^{+\infty} x^{n+m+1} \sum_{k=0}^{m} a_k F(n+k)$$

$$\phi(x) = F(0)+xF(1)+\ldots+x^m F(m)+\sum_{k=0}^{m} a_k x^k \sum_{k=0}^{+\infty} F(k)x^k = F(0)+xF(1)+\ldots+x^m F(m)+\sum_{k=0}^{m} a_k x^k \phi(x)$$

Si è ottenuta un'uguaglianza tra un polinomio che moltiplica $\phi(x)$ e un altro polinomio, e dunque ϕ è una funzione razionale.

$2 \implies 1$ Consideriamo una funzione razionale

$$\phi(x) = \frac{p(x)}{q(x)} \quad \text{con } q(0) \neq 0$$

Una funzione razionale è analitica nel suo dominio, cosa per cui rimando a 5.7. Dunque esiste un intorno del punto $x = 0$ in cui la funzione coincide con il suo polinomio di Taylor.

$$\phi(x) = \frac{p(x)}{q(x)} = \sum_{k=0}^{+\infty} F(k)x^k$$

Moltiplichiamo per $q(x) = \sum_{j=0}^{r} a_j x^j$ ambo i membri. Otteniamo

$$p(x) = q(x) \sum_{k=0}^{+\infty} F(k)x^k = \sum_{j=0}^{r} a_j x^j \sum_{k=0}^{+\infty} F(k)x^k$$

Eseguendo il prodotto di Cauchy delle due sommatorie otteniamo

$$p(x) = \sum_{k=0}^{+\infty} x^k \sum_{j=0}^{\min r,k} a_j F(k-j)$$

Alla LHS abbiamo un polinomio di un certo grado m. Allora tutti i coefficienti della serie oltre il grado m devono essere nulli, dato che l'unica serie di Taylor di un polinomio è esso stesso. Dunque, per i coefficienti di ordine $k > \max(r, m)$ vale la seguente condizione

$$\sum_{j=0}^{r} a_j F(k-j) = 0$$

Che è un'equazione lineare omogenea alle differenze finite di ordine r.

□

17.2 Cardinalità delle partizioni di $[n]$

Si vuole conoscere quante siano le relazioni di equivalenza, ossia le partizioni, di un insieme finito di n elementi. Mostreremo la seguente formula, detta formula di Dobinski:

$$PAR(n) = \frac{1}{e} \sum_{k=0}^{+\infty} \frac{k^n}{k!} \tag{18}$$

Dimostrazione Partiamo da qualcosa che sembra non c'entrare una sega. Consideriamo le derivate successive della funzione e^{e^x}. Indicheremo con $\longrightarrow$ la derivazione.

$$e^{e^x} \longrightarrow e^{e^x} e^x \longrightarrow e^{e^x} e^x + e^{e^x} e^x e^x \longrightarrow e^{e^x} e^x + e^{e^x} e^x e^x + e^{e^x} e^x e^x + e^{e^x} e^x e^x + e^{e^x} e^x e^x e^x \ldots$$

Perché tutto questo delirio? Si può osservare che il numero di termini al passo n-esimo, contati con la loro molteplicità, è pari al numero di partizioni di n elementi. Inoltre, ogni termine ha tanti fattori della forma e^x quanti sono gli insiemi non vuoti della partizione che rappresenta.
Per dimostrarlo, procediamo per induzione su n. Per il caso $n = 0$ la tesi è banalmente vera. Per il passo induttivo, osserviamo che ogni termine ha un prefisso e^{e^x} che moltiplica una catena di e^x, e deve essere derivato con la formula del prodotto. Derivare il prefisso aggiunge un e^x alla catena. Questo equivale ad aggiungere un nuovo insieme non vuoto alla partizione, che corrisponde al singoletto $\{n + 1\}$, e a tenere gli altri insiemi com'erano. Derivare gli e^x, invece, crea tante copie del termine quanti gli elementi della catena. Ciò equivale ad aggiungere l'elemento $n + 1$ a un insieme non vuoto che era già all'interno della partizione. Si osserva che i due procedimenti descritti sono tutti e soli i modi di formare delle partizioni di $n + 1$ elementi date le partizioni di n elementi, e dunque la tesi è vera.[19] Ora, indicata con D^n la derivata n-esima, abbiamo che

$$D^n e^{e^x} = e^{e^x} \sum_{P \in PAR(n)} \prod_{J \in P} e^x$$

Se calcoliamo la formula in $x = 0$, otteniamo

$$D^n e^{e^x}|_{x=0} = e\,|PAR(n)|$$

Dunque, il polinomipo di Taylor di e^{e^x} è

$$T[0, e^{e^x}](x) = \sum_{n=0}^{+\infty} e\,|PAR(n)| \frac{x^n}{n!}$$

[19]Quanto appena mostrato segue dalla cosiddetta formula di Faà di Bruno, la quale afferma che la derivata di una funzione composta vale:

$$(f \circ g)^{(n)} = \sum_{P \in PAR(n)} f^{(|P|)} \prod_{A \in P} g^{(|A|)}$$

Dove f è derivata rispetto a g. Questa formula indecente, la cui dimostrazione sostanzialmente coincide con quanto visto per e^{e^x}, è una delle pochissime cose per cui si ricorda il beato Faà di Bruno. Che modo orribile di passare alla storia.

D'altra parte, se sviluppiamo e^{e^x} come serie esponenziale, otteniamo

$$e^{e^x} = \sum_{k=0}^{+\infty} \frac{e^{kx}}{k!} = \sum_{k=0}^{+\infty}\sum_{n=0}^{+\infty} \frac{(kx)^n}{n!k!}$$

La serie è assolutamente sommabile, dato che fissato x il valore di e^{e^x} è finito. Allora è possibile invertire le sommatorie, ottenendo

$$e^{e^x} = \sum_{n=0}^{+\infty}\sum_{k=0}^{+\infty} \frac{(kx)^n}{n!k!} = \sum_{n=0}^{+\infty} \frac{x^n}{n!} \sum_{k=0}^{+\infty} \frac{k^n}{k!}$$

Comparando i due sviluppo di Taylor, tutti i coefficienti devono essere uguali. Allora

$$\sum_{n=0}^{+\infty} e\, |PAR(n)| \frac{x^n}{n!} = \sum_{n=0}^{+\infty} \frac{x^n}{n!} \sum_{k=0}^{+\infty} \frac{k^n}{k!}$$

Dall'uguaglianza termine a termine segue la tesi:

$$|PAR(n)| = \frac{1}{e} \sum_{k=0}^{+\infty} \frac{k^n}{k!}$$

□

17.3 Numero di partizioni in un numero fissato di classi

Consideriamo lo stesso problema del paragrafo precedente, ma imponendo ora che l'insieme $[n]$ risulti partizionato in k insiemi non vuoti. Definiamo allora $p(n,k)$ come il numero di partizioni di n elementi in k insiemi. Adotteremo la convenzione che

- $p(0,0) = 1$;
- Se $n \neq 0$, $p(n,0) = 0$;
- Se $k > n$, $p(n,k) = 0$.

Cerchiamo una relazione ricorsiva che ricalchi quella delle partizioni. Abbiamo che

$$\begin{cases} p(1,1) = 1 \\ p(n+1,k) = kp(n,k) + p(n,k-1) \end{cases}$$

La prima è ovvia. Per quanto riguarda la seconda, se vogliamo dividere un insieme da $n+1$ elementi in k blocchi, possiamo o suddividere i primi n elementi e dopo scegliere dove posizionare l'$n+1$-esimo, oppure isolare quest'ultimo e suddividere i restanti n in $k-1$ insiemi. Moltiplichiamo per x^k e sommiamo per $k \in \mathbb{N}$, ottenendo (dopo aver soppresso i termini superflui)

$$\sum_{k=0}^{+\infty} p(n+1,k)x^k = \sum_{k=0}^{+\infty} kp(n,k)x^k + \sum_{k=0}^{+\infty} p(n,k-1)x^k$$

Riscalando adeguatamente gli indici e mantenendo i termini oltre n (che sono nulli per convenzione, ma ci faranno comodo) otteniamo

$$\sum_{k=0}^{+\infty} p(n+1,k)x^k = x\left(\sum_{k=0}^{+\infty} kp(n,k)x^{k-1} + \sum_{k=0}^{+\infty} p(n,k)x^k\right)$$

Ora, chiamiamo $P_n(x) = \sum_{k=1}^{+\infty} p(n,k)x^k$ il polinomio i cui coefficienti sono le partizioni di n elementi. Possiamo allora scrivere, osservando che il primo termine alla RHS è la derivata del secondo, che

$$P_{n+1}(x) = x\left(P_n(x)' + P_n(x)\right)$$

Moltiplichiamo per un fattore integrante in modo da completare la derivata del prodotto. Otteniamo

$$e^x P_{n+1}(x) = x\,(e^x P_n(x)' + e^x P_n(x)) = x\,(e^x P_n(x))'$$

Per rendere la trattazione più agevole, possiamo allora studiare la successione $S_n(x) = e^x P_n(x)$, che rispetta la regola ricorsiva

$$\begin{cases} S_0(x) = e^x = \sum_{k=0}^{+\infty} \frac{x^k}{k!} \\ S_{n+1}(x) = xS_n'(x) \end{cases}$$

Siamo ora in grado di esplicitare successivamente S_n, osservando che ogni volta che si deriva e si moltiplica per x l'unico risultato è moltiplicare ogni termine per il corrispondente k. Otteniamo allora che

$$S_n(x) = \sum_{k=0}^{+\infty} \frac{k^n x^k}{k!}$$

Otteniamo infine $P_n(x) = e^{-x} S_n(x)$ dal prodotto di Cauchy delle due serie:

$$P_n(x) = e^{-x} S_n(x) = \sum_{k=0}^{n} x^k \sum_{j=0}^{k} \frac{(-1)^{n-j} j^n}{j!(n-j)!}$$

Da cui

$$p(n,k) = \sum_{j=0}^{k} \frac{(-1)^{n-j} j^n}{j!(n-j)!} = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{n-j} \binom{k}{j} j^n$$

Il lettore più attento (anche del sottoscritto) osserverà una somiglianza con il numero di funzioni suriettive da un insieme di n elementi a uno di k, di cui si è discusso in 1.5. Infatti, dette $S(n,k)$ queste ultime, vale che

$$p(n,k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{n-j} \binom{k}{j} j^n = \frac{S(n,k)}{k!}$$

Il che ha senso, in quanto per determinare una funzione surgettiva si partiziona l'insieme $[n]$ in k insiemi controimmagine, e si determina il valore che la funzione assume su ciascuno di essi; viceversa, data una funzione surgettiva, essa determina una partizione del dominio che non dipende dai valori che la funzione assume.
Si osserva che si può riottenere in modo rapido il numero di partizioni di $[n]$ dal conto fatto. Infatti, se si calcola $P_n(1)$ si ottiene esattamente la somma su k dei $p(n,k)$. Allora

$$PAR(n) = \sum_{k=0}^{n} p(n,k) = P_n(1) = e^{-1} S_n(1) = \frac{1}{e} \sum_{k=0}^{+\infty} \frac{k^n}{k!}$$

Che è la formula di Dobinski.

□

17.4 (*) Stima del resto della formula di Dobinski

Dato che la cardinalità delle partizioni è un numero intero, è sufficiente trovare da che termine N in poi la coda della serie è minore di uno. Infatti, per calcolare il numero delle partizioni basterà fermarsi a tale termine e arrotondare per eccesso.
Fissato un naturale n di cui si calcolano le partizioni, escludiamo innanzitutto i valori di N per cui il termine generale della serie è maggiore di 1, in quanto sicuramente il resto sarà minore di 1 solo dopo tutti i termini maggiori di 1. Imponiamo allora

$$\frac{N^n}{N!} < 1$$

Per la formula di Stirling, sappiamo che

$$\frac{N^N}{e^N} \leq N!$$

Imponiamo allora
$$N^n \leq \frac{N^N}{e^N} \leq N!$$
Passando ai logaritmi a destra, otteniamo
$$n \log N \leq N \log N - N \leq N!$$
Dividendo per $\log N$ otteniamo
$$n \leq N\left(1 - \frac{1}{\log N}\right) \leq N$$
Abbiamo dunque $N \geq n$. Ora, prendiamo in considerazione il resto N-esimo della serie:
$$R_N = \frac{1}{e}\sum_{k=N+1}^{+\infty} \frac{k^n}{k!} = \frac{1}{e}\sum_{j=0}^{+\infty} \frac{(N+j)^n}{(N+j)!} \leq \frac{N^n}{N!e}\sum_{j=0}^{+\infty}\frac{1}{j!}$$
L'ultima disuguaglianza segue dal fatto che, comunque si scelga $j \geq 0$, se $n \leq N$ allora
$$\frac{(N+j)^n}{(N+j)!} \leq \frac{N^n}{N!\,j!} \quad \longrightarrow \quad \left(\frac{N+j}{N}\right)^n \leq \binom{N+j}{N}$$
Che è vero perché
$$\left(\frac{N+j}{N}\right)^n = \left(1+\frac{j}{N}\right)^n \leq \left(1+\frac{j}{N}\right)^N \leq \frac{j+1}{1}\frac{j+2}{2}\cdots\frac{j+N}{N} = \binom{N+j}{N}$$
Tornando al resto, la serie alla fine è esattamente la serie esponenziale di $\exp(1) = e$. Allora possiamo semplificare e e ottenere
$$R_N = \frac{N^n}{N!e}\sum_{j=0}^{+\infty}\frac{1}{j!} = \frac{N^n}{N!}$$
Dato che abbiamo già imposto che il termine alla RHS sia minore di 1, abbiamo concluso. Dunque, basta calcolare un numero $N \geq n$ di termini della serie.

□

17.5 Problema del cambio della moneta

Supponiamo di avere una quantità illimitata di banconote di r taglie intere distinte e coprime, siano $p_1, \ldots, p_r$. Fissato un intero $n \gg 1$, in quanti modi è possibile ottenerlo come somma di un certo numero di banconote di quella taglia?
Formalmente, sia $\mathbf{p} = (p_1, \ldots, p_r) \in (\mathbb{N}*)^r$, dove $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, tale che $MCD(p_1, \ldots, p_r) = 1$. Detto $\mathbf{p}\cdot\mathbf{k}$ il prodotto scalare canonico, si vuole stimare asintoticamente
$$a(n) = \#\left\{\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{N}^r \text{ t.c. } \mathbf{p}\cdot\mathbf{k} = \sum_{i=1}^{r} k_i p_i = n\right\}$$
Consideriamo la serie generatrice $f(x) = \sum_{n\in\mathbb{N}} a(n)x^n$.
$$\sum_{n\in\mathbb{N}} a(n)x^n = \sum_{n\in\mathbb{N}} \sum_{\mathbf{k}\in\mathbb{N}^r}^{\mathbf{p}\cdot\mathbf{k}=n} x^{\mathbf{p}\cdot\mathbf{k}} = \sum_{\mathbf{k}\in\mathbb{N}^r} x^{\mathbf{p}\cdot\mathbf{k}}$$
Ora, supponiamo che sia $x \in [0,1)$ e troviamo uno sviluppo per la sommatoria, che faccia uso delle proprietà di commutatività delle somme per reali non negativi, di cui a 15.1. Dopodiché potremo estendere il risultato a $x \in \mathbb{C}$, con $|x| < 1$, essendo la sommatoria assolutamente sommabile.
$$\sum_{n\in\mathbb{N}} a(n)x^n = \sum_{\mathbf{k}\in\mathbb{N}^r} x^{\mathbf{p}\cdot\mathbf{k}} = \prod_{i=1}^{r}\sum_{k_i\in\mathbb{N}} x^{p_i k_i} = \prod_{i=1}^{r}\frac{1}{1-x^{p_i}} \tag{19}$$

Ora, sviluppiamo il denominatore, composto da polinomi le cui radici sono le radici p_i-esime dell'unità. L'unico fattore comune a tutti è $(1-x)$, dato che gli esponenti sono coprimi. Risulta allora che, dette ω_j le radici dell'unità,

$$\prod_{i=1}^{r}(1-x^{p_i}) = (1-x)^r \prod_{j=1}^{N}(\omega_j - x)^{r_j}$$

per opportuni coefficienti r_j. Si osserva che $r_j < r$, altrimenti ω_j sarebbe una radice comune. Inoltre, dato che il prodotto delle radici dello stesso ordine fa 1 (escludendo eventualmente $\omega = -1$), possiamo dividere per $\prod_{j=1}^{N}(\omega_j)^{r_j}$, ottenendo, a meno di un segno

$$\prod_{i=1}^{r}(1-x^{p_i}) = \pm(1-x)^r \prod_{j=1}^{N}(1-\frac{x}{\omega_j})^{r_j}$$

Sostituiamo il denominatore nell'equazione 19. La RHS ammette uno sviluppo in frazioni semplici del tipo

$$\prod_{i=1}^{r}\frac{1}{1-x^{p_i}} = \frac{a_r}{(1-x)^r} + \sum_{h=1}^{r-1}\frac{a_h}{(1-x)^h} + \sum_{j=1\dots N}^{k=1\dots r_j}\frac{c_{jk}}{(1-\frac{x}{\omega_j})^k} \tag{20}$$

Per certe costanti c_{jk} e a_h. Il segno che eventualmente ci siamo persi nel passaggio precedente viene inglobato nei c_{jk}, dunque non ce ne dobbiamo più preoccupare. Possiamo ulteriormente sviluppare la prima frazione in accordo con quanto mostrato in 16.5:

$$\prod_{i=1}^{r}\frac{1}{1-x^{p_i}} = a_r \sum_{m=0}^{+\infty}\binom{-r}{m}(-x)^n + \sum_{h=1}^{r-1}\frac{a_h}{(1-x)^h} + \sum_{j=1\dots N}^{k=1\dots r_j}\frac{c_{jk}}{(1-\frac{x}{\omega_j})^k}$$

con la notazione

$$\binom{-r}{m} = \frac{(-r)(-r-1)\dots(-r-m+1)}{m!} = (-1)^m\binom{r+m-1}{m} = (-1)^m\frac{(m+r-1)!}{(r-1)!m!}$$

Per quanto visto in 20.7, se m è sufficientemente grande possiamo approssimare il binomiale come

$$\binom{-r}{m} \approx (-1)^m\frac{m!m^{r-1}}{m!(r-1)!}(1+o(1)) = (-1)^m\frac{m^{r-1}}{(r-1)!}(1+o(1))$$

Riassumendo, abbiamo

$$\sum_{n\in\mathbb{N}} a(n)x^n = \prod_{i=1}^{r}\frac{1}{1-x^{p_i}} \approx \sum_{m=0}^{+\infty}\frac{a_r\, m^{r-1}}{(r-1)!}x^m + \sum_{h=1}^{r-1}\frac{a_h}{(1-x)^h}\sum_{j=1\dots N}^{k=1\dots r_j}\frac{c_{jk}}{(1-\frac{x}{\omega_j})^k} \tag{21}$$

Il secondo e il terzo termine si possono trascurare per $m \gg 1$. Infatti, sviluppando in modulo ciascuna frazione si otterranno dei coefficienti binomiali, esattamente come sopra, ma del tipo $\binom{r_j}{m}$ oppure $\binom{h}{m}$, con h, $r_j < r$. Reiterando le considerazioni precedenti, otterremo che questi dipendono da m con una potenza trascurabile rispetto a m^r. Allora otteniamo che

$$\sum_{n\in\mathbb{N}} a(n)x^n \approx \sum_{m=0}^{+\infty}\frac{a_r m^{r-1}}{(r-1)!}x^m$$

Da cui

$$a(m) \approx \frac{a_r m^{r-1}}{(r-1)!}$$

Resta da determinare a_r. Moltiplichiamo la prima parte dell'equazione 20 per $(1-x)^r$ e facciamo tendere $x \to 1$. Osserviamo che i termini della seconda e della terza sommatoria sono infinitesimi per $x \to 1$. Allora

$$a_r = \frac{1}{\prod_{i=1}^{r}\frac{1-x^{p_i}}{1-x}} + o(1)$$

Passando al limite otteniamo

$$a_r = \frac{1}{\prod_{i=1}^r p_i}$$

Dunque

$$a(m) \approx \frac{m^{r-1}}{(r-1)! \prod_{i=1}^r p_i}$$

Che è la cosiddetta stima di Schur sui coefficienti del problema di cambio della moneta.
Ora, sia $A(m) = \sum_{i=1}^m a(i)$, cioè il numero di modi di sommare delle banconote per ottenere una cifra minore o uguale a m. In altri termini

$$A(n) = \# \left\{ \mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{N}^r \text{ t.c. } \mathbf{p} \cdot \mathbf{k} = \sum_{i=1}^r k_i p_i \leq n \right\} \tag{22}$$

Osserviamo che

$$(1-x) \sum A(m)x^m = \sum (A(m) - A(m-1))x^m = \sum a(m)x^m$$

Da questa verità generale si ottiene che

$$\sum A(m)x^m = \frac{\sum a(m)x^m}{1-x}$$

Ripetendo esattamente lo stesso conto di prima, ma considerando che al denominatore stavolta c'è $(1-x)^{r+1}$ si ottiene

$$A(m) \approx \frac{m^r}{r! \prod_{i=1}^r p_i}$$

Se ci bastasse una stima un po' più rozza, potremmo osservare che, nello spazio $(\mathbb{N}^*)^r$, l'equazione tra parentesi in 22 è quella di un semispazio r-dimensionale, di cui si vogliono contare i punti a coordinate naturali. Una stima di $A(n)$ può allora essere il volume del solido (detto simplesso) delimitato da tale piano e dai piani coordinati. Tale simplesso interseca ciascun asse x_i,in cui tutte le altre coordinate sono nulle, nel punto n/p_i. Si può allora dimostrare, con il principio di cavalieri, che il volume del solido vale

$$V = \frac{1}{r!} \frac{m}{p_1} \cdots \frac{m}{p_r} = \frac{m^r}{r! \prod_{i=1}^r p_i}$$

Si rimanda a 8.2 per la dimostrazione.

17.6 Numeri di Catalan

Consideriamo un poligono regolare di n lati e chiediamoci in quanti modi è possibile triangolarlo, ossia scomporlo in somma di triangoli i cui vertici sono vertici del poligono. Sia $C(n)$ il numero di modi di scomporre un poligono di $n+2$ lati. Osserviamo che

$$\begin{cases} C(0) = 1 \\ C(n+1) = \sum_{k=0}^n C(k)C(n-k) \end{cases}$$

La prima relazione dice che un segmento è già un triangolo, sebbene sia degenere. La seconda segue da questa costruzione: fissiamo un lato AB su un poligono di $n+3$ lati e facciamo variare il terzo vertice del triangolo di base AB tra gli $n+1$ restanti, che numeriamo da 0 a n. Scegliere il vertice k-esimo significa suddividere il poligono in un triangolo, un poligono da $k+1$ vertici (che sono A, k e i $k-1$ prima) e un poligono da $n-k+1$ (rispettivamente B, k e i vertici fino a n). Dato che questi ultimi si triangolano in $C(k)$ e $C(n-k)$ modi in maniera indipendente, segue la tesi.
Ora, consideriamo la serie generatrice

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

Per la relazione ricorsiva si ha che

$$f(x) = C(0) + \sum_{n=0}^{+\infty} C(n+1)x^{n+1} = 1 + x\sum_{n=0}^{+\infty} x^n \sum_{k=0}^{n} C(k)C(n-k)$$

Riconosciamo nell'ultima serie il prodotto di Cauchy della funzione con se stessa. Dunque otteniamo la seguente equazione

$$f(x) = 1 + xf^2(x)$$

Da cui

$$f(x) = \frac{1 \pm \sqrt{1-4x}}{2x}$$

L'ambiguità nel segno si risolve imponendo che in $x = 0$ valga $f(0) = C(0) = 1$. Dunque deve esistere il limite per $x \to 0$, da cui il numeratore deve essere infinitesimo. Otteniamo allora che

$$f(x) = \frac{1 - \sqrt{1-4x}}{2x}$$

Ora, sviluppiamo in serie la funzione ottenuta. Dal fatto che $\sqrt{1-4x}$è la serie binomiale di $\alpha = 1/2$[20] calcolata in $-4x$ otteniamo

$$f(x) = \frac{1-\sqrt{1-4x}}{2x} = \frac{1 - \sum_{k=0}^{+\infty}\binom{1/2}{k}(-4x)^k}{2x} = \sum_{k=1}^{+\infty}\binom{1/2}{k}\frac{4^k}{2}(-x)^{k-1} = \sum_{k=0}^{+\infty}\binom{1/2}{k+1}\frac{4^{k+1}}{2}(-x)^k$$

Da cui, comparando i coefficienti dello sviluppo,

$$C(k) = \binom{1/2}{k+1}\frac{(-1)^k 4^{k+1}}{2} = \frac{(1/2)\dots(1/2-k)}{(k+1)!}\frac{(-1)^k 4^{k+1}}{2} = \frac{(2k-1)\dots(1)}{(k+1)!}2^k =$$

$$= \frac{(2k-1)!!}{(k+1)!}2^k = \frac{(2k)!}{(k+1)!k!}$$

Da cui otteniamo che il k-esimo numero di Catalan vale

$$C(k) = \frac{1}{k+1}\binom{2k}{k}$$

[20] Vedasi 16.3.

18 Integrale di Riemann-Darboux

La sagra dell'estremo superiore.

Memorie dall'infanzia rurale di Pietro Majer.

Suddivisione alla Riemann Dato un intervallo chiuso e limitato $[a, b]$, una sua suddivisione, o partizione, p è un insieme finito di punti $a = x_0 < x_1 < \ldots < x_n = b$. L'intervallo $[x_{k-1}, x_k]$ verrà indicato con I_k e la sua ampiezza con $|I_k|$. La massima ampiezza di un intervallo di una partizione si dice parametro di finezza della partizione, e si indicherà con $|p|$. Sia P l'insieme delle suddivisioni.

Somma integrale

Piccole somme crescono

Memorie dall'infanzia rurale di Pietro Majer.

Data una funzione $f : [a, b] \longrightarrow \mathbb{R}$ e una partizione p di $[a, b]$, si definisce somma integrale inferiore su p la quantità

$$s(f, p) = \sum_{I_k \in p} |I_k| \inf_{x \in I_k} f(x)$$

Analogamente, si definisce la somma integrale superiore su p come

$$S(f, p) = \sum_{I_k \in p} |I_k| \sup_{x \in I_k} f(x)$$

Ci restringiamo fin da subito al caso in cui la funzione in considerazione è limitata e definita in un compatto, ed estenderemo poi la nozione a tutte le funzioni.
Per le somme integrali valgono le seguenti proprietà, di immediata verifica:

1. $s(f, p) = -S(-f, p)$. Questa proprietà segue dal fatto che l'inf di un insieme è definito come il sup dell'insieme degli opposti. In virtù di tale lemmino è possibile enunciare le altre proprietà solo per le somme inferiori per poi trasferirle alle somme superiori.

2. Se $p \subseteq q$, $s(f, p) \leq s(f, q)$ É sufficiente mostrare questa proprietà nel caso in cui si aggiunge alla partizione minimale $\{a, b\}$ un solo punto, nel qual caso risulta ovvia la dimostrazione, e procedere induttivamente sul numero di punti della partizione.

3. Analogamente, Se $p \subseteq q$, $S(f, p) \leq S(f, q)$

4. $\forall p, q$, $s(f, p) \leq S(f, q)$.
   Basta osservare che
   $$s(f, p) \leq s(f, p \cup q) \leq S(f, p \cup q) \leq S(f, q)$$

Integrale Si definisce integrale inferiore la quantità

$$_*\int_a^b f(x)\, dx = \sup_{p \in P} s(f, p)$$

E analogamente l'integrale superiore come

$$^*\int_a^b f(x)\, dx = \inf_{p \in P} S(f, p)$$

Le proprietà appena viste per le somme integrali passano al limite. Inoltre, sono rilevanti le seguenti tre proprietà dell'integrale superiore, inteso come una funzione dallo spazio delle funzioni su $[a, b]$ in $\mathbb{R}$:

- Omogeneità: $\forall \lambda \in \mathbb{R}$, ${}^*\int_a^b \lambda f(x)\,dx = \lambda {}^*\int_a^b f(x)\,dx$;
- Additività sugli intervalli: ${}^*\int_a^b f\,dx + {}^*\int_b^c f\,dx = {}^*\int_a^c f\,dx$
Infatti, la somma degli integrali equivale a fare l'inf delle partizioni in cui compare il punto b. Allora, da un lato, visto che si sta facendo l'inf di un sottoinsieme,
$$\sup_{p\in P,\, c\in p} S(f,p) = {}^*\int_a^b f\,dx + {}^*\int_b^c f\,dx \geq {}^*\int_a^c f\,dx = \sup_{p\in P} S(f,p)$$
Dall'altro, presa comunque una partizione $q \in P$, si ha che $S(f,q) \geq S(f, g \cup \{c\})$, da cui
$$\sup_{p\in P,\, c\in p} S(f,p) \leq \sup_{p\in P} S(f,p)$$
Dalle due disuguaglianze segue la tesi.
- Subadditività sull'argomento: ${}^*\int_a^b (f+g)\,dx \leq {}^*\int_a^b f\,dx + {}^*\int_a^b g\,dx$
Per mostrarlo, ricordiamo che, per una nota proprietà dell'estremo superiore, vale che
$$\sup_{x\in[a,b]} f(x) + g(x) \leq \sup_{x\in[a,b]} f(x) + \sup_{x\in[a,b]} g(x)$$
Da cui
$${}^*\int_a^b (f+g)\,dx \leq S(f+g, p\cup q) \leq S(f, p\cup q) + S(g, p\cup q) \leq S(f,p) + S(g,q)$$
Passando all'inf prima su p e poi su q otteniamo la disuguaglianza voluta.

□

Analogamente, l'integrale inferiore è omogeneo, additivo sugli intervalli e superadditivo sull'argomento, cioè ${}_*\int_a^b (f+g)\,dx \geq {}_*\int_a^b f\,dx + {}_*\int_a^b g\,dx$

18.1 Integrabilità e definizioni equivalenti

Data una funzione limitata f, le seguenti tre condizioni sono equivalenti:

1. L'integrale superiore e l'integrale inferiore coincidono;
2. detta $\rho(p) = S(f,p) - s(f,p)$ la differenza tra le somme sulla stessa partizione, che chiameremo lo scarto di f su p, vale che $\inf_{p\in P} \rho(p) = 0$. In altri termini, $\forall\varepsilon$ esiste una partizione $p \in P$ tale che $\rho(p) < \varepsilon$;
3. $\lim_{|p|\to 0} \rho(p) = 0$. In altri termini, $\forall\varepsilon\ \exists\delta$ tale che per ogni partizione di finezza minore di δ si ha $\rho(p) < \varepsilon$.

Se vale una (e dunque tutte) di queste proprietà, la funzione si dice integrabile secondo Riemann-Darboux. La seconda condizione venne formulata da Darboux pochi anni dopo la pubblicazione dei lavori di Riemann, il quale aveva invece dato la terza.
Si osserva che $\rho(p) = S(f,p) - s(f,p) = \sum_{I_k\in p} |I_k| \sup_{x,y\in I_k} (f(x) - f(y))$. Si definisce oscillazione di f sull'intervallo I_k la quantità $osc(f, I_k) = \sup_{x,y\in I_k} (f(x) - f(y))$.

Dimostrazione

- ($1 \iff 2$) Se $\inf_{p\in P} \rho(p) = 0$, le somme superiori e inferiori hanno lo stesso limite. Viceversa, se la funzione è integrabile, allora ci sono due partizioni q e r tali che $S(f,q) - s(f,r) < \varepsilon$. Adesso, posto $p = q \cup r$, otteniamo che $S(f,p) - s(f,p) \leq S(f,q) - s(f,r) < \varepsilon$.
- ($3 \Longrightarrow 2$) Ovvio dato che si stanno esibendo delle partizioni tali che $\inf_{p\in P} \rho(p) = 0$.

- ($3 \Longleftarrow 2$) Sia $\varepsilon > 0$ e sia $p_\varepsilon \in P$ tale che $\rho(p_\varepsilon) < \varepsilon$. Definiamo
$$\delta = \min\left\{ |I_{k_\varepsilon}|, \frac{\varepsilon}{N_\varepsilon osc(f, [a,b])} \right\}$$
Dove I_{k_ε} sono gli intervalli di p_ε e N_ε è il loro numero. Consideriamo ora una partizione q i cui intervalli J_k hanno finezza massima $|q| < \delta$. Alcuni degli intervalli di q conterranno un punto di p_ε, e tra l'altro uno solo essendo la densità di q minore di δ. Consideriamo allora $\rho(q)$ e spezziamola in due parti:
$$\rho(q) = \sum_{J_k \cap p_\varepsilon \neq \emptyset} |J_k| osc(f, J_k) + \sum_{J_k \cap p_\varepsilon = \emptyset} |J_k| osc(f, J_k)$$
Il primo termine è una sommatoria di esattamente N_ε elementi, e può essere stimato con $N_\varepsilon |q| osc(f, [a,b]) \leq N_\varepsilon \delta osc(f, [a,b])$. Per quanto riguarda il secondo, esso è una parte di una somma $\rho(p)$ a cui sono stati tolti degli intorni dei punti di p_ε. Allora è possibile completarla aggiungendo questi ultimi e dunque maggiorando il secondo termine con $\rho(q \cup p_\varepsilon)$. A sua volta, in quanto la precisione non può che diminuire se si tolgono punti della partizione, avremo che $\rho(q \cup p_\varepsilon) \leq \rho(p_\varepsilon)$. Mettendo insieme le cose, otteniamo
$$\rho(q) \leq N_\varepsilon \delta osc(f, [a,b]) + \rho(p_\varepsilon) < 2\varepsilon$$
Per come si sono scelti δ e p_ε.

□

Conseguenza della terza affermazione è che se una funzione è integrabile allora il suo integrale si può calcolare restringendosi alle partizioni di punti equispaziati (cosa che legittima il dx dentro l'integrale), o in generale con qualsiasi insieme di partizioni denso in $[a,b]$.

18.2 Proprietà dell'integrale

Osserviamo che, unendo le proprietà di subadditività dell'integrale superiore e la corrispondente superadditività dell'integrale inferiore, troviamo che l'integrale è additivo, ossia la somma di funzioni integrabili è integrabile. Si può dire di più: l'integrale è un operatore lineare, ossia
$$\int_a^b (\alpha f + \beta g)\, dx = \alpha \int_a^b f\, dx + \beta \int_a^b g\, dx$$
Sia allora $R([a,b])$ l'insieme delle funzioni Riemann-integrabili su $[a,b]$. Esso è uno spazio vettoriale su $\mathbb{R}$, e l'integrale ne è un funzionale in quanto applicazione lineare dallo spazio al campo.
Anche la proprietà di additività sugli intervalli passa all'integrale. In particolare, se si definisce l'inegrale inverso come $\int_b^a f\, dx = -\int_a^b f\, dx$, otteniamo la seguente formuletta ciclica molto carina:
$$\forall a, b, c \in \mathbb{R} \qquad \int_a^b f\, dx + \int_b^c f\, dx + \int_c^a f\, dx = 0$$

18.3 Integrabilità delle funzioni continue e/o monotone

Teorema Una funzione $f : [a,b] \longrightarrow \mathbb{R}$ limitata e continua è Riemann-integrabile.

Dimostrazione Mostriamo che $\inf_{p \in P} \rho(p) = 0$. Dato che f è continua e definita su un compatto, è uniformemente continua[21] e limitata[22]. Si può allora stimare $\rho(p)$ con il modulo di continuità della funzione:
$$\rho(p) \leq \sum_{I_k \in p} |I_k| \omega(|I_k|) \leq \omega(|p|) \sum_{I_k \in p} |I_k| = \omega(|p|)(b-a)$$
Dato che il modulo di continuità è continuo in zero, basta scegliere partizioni di finezza piccola a piacere per rendere arbitrariamente piccolo $\rho(p)$.

□

[21] Per teorema di Heine-Cantor, 7.10
[22] Per teorema di Weierstrass, 7.9

Teorema Una funzione limitata e monotona è Riemann-integrabile.

Dimostrazione Senza perdita di generalità, supponiamo che sia debolmente crescente. Allora, stimando ogni intervallo con la finezza della partizione, otteniamo

$$\rho(p) = \sum_{I_k \in p} |I_k|(f(x_k) - f(x_{k-1}) \leq |p| \sum_{I_k \in p} (f(x_{k-1}) - f(x_k) = |p|(f(b) - f(a))$$

Che è una quantità piccola a piacere.

□

18.4 Integrabilità della funzione composta

Sia $f : [a,b] \longrightarrow [c,d]$ una funzione Riemann-integrabile e limitata, e sia $\phi : [c,d] \longrightarrow \mathbb{R}$ una funzione continua. Allora la composizione $\phi \circ f$ è Riemann-integrabile.

Dimostrazione Innanzitutto, ϕ è continua su un compatto, e dunque è uniformemente continua e limitata. Perciò ammette un modulo di continuità finito $\omega(t) = \sup\{\phi(x) - \phi(y) \,|\, d(x,y) \leq t\}$. Ora, cerchiamo di rendere infinitesima $\rho(\phi(f), p)$, con $p \in P$.

$$\rho(\phi(f), p) = \sum_{I_k \in p} |I_k| \sup_{x,y \in I_k} (\phi(f(x)) - \phi(f(y))) \leq \sum_{I_k \in p} |I_k|\omega(\sup_{x,y \in I_k} |x - y|) = \sum_{I_k \in p} |I_k|\omega(osc(f, I_k))$$

Ora, fissiamo un certo valore $\eta > 0$ reale, e scomponiamo la somma mettendo da parte gli intervalli la cui oscillazione è minore di η.

$$\rho(\phi(f), p) \leq \sum_{I_k \in p}^{osc(f,I_k) \leq \eta} |I_k|\omega(osc(f, I_k)) + \sum_{I_k \in p}^{osc(f,I_k) > \eta} |I_k|\omega(osc(f, I_k))$$

Il prrimo termine si minora ponendo $\omega(osc(f, I_k)) \leq \omega(\eta)$, dato che ω è crescente. Allora abbiamo

$$\rho(\phi(f), p) \leq \omega(\eta) \sum_{I_k \in p}^{osc(f,I_k) \leq \eta} |I_k| + \sum_{I_k \in p}^{osc(f,I_k) > \eta} |I_k|\omega(osc(f, I_k)) \leq \omega(\eta)(b-a) + \sum_{I_k \in p}^{osc(f,I_k) > \eta} |I_k|\omega(osc(f, I_k))$$

Per quanto riguarda il secondo termine, osserviamo che

$$\rho(f, p) \geq \sum_{I_k \in p}^{osc(f,I_k) > \eta} |I_k| osc(f, I_k) \geq \eta \sum_{I_k \in p}^{osc(f,I_k) > \eta} |I_k|$$

Da cui

$$\sum_{I_k \in p}^{osc(f,I_k) > \eta} |I_k| \leq \rho(f, p)/\eta$$

Sostituendo nel secondo termine, e minorando $\omega(osc(f, I_k))$ con $\omega(osc(f, [a,b]))$, che è un valore finito per costruzione, otteniamo

$$\rho(\phi(f), p) \leq \omega(\eta)(b - a) + \omega(osc(f, [a,b]))\rho(f, p)/\eta$$

Ora, la disuguaglianza vale per ogni η e per ogni partizione p. Fissato $\varepsilon > 0$, scegliamo η in modo che $\omega(\eta)(b - a) < \varepsilon/2$. Un valore del genere esiste per la continuità di ω in $t = 0$. Ora, fissato quell'η, scegliamo la partizione in modo che $\omega(osc(f, [a,b]))\rho(f, p)/\eta < \varepsilon/2$, cosa che si può fare dato che f è integrabile, ossia $\rho(f, p)$ è arbitrariamente piccola. Otteniamo allora $\rho(\phi(f), p) < \varepsilon$, da cui la tesi.

□

Osservazione Dal teorema precedente segue che, se f è integrabile, lo è anche f^2. Da questo ricaviamo, scrivendo $fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$, che anche il prodotto di due funzioni integrabili è integrabile. Ciò porta a dire che l'insieme delle funzioni Riemann-integrabili su $[a, b]$ è un'algebra, ossia è sia uno spazio vettoriale (in quanto chiuso per somma e moltiplicazione per scalari) sia un anello (in quanto chiuso per somma e prodotto).

Osservazione Se la funzione ϕ non è continua, non è detto che la composizione sia integrabile. Consideriamo ad esempio

$$\phi(x) = \chi_{\mathbb{R}^+} = \begin{cases} 0 & \text{se } x \leq 0; \\ 1 & \text{se } x > 0 \end{cases}$$

$$f(x) = \begin{cases} 0 & \text{se } x \notin \mathbb{Q} \cap [0,1]; \\ 1/q & \text{se } x = p/q \, \text{con } p, q \text{ coprimi} \end{cases}$$

Vedremo a breve che f è integrabile, e ovviamente lo è anche ϕ. Ciò nonostante, la loro composizione non è integrabile, in quanto risulta essere la funzione di Dirichlet:

$$\phi(f(x)) = \chi_{\mathbb{Q}\cap[0,1]} = \begin{cases} 0 & \text{se } x \notin \mathbb{Q} \cap [0,1]; \\ 1 & \text{se } x \in \mathbb{Q} \cap [0,1] \end{cases}$$

18.5 Passaggio al limite sotto il segno di integrale

Sia f_n una successione di funzioni limitate e Riemann-integrabili su $[a, b]$. Se la successione converge uniformemente a una funzione f, allora essa è integrabile, e il suo integrale è il limite degli integrali. In altri termini, si possono scambiare il limite uniforme e l'integrale.

$$\exists \int_a^b f(x)\, dx = \int_a^b \lim_{n\to+\infty} f_n(x)\, dx = \lim_{n\to+\infty} \int_a^b f_n(x)\, dx$$

Dimostrazione É innanzitutto evidente che anche la funzione limite sarà limitata, in quanto costretta a stare definitivamente a meno di ε dalla successione. Per quanto riguarda il suo integrale, non sappiamo ancora se esista o meno, ma possiamo considerare l'integrale superiore e inferiore, che sono sempre ben definiti. Per subadditività dell'integrale superiore abbiamo:

$$^*\!\int_a^b f\, dx \leq {}^*\!\int_a^b f_n\, dx + {}^*\!\int_a^b (f - f_n)\, dx \leq {}^*\!\int_a^b f_n\, dx + S(f - f_n,\ \{a, b\}) = {}^*\!\int_a^b f_n\, dx + (b - a) d_\infty(f, f_n)$$

Dunque

$$^*\!\int_a^b f\, dx - {}^*\!\int_a^b f_n\, dx \leq (b - a) d_\infty(f, f_n)$$

Ripetendo lo stesso ragionamento scambiando f e f_n possiamo mettere il modulo alla LHS. Osserviamo innanzitutto che, essendo f_n integrabile, si può sostituire l'integrale superiore con l'integrale inferiore. Inoltre, dato che, per convergenza uniforme, la distanza infinito tra il limite e la successione è infinitesima, abbiamo che

$$\forall \varepsilon \exists n \text{ t.c. } \left| {}^*\!\int_a^b f\, dx - \int_a^b f_n\, dx \right| < \varepsilon \quad \Longrightarrow \quad \lim_{n\to+\infty} \int_a^b f_n\, dx = {}^*\!\int_a^b f\, dx$$

Ora, reiteriamo il ragionamento sostituendo le funzioni con le loro opposte, e applicando l'identità ${}^*\!\int_a^b f = -{}_*\!\int_a^b -f$. Abbiamo allora che

$$\forall \varepsilon \exists n \text{ t.c. } \left| {}_*\!\int_a^b f\, dx - \int_a^b f_n\, dx \right| < \varepsilon \quad \Longrightarrow \quad \lim_{n\to+\infty} \int_a^b f_n\, dx = {}_*\!\int_a^b f\, dx$$

Per unicità del limite, segue che l'integrale superiore e inferiore sono uguali. Dunque la funzione limite è integrabile, e il suo integrale è il limite degli integrali.

□

Esempio La funzione

$$f(x) = \begin{cases} 1/q & \text{se } x = p/q \text{ con } p, q \text{ coprimi} \\ 0 & \text{se } x \notin \mathbb{Q} \cap [0,1] \end{cases}$$

è il limite uniforme delle funzioni

$$f_n(x) = \begin{cases} 1/q & \text{se } x = p/q \text{ con } p, q \text{ coprimi e } q \leq n; \\ 0 & \text{altrimenti} \end{cases}$$

in quanto, al passo n-esimo, tutti i punti sono a distanza al massimo $\frac{1}{n+1}$ dai rispettivi valori della funzione. Ciascuna di queste funzioni è integrabile e ha integrale nullo, in quanto la funzione differisce da $y = 0$ solo in un numero finito di punti. Allora la funzione limite è integrabile e il suo integrale è nullo.

Osservazione La convergenza puntuale non è sufficiente. Ad esempio, data una numerazione dei razionali q_n, si definisce la successione di funzioni

$$f_n(x) = \begin{cases} 1 & \text{se } x = q_k \text{ con } q \leq n; \\ 0 & \text{altrimenti} \end{cases}$$

Ciascuna di queste funzioni è integrabile, ma il limite è la funzione di Dirichlet e non lo è.

Osservazione Possono capitare cose anche più strane, ad esempio che una successione di funzioni di integrale 1 converga puntualmente alla funzione identicamente nulla. Ad esempio, si consideri la successione:

$$f_n(x) = \begin{cases} 4n^2 x & \text{se } x \leq 1/2n; \\ 4n - 4n^2 x & \text{se } 1/2n < x \leq 1/n; \\ 0 & \text{se } 1/n < x \leq 1. \end{cases}$$

Le funzioni descrivono dei triangoli di base $1/n$ e altezza $2n$ con l'asse x. Perciò ciascuna di esse ha integrale pari all'area del triangolo, che vale 1. Tuttavia, il limite della successione è la funzione identicamente nulla, perché da un lato $f_n(0) = 0 \; \forall n \in \mathbb{N}$; dall'altro ogni $x > 0$ è definitivamente maggiore di $1/n$ e dunque la sua immagine sarà definitivamente nulla.

18.6 Funzione integrale

Data una funzione $f \in R([a,b])$, si definisce la sua funzione integrale

$$F(x) = \int_a^x f(t)\, dt$$

Si osserva innanzitutto che, se f è limitata, la funzione integrale è lipschitziana, e dunque uniformemente continua, di costante $\|f\|_\infty$, in quanto

$$|F(x) - F(y)| = |\int_x^y f(t)\, dt| \leq |x - y| \|f\|_\infty$$

18.7 Teorema fondamentale del calcolo integrale

Sia f una funzione integrabile in $[a,b]$ e continua in un punto $x_0 \in [a,b]$. Allora la corrispondente funzione integrale $F(x)$ è derivabile in x_0 e la sua derivata coincide con $f(x_0)$. In altri termini

$$\frac{d}{dx} \int_a^x f(t)\, dt \restriction_{x=x_0} = f(x_0)$$

Dimostrazione Osserviamo che, $\forall x \in [a, b]$, vale che

$$F(x) - F(x_0) - f(x_0)(x - x_0) = \int_{x_0}^{x} f(t)\,dt - f(x_0)\int_{x_0}^{x} dt = \int_{x_0}^{x} (f(t) - f(x_0))$$

Mostriamo che questa quantità è $o(x - x_0)$ per $x \to x_0$. Questo permetterà di concludere che la derivata di $F(x)$ è $f(x_0)$. Ora, stimiamo l'integrale, in modulo, con la massima somma superiore:

$$\left| \int_{x_0}^{x} (f(t) - f(x_0)) \right| \leq |x - x_0| \sup_{t \in [x, x_0]} |f(t) - f(x_0)|$$

Ora, dato che f è continua per $x \to x_0$, il sup è un infinitesimo se si sceglie x abbastanza vicino. Allora

$$|F(x) - F(x_0) - f(x_0)(x - x_0)| \leq |x - x_0| \sup_{t \in [x, x_0]} |f(t) - f(x_0)| = |x - x_0| o(1) = o(x - x_0)$$

Da cui

$$F(x) = F(x_0) + f(x_0)(x - x_0) + o(x - x_0) \quad \longrightarrow \quad \frac{dF(x)}{dx} \restriction_{x=x_0} = f(x_0)$$

□

Osservazione In generale, se la funzione non è continua in x_0 il teorema non vale. Ad esempio, l'integrale di $f(x) = \operatorname{sgn}(x)$ è la funzione $F(x) = |x|$, non derivabile in $x = x_0$.

Corollario Ogni funzione continua f ammette la sua funzione integrale come primitiva (o antiderivata), ossia la funzione integrale risolve l'equazione differenziale $g' = f$. C'è di più: ogni primitiva di una funzione continua è la sua funzione integrale a meno di una costante additiva.

18.8 Variante del TFCI

Il teorema fondamentale del calcolo integrale vale anche con ipotesi leggermente più deboli.
Sia $f : [a, b] \longrightarrow \mathbb{R}$ derivabile e con derivata Riemann-integrabile su $[a, b]$. Allora vale che

$$f(x) - f(a) = \int_{a}^{x} f'(t)\,dt$$

Dimostrazione Consideriamo una partizione $p = \{y_k\}_{k=0,\ldots,n}$ di $[a, x]$. Allora la differenza tra le immagini agli estremi si telescopizza come

$$f(x) - f(a) = \sum_{k=0}^{n-1} f(y_{k+1}) - f(y_k)$$

Applicando il teorema di Lagrange all'intervallo $[y_k, y_{k+1}]$, esiste un c_k tale che

$$f(x) - f(a) = \sum_{k=0}^{n-1} f(y_{k+1}) - f(y_k) = \sum_{k=0}^{n-1} f'(c_k)(y_{k+1} - y_k)$$

Allora possiamo stringere $f(x) - f(a)$ tra la somma superiore e la somma inferiore su p della derivata, che è Riemann-integrabile. Passando al limite otteniamo allora:

$$s(f', p) \leq f(x) - f(a) \leq S(f', p) \quad \longrightarrow \quad f(x) - f(a) = \int_{a}^{x} f'(t)\,dt$$

□

18.9 Integrazione per parti

> I parti, noti per la loro abilità nel lanciare frecce a cavallo contro i loro inseguitori. Mi chiederete che c'entra. Boh.
>
> *Crasso prima della disfatta di Carre del 53 a.C.*

Dalla regola di Leibniz per la derivata del prodotto

$$(fg)' = f'g + fg'$$

Se le funzioni sono continue, derivabili e con derivata continua (o anche solo con derivata Riemann-integrabile), segue la formula di integrazione per parti, per banale applicazione del teorema fondamentale del calcolo integrale:

$$\int_a^b f'g = fg \restriction_a^b - \int_a^b fg'$$

18.10 Integrazione per sostituzione

Sia $g : [a,b] \longrightarrow \mathbb{R}$ derivabile e sia $f : Im(g) \longrightarrow \mathbb{R}$ continua. Allora

$$\int_{g(a)}^{g(b)} f(x)\,dx = \int_a^b f(g(x))g'(x)\,dx$$

Per mostrarlo osserviamo che le seguenti funzioni integrali coincidono

$$\int_{g(a)}^{g(x)} f(t)\,dt = \int_a^x f(g(t))g'(t)\,dt$$

La prima è la composizione della funzione integrale di f con g, ossia $F(g(x))$. Per regola della derivata della composizione, $\frac{d}{dx}F(g(x)) = f(g(x))g'(x)$. La seconda è invece la funzione integrale di $f(g(x))g'(x)$, e dunque deve differire dalla prima solo per una costante. Dato che in $x = a$ le due funzioni valgono entrambe zero, devono coincidere.

□

18.11 Sostituzioni notevoli

Sostituzioni di Eulero

> Ovviamente nessuno si ricorda a memoria tutte le sostituzioni. Se aprendo il frigo, un giorno, avrete l'urgenza di integrare una funzione razionale, potrete andarvi a riprendere la formula.
>
> *Da un prontuario di cucina.*

Consideriamo un integrando del tipo $R(x, \sqrt{ax^2+bx+c})$, dove $R(x)$ è una funzione razionale. Per ricondurre l'integrale a una funzione razionale $Q(x)$ si possono eseguire le seguenti sostituzioni, dipendenti dal trinomio di secondo grado sotto radice.

Prima sostituzione di Eulero Supponiamo che il coefficiente direttore a sia positivo. Imponiamo il seguente cambio di variabile:

$$\sqrt{ax^2+bx+c} = (\sqrt{a}x + t) \tag{23}$$

Elevando al quadrato e semplificando il termine di grado 2, si esplicita x come una funzione razionale in t, ossia

$$x = \frac{t^2 - c}{b - 2\sqrt{a}t}$$

Seconda sostituzione di Eulero Supponiamo che il termine noto c sia positivo. Imponiamo il seguente cambio di variabile:

$$\sqrt{ax^2 + bx + c} = xt + \sqrt{c} \tag{24}$$

Elevando al quadrato, si elide c. Allora, dividendo per x[23] e spostando tutto a destra otteniamo

$$x = \frac{-b + 2\sqrt{c}t}{a - t^2}$$

Terza sostituzione di Eulero Supponiamo che il coefficiente direttore a sia negativo. Allora necessariamente il discriminante Δ dell'equazione associata deve essere positivo o nullo, dato che altrimenti la funzione sotto radice sarebbe sempre negativa. Supponendo che le radici siano λ e μ possiamo sostituire

$$\sqrt{a(x - \lambda)(x - \mu)} = (x - \lambda)t \tag{25}$$

Da cui

$$x = \frac{a\mu - \lambda t^2}{a - t^2}$$

18.12 Caratterizzazione delle funzioni integrabili

In questo paragrafo utilizzeremo alcune nozioni basilari di teoria della misura[24] per caratterizzare le funzioni Riemann-integrabili. Iniziamo con alcune definizioni e lemmi propedeutici.

Oscillazione Si dice oscillazione di una funzione in un insieme I la differenza tra l'estremo superiore e inferiore delle immagini, ossia il diametro delle immagini.

$$osc(f, I) = \sup_{x,y \in I} f(x) - f(y) = diam(F(I))$$

Similmente, si definisce l'oscillazione in un punto come l'estremo inferiore delle oscillazioni sui suoi intorni.

$$osc(f, x) = \inf_{U \text{ intorno di } x} osc(f, U)$$

Lemma 1 Una funzione è continua in x_0 se e solo se $osc(f, x_0) = 0$. Passando alla contronominale, l'insieme dei punti di discontinuità di una funzione è definito come

$$D(f) = \{x \in [a, b] \text{ t.c. } osc(f, x) > 0\}$$

La dimostrazione è una semplice riscrittura di cosa significhi essere continua.

Lemma 2 I sottolivelli della funzione oscillazione sono aperti. In altri termini, comunque si scelga $\lambda \in \mathbb{R}$, l'insieme dei punti del dominio la cui oscillazione è inferiore a λ è un aperto.
La dimostrazione è banale, perché ogni punto avente oscillazione inferiore ammette un intorno in cui l'oscillazione è minore di λ. Ma allora tutti i punti dell'intorno in questione hanno oscillazione minore di λ, e dunque stanno nel sottolivello.

[23] In teoria sarebbe illecito dividere per qualcosa di cui non sappiamo le condizioni di esistenza. Tuttavia l'analisi risponde con sufficienza che se si sottrae un solo punto al dominio della funzione essa rimane integrabile con lo stesso integrale. E poi, di fatto, se guardiamo la sostituzione dimenticandoci di come l'abbiamo ottenuta osserviamo che è possibile riottenere, per un adeguato t, il valore $x = 0$. L'osservazione si propaga anche ala terza formula di Eulero.

[24] Si rimanda a 9.1.

Lemma 3 Se $\forall x \in [a,b]$ $osc(f,x) < \lambda$, allora esiste una suddivisione dell'intervallo tale che $osc(f, I_k) < \lambda$ per ogni $k \leq n$.
Basta ricoprire il dominio con intorni di centro x su cui la funzione ha oscillazione minore di λ, estrarne un numero finito per compattezza e determinare una partizione prendendo dei punti nelle reciproche intersezioni.

Teorema di Vitali-Lebesgue Una funzione $f : [a,b] \longrightarrow \mathbb{R}$ è Riemann-integrabile se e solo se è limitata e l'insieme dei suoi punti di discontinuità $D(f)$ è trascurabile.
($\Longrightarrow$) Sia f integrabile, e dunque sicuramente limitata. Fissato un $n \in \mathbb{N}$, sia p una partizione il cui scarto sia minore di $\frac{1}{n^2}$, ossia $\rho(f,p) < \frac{1}{n^2}$. Ora, consideriamo l'insieme degli intervalli della partizione in cui l'oscillazione supera $\frac{1}{n}$

$$K_n = \left\{ k \in [n] \text{ t.c. } osc(f, I_k) > \frac{1}{n} \right\}$$

Per definizione di scarto, avremo che

$$\frac{1}{n^2} > \rho(f,p) = \sum_{k=1}^{h} |I_k| osc(f, I_k) \geq \sum_{k \in K_n} |I_k| osc(f, I_k) \geq \frac{1}{n} \sum_{k \in K_n} |I_k|$$

$$\sum_{k \in K_n} |I_k| \leq \frac{1}{n}$$

Ora, consideriamo l'insieme dei punti di discontinuità la cui oscillazione è almeno $\frac{1}{n}$. Alcuni di essi saranno contenuti negli intervalli indicizzati su K_n, altri potrebbero essere gli estremi di altri intervalli. In ogni caso, vale che

$$D_n(f) = \left\{ x \in [a,b] \text{ t.c. } osc(f,x) \geq \frac{1}{n} \right\} \subseteq p \cup \bigcup_{k \in K_n} I_k$$

L'insieme a destra è un'unione finita di intervalli (alcuni dei quali di ampiezza nulla) la cui lunghezza totale è minore di $\frac{1}{n}$. Ciò non basta ancora a concludere che $D_n(f)$ è trascurabile, perché bisogna poterlo racchiudere in un unione di intervalli di ampiezza qualsiasi. Tuttavia, possiamo osservare che, se $m > n$, $D_n(f) \subseteq D_m(f)$ e che quest'ultimo è contenuto in un unione di intervalli la cui somma delle ampiezze è minore di $\frac{1}{m}$. Data l'arbitrarietà della scelta di m, $D_n(f)$ è trascurabile. Ora, per dire che lo è anche $D(f)$ basta osservare che questo è l'unione dei $D_n(f)$, e che l'unione numerabile di insiemi trascurabili è ancora trascurabile.
($\Longleftarrow$) Sia f limitata, di oscillazione massima M. Consideriamo ancora l'insieme $D_n(f)$, che essendo un sottoinsieme di un insieme trascurabile è ancora trascurabile. Allora ricopriamolo con un insieme numerabile di intervalli aperti contenuti in $[a,b]$ la cui somma delle ampiezze è minore di $\frac{1}{n}$. Ora, $D_n(f)$ è chiuso, perché il lemma 2 ci dice che il suo complementare è aperto, e limitato essendo un sottoinsieme di $[a,b]$. Allora per teorema di Heine-Borel è compatto. Dunque, si può estrarre un sottoricoprimento finito, che possiamo chiudere. Siano I_k gli insiemi del ricoprimento, che possiamo considerare disgiunti a meno di unirli. Ora, "completiamo" la partizione aggiungendo gli intervalli J_k che coprano i "buchi" lasciati dagli I_k. Se mostriamo che lo scarto della funzione su una partizione p così ottenuta è infinitesimo abbiamo vinto.

$$\rho(f,p) = \sum_k |I_k| osc(f, I_k) + \sum_k |J_k| osc(f, J_k) \leq M \sum_k |I_k| + \frac{1}{n} \sum_k |J_k| \leq \frac{1}{n}(M + b - a)$$

□

Corollari Le seguenti osservazioni seguono tutte dalla nuova caratterizzazione dell'integrabilità.

1. $R([a,b])$ è un'algebra, perché i punti di discontinuità della somma e del prodotto sono al più contenuti nell'unione dei punti di discontinuità.

2. Una funzione continua è integrabile.

3. Una funzione monotona è integrabile, perché il suo insieme di punti di discontinuità è numerabile e dunque trascurabile.

4. Se ϕ è continua e f è integrabile, anche $\phi \circ f$ è integrabile, avendo gli stessi punti di discontinuità.

5. (Cambio di variabile) Se $g : [c, d] \longrightarrow [a, b]$ è un omeomorfismo con inversa lipschitziana di costante L e f è integrabile, allora $f \circ g$ è integrabile. Infatti, se si ricopre l'insieme dei punti di discontinuità di f con un insieme numerabile di intervalli di ampiezza totale ε, portando indietro gli intervalli per mezzo di g^{-1} si ottiene che l'insieme di punti di discontinuità di $f \circ g$ è ricoperto da un insieme numerabile di intervalli di ampiezza totale al più $L\varepsilon$.

6. Se una successione di funzioni integrabili converge uniformemente, il limite è integrabile. Infatti, dato che se tutte le funzioni sono continue in x_0 anche il limite lo è[25], l'insieme dei punti di discontinuità è al più l'unione numerabile dei punti di discontinuità delle funzioni, e dunque è trascurabile a sua volta.

[25] Vedasi 21.4.

19 Calcolo di integrali

19.1 (*)$\int_0^\pi \sin^n x\, dx$

Cerchiamo una formula ricorsiva, definendo $I(n) = \int_0^\pi \sin^n x\, dx$. Integriamo per parti:

$$\int_0^\pi \sin^n x\, dx = \int_0^\pi \sin^{n-1} x \sin x\, dx = -\sin^{n-1} x \cos x \restriction_0^\pi + \int_0^\pi (n-1) \sin^{n-2} x \cos^2 x\, dx$$

Dalla prima relazione fondamentale della goniometria, $\cos^2 x = 1 - \sin^2 x$. Allora

$$\int_0^\pi \sin^n x\, dx = \int_0^\pi (n-1) \sin^{n-2} x\, dx - \int_0^\pi (n-1) \sin^n x\, dx$$

Da cui, portando a sinistra l'integrale,

$$I(n) = \frac{n-1}{n} I(n-2)$$

Basta ora calcolare i primi due valori della successione $I(n)$, ottenendo:

$$\begin{cases} I(0) = \pi \\ I(1) = 2 \\ I(n) = \frac{n-1}{n} I(n-2) \end{cases}$$

Si può ora ricavare una formula esplicita per la successione ricorsiva, continuando a sostituire. Si ottengono due espressioni distinte, una per i pari e una per i dispari:

$$\begin{cases} I(2n) = \frac{(2n-1)!!}{(2n)!!} I(0) = \frac{(2n-1)!!}{(2n)!!} \pi \\ I(2n+1) = \frac{(2n)!!}{(2n+1)!!} I(1) = \frac{(2n)!!}{(2n+1)!!} 2 \end{cases}$$

Si osservi che i termini dispari sono tutti razionali, mentre quelli pari sono multipli razionali di π. Si osserva inoltre che la successione di funzioni $\sin^n x$ è puntualmente decrescente in n nell'intervallo $[0, \pi]$, in quanto ogni valore che il seno assume è minore di 1 e dunque le sue potenze decrescono. Ciò significa anche che la successione degli integrali è debolmente decrescente. Da queste informazioni si ricava una stima asintotica per π, che di fatto è ancora il prodotto di Wallis di cui si è già trattato in 14.4.

$$I(2n+1) \leq I(2n) \leq I(2n-1) \quad \longrightarrow \quad \frac{(2n)!!}{(2n+1)!!} 2 \leq \frac{(2n-1)!!}{(2n)!!} \pi \leq \frac{(2n-2)!!}{(2n-1)!!} 2$$

Da cui

$$\frac{((2n)!!)^2}{(2n+1)!!(2n-1)!!} \leq \pi/2 \leq \frac{(2n)!!(2n-2)!!}{((2n-1)!!)^2}$$

Poiché evidentemente[26] le due successioni hanno lo stesso limite, per teorema della Madama otteniamo

$$\pi/2 = \lim_{n\to+\infty} \frac{(2n)!!(2n-2)!!}{((2n-1)!!)^2} = \frac{2}{1}\frac{2}{3}\frac{4}{3}\frac{4}{5}\cdots$$

□

19.2 Teorema del prosciutto (Principio di Cavalieri)

[27] Il principio di Cavalieri afferma che il volume di un solido è dato dall'integrale lungo una direzione X delle sezioni del corpo sul piano perpendicolare a X. In altri termini, si scompone il

[26]Se dobbiamo essere precisi, le due successioni tendono allo stesso limite perché sono asintoticamente equivalenti, dato che il loro rapporto tende a 1.

[27]Il nome fantasioso, suggeritomi da un amico, si riferisce al fatto che, facendo ruotare completamente una curva attorno all'asse X, si ottiene un "prosciutto" il cui volume è calcolato "affettandolo".

solido in "mattoncini" di altezza dx e base $S(x)$, e si ragiona analogamente a quanto già visto nella costruzione dell'integrale di Riemann.

Come applicazione particolare, data una funzione $f(x) : [a, b] \longrightarrow \mathbb{R}$ continua, il volume del solido di rotazione ottenuto ruotando il grafico di 2π attorno all'asse X è l'integrale in dx della sezione, che in questo caso è un cerchio di raggio $f(x)$:

$$V = \int_a^b S(x)\,dx = \pi \int_a^b f^2(x)\,dx$$

Il teorema si può generalizzare a n dimensioni. Definiamo il solido n-dimensionale di rotazione attorno all'asse X_1 della funzione $f(x_1) : [a, b] \longrightarrow \mathbb{R}$ come il luogo dei punti di $\mathbb{R}^n$ la cui distanza dall'asse X_1 è minore o uguale al corrispondente valore della funzione:

$$F_n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \text{ t.c. } a \le x_1 \le b, \quad d(X, (x_1, \dots, y_n)) \le f(x_1)\}$$

Allora il volume del solido è l'integrale in dx_1 della sezione $n-1$-dimensionale, che è una palla euclidea di raggio $f(x_1)$. Per un'applicazione si veda 8.1.

19.3 Test integrale per serie

Sia $f : [0, +\infty) \longrightarrow \mathbb{R}$ decrescente (debolmente) e infinitesima. Allora l'integrale $\int_0^{+\infty} f(x)\,dx$ converge se e solo se converge la serie $\sum_{k=0}^{+\infty} f(k)$. Più precisamente, vale la seguente disuguaglianza:

$$\sum_{k=1}^{+\infty} f(k) \le \int_0^{+\infty} f(x)\,dx \le \sum_{k=0}^{+\infty} f(k)$$

La dimostrazione si basa sul fatto che, sul compatto $[0, n]$, se si considera la partizione $\{0, \dots, n\}$ l'integrale è compreso tra la somma superiore $\sum_{k=0}^{n-1} f(k)$ e la somma inferiore $\sum_{k=1}^{n} f(k)$.

19.4 Criterio di Abel-Dirichlet integrale

Sia $f : [0, +\infty) \longrightarrow \mathbb{R}$ decrescente (debolmente) e infinitesima e sia $g : [0, +\infty) \longrightarrow \mathbb{R}$ integrabile. Se gli integrali parziali di g sono equilimitati, ossia $\sup \left| \int_0^x g(x)\,dx \right| < M$ per qualche M, allora l'integrale del prodotto $\int_0^{+\infty} f(x)g(x)\,dx$ converge.

Dimostrazione Fissiamo un ε. Suddividiamo $\mathbb{R}$ in intervalli della forma $[n, n+1]$. Su ciascuno di questi determiniamo una partizione p_n in modo che la somma superiore disti meno di $\varepsilon 2^{-n}$ dall'integrale. Abbiamo allora che

$$\int_0^n f(x)g(x)\,dx \le \sum_{i=0}^{n} \sum_{I_k \in p_i} |I_k| \sup_{x \in I_k} f(x)g(x) \le \sum_{i=0}^{n} \sum_{I_k \in p_i} |I_k| f(x_k) \sup_{x \in I_k} g(x)$$

Dato che $|I_k| \le 1$ possiamo trascurare la dimensione dei sottintervalli. Numerando i punti delle partizioni otteniamo

$$\int_0^n f(x)g(x)\,dx \le \sum_{k=0}^{N} f(x_k) \sup_{x \in I_k} g(x)$$

Ora, definiamo $F(k) = f(x_k)$ e $G(k) = \sup_{x \in I_k} g(x)$. La prima successione è decrescente e infinitesima, al seconda ha le somme parziali limitate. Dunque per criterio di Abel-Dirichlet per le serie[28] la serie converge. La convergenza dell'integrale è garantita dal fatto che è possibile ripetere esattamente lo stesso procedimento con le somme inferiori. Inoltre, la possibilità di scegliere ε arbitrariamente conduce a stime vicine a piacere al valore dell'integrale, che deve quindi convergere.

□

[28] vedi 16.2

20 Funzione Gamma

20.1 Primi cenni sulla funzione fattoriale

Consideriamo l'integrale

$$F(x) = \int_0^{+\infty} t^x e^{-t}\,dt = \lim_{a\to 0}\lim_{b\to+\infty}\int_a^b t^x e^{-t}\,dt$$

e chiediamoci innanzitutto per quali valori di x converga. Per determinarlo, basta sovrastimarlo con un integrale convergente, dato che l'integrando è positivo. Osserviamo che, se $x<0$, l'integrale è improprio sia per $a\to 0$ sia per $b\to+\infty$. Spezziamolo allora in due integrali, uno su $]0,1]$ e uno su $[1,+\infty[$.
Per quanto riguarda il primo, possiamo limitare e^{-t} tra e^{-1} e 1. Abbiamo allora che

$$e^{-1}\int_0^1 t^x\,dt \le F(x) \le \int_0^1 t^x\,dt = \frac{1}{x+1}$$

Entrambi gli integrali convergono se e solo se $x>-1$. Inoltre, questa stima permette di mostrare che la funzione $F(x)$ ha un asintoto verticale in $x=-1$, dove il suo ordine di crescita è quello di un'iperbole ($\mathcal{O}(x^{-}1)$).
Consideriamo ora il secondo integrale. Osserviamo che converge sempre, in quanto

$$F(x) = \int_1^{+\infty} t^x e^{-t}\,dt = \int_1^{+\infty}\left(t^x e^{-t/2}\right)e^{-t/2}\,dt \le \max_{t\ge 1}\left(t^x e^{-t/2}\right)\int_1^{+\infty} e^{-t/2}\,dt$$

Dunque sappiamo che la funzione $F(x)$ è definita su $(-1,+\infty)$. Ora, cerchiamo di determinarne l'andamento. Supponiamo $x>0$ e integriamo per parti la scrittura della funzione.

$$F(x) = \int_0^{+\infty} t^x e^{-t}\,dt = -t^x e^{-t}\restriction_{t\to 0}^{t\to+\infty} + \int_0^{+\infty} x t^{x-1}e^{-t} \longrightarrow x\int_0^{+\infty} t^{x-1}e^{-t} = xF(x-1)$$

L'equazione funzionale $F(x)=xF(x-1)$ è la stessa soddisfatta sui naturali dal fattoriale. Dal momento che $F(0)=\int_0^{+\infty} e^{-t}\,dt = 1$, abbiamo che la funzione $F(x)$ coincide con $n!$ se ristretta ai naturali.
Ora, mostriamo che questa funzione è di classe C^∞.[29] Per farlo, definiamo le seguenti funzioni, che mostreremo essere uguali alle derivate successive.

$$F_m(x) = \int_0^{+\infty}\frac{d^m}{dx^m}t^x e^{-t}\,dt = \int_0^{+\infty}\log^m(t)t^x e^{-t}\,dt$$

Osserviamo innanzitutto che $F_0=F$. Mostriamo che tali funzioni sono definite su $(-1,+\infty)$, ossia che l'integrale converge. Per $b\to+\infty$ è banalmente vero, dato che sia il logaritmo che t^x sono dominati dall'esponenziale. Per $a\to 0$ mostriamo che l'integrale è assolutamente convergente. Per farlo, basta mostrare che l'integrale tra zero e uno è convergente.

$$\int_0^1 \left|\log^m(t)\right| t^x e^{-t}\,dt = \int_0^1 (-\log^m(t))t^x e^{-t}\,dt$$

Eseguiamo il cambio di variabile $-\log(t)=s$, da cui $dt=-e^{-s}ds$.

$$F_m(x) = \int_0^1 s^m e^{-xs}e^{-e^{-s}}(-e^{-s})\,ds = \int_0^1 s^m e^{-(x+1)s}e^{-e^{-s}}\,ds$$

Come prima, questo integrale converge per $x>-1$. Adesso, mostriamo che gli $F_m(x)$ sono ciascuno la derivata del precedente, da cui seguirà che $F(x)$ è C^∞. Dobbiamo cioè mostrare che, per $h\to 0$

$$F_m(x+h) = F_m(x) + hF_{m+1}(x) + o(h)$$

[29] Volendo, si possono saltare un sacco di passaggi tramite il teorema DUTIS, di cui a 23.13.

Per farlo, partiamo dallo sviluppo di t^h al secondo ordine, con resto di Lagrange.

$$t^h = 1 + h\log(t) + h^2 \frac{t^c \log^2(t)}{2} \quad \text{con } c \in (0, h)$$

Moltiplichiamo per $\log^m(t) t^x e^{-t}$ da ambo i membri e integriamo in dt. Otteniamo

$$\int_0^{+\infty} \log^m(t) t^{x+h} e^{-t} dt = \int_0^{+\infty} \log^m(t) t^x e^{-t} dt + h \int_0^{+\infty} \log^{m+1}(t) t^x e^{-t} dt + \frac{h^2}{2} \int_0^{+\infty} \log^{m+2}(t) t^{x+c} e^{-t} dt$$

I primi tre termini si riscrivono come

$$F_m(x+h) = F_m(x) + hF_{m+1}(x) + \frac{h^2}{2} \int_0^{+\infty} \log^{m+2}(t) t^{x+c} e^{-t}\, dt$$

Resta ora da dimostrare che l'ultimo termine è $o(h)$ per $h \to 0$, ma questo è ovvio dato che l'integrale è $F_{m+2}(x+c)$ e dunque è sempre finito.
Osserviamo inoltre che, essendo $F''(x) = F_2(x) > 0$ su tutto il dominio, la funzione è strettamente convessa.

20.2 Funzione Digamma

Consideriamo la funzione $g(x) = (\log F(x))' = \frac{F'(x)}{F(x)}$ Poiché $F(x) = xF(x-1)$ anche g soddisfa un'equazione ricorsiva per $x > 0$, del tipo

$$g(x) = (\log F(x))' = (\log x)' + (\log F(x-1))' = \frac{1}{x} + g(x-1)$$

Se ci scordiamo per un attimo che g è determinata da F, notiamo che l'equazione funzionale precedente è risolta da un'infinità di funzioni, dato che basta definire g come si vuole in $(-1, 0]$ e solo dopo determinare i valori positivi. Mostriamo tuttavia che esiste un'unica funzione g crescente in senso stretto e nulla in $x = 0$.

Esistenza Consideriamo la serie

$$g(x) = \sum_{k=1}^{+\infty} \frac{1}{k} - \frac{1}{k+x}$$

La funzione appena mostrata converge su tutto il dominio. Infatti:

$$g(x) = \sum_{k=1}^{+\infty} \frac{1}{k} - \frac{1}{k+x} = \sum_{k=1}^{+\infty} \frac{x}{k(k+x)} = x \sum_{k=1}^{+\infty} \frac{1}{k(k+x)} \leq x \left(\frac{1}{k(k+x)} + \sum_{k=2}^{+\infty} \frac{1}{k(k-1)} \right)$$

La seconda somma converge essendo la serie di Mengoli. Inoltre, g è nulla in zero, banalmente crescente, poiché lo sono i singoli termini, e per x positivo soddisfa l'equazione funzionale, dato che

$$g(x) = \sum_{k=1}^{+\infty} \frac{1}{k} - \frac{1}{k+x} = \frac{1}{x} + \sum_{k=1}^{+\infty} \frac{1}{k} - \frac{1}{k-1+x} = \frac{1}{x} + g(x-1)$$

Unicità Sui naturali la funzione g deve soddisfare le condizioni

$$\begin{cases} g(0) = 0 \\ g(n) = \frac{1}{n} + g(n-1) = \sum_{k=1}^{n} \frac{1}{k} \end{cases}$$

Ora, consideriamo un'altra soluzione h che sia crescente e nulla in zero. Per quanto visto, h sarà anch'essa univocamente determinata sui naturali. Se identifichiamo con $\lfloor x \rfloor$ la parte intera (inferiore), la differenza tra le due funzioni si stima come

$$g(x) - h(x) \leq g(\lfloor x \rfloor + 1) - h(\lfloor x \rfloor) = g(\lfloor x \rfloor + 1) - g(\lfloor x \rfloor) = \frac{1}{\lfloor x \rfloor + 1}$$

La disuguaglianza segue dalla monotonia delle funzioni. Si vede che la differenza tende a zero. Per simmetria, si può dire che $|g(x) - h(x)| = o(1)$ per $x \to +\infty$. Tuttavia

$$g(x) - h(x) = g(x-1) - h(x-1) + \frac{1}{x} - \frac{1}{x} = g(x-1) - h(x-1)$$

ossia la differenza tra le funzioni è periodica. Dovendo essere infinitesima all'infinito, essa è nulla, ossia $h \equiv g$.

□

La funzione g determinata è concava, perché somma di funzioni concave[30]. Inoltre è analitica, ossia è di classe C^∞ e ogni punto ammette un intorno in cui la funzione coincide con il suo sviluppo in serie di Taylor. In altri termini, è localmente uguale a una serie di potenze con centro in ogni suo punto. É possibile dimostrare[31] che una serie di potenze è analitica nel suo dominio, dunque basta mostrare che per ogni punto y esiste un intorno in cui g è una serie di potenze (sviluppata in un qualche centro). Per mostrarlo, osserviamo innanzitutto che se $|x| < 1$ è possibile sviluppare ogni termine come somma geometrica:

$$g(x) = \sum_{k=1}^{+\infty} \frac{1}{k} - \frac{1}{k+x} = \sum_{k=1}^{+\infty} \frac{1}{k}\left(1 - \frac{1}{1+\frac{x}{k}}\right) = \sum_{k=1}^{+\infty} \frac{1}{k}\left(1 - \sum_{n=0}^{+\infty}\left(\frac{-x}{k}\right)^n\right) =$$

$$= \sum_{k=1}^{+\infty}\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}x^n}{k^{n+1}} = \frac{1}{x}\sum_{k=1}^{+\infty}\sum_{n=2}^{+\infty} \frac{(-1)^n x^n}{k^n}$$

La somma è assolutamente convergente. Infatti:

$$\frac{1}{x}\sum_{k=1}^{+\infty}\sum_{n=2}^{+\infty} \frac{|x|^n}{k^n} = \frac{1}{x}\sum_{k=1}^{+\infty} \frac{1}{1-\frac{|x|}{k}} - 1 - \frac{|x|}{k} = \frac{1}{x}\sum_{k=1}^{+\infty} \frac{|x|^2}{k^2\left(1-\frac{|x|}{k}\right)}$$

Abbiamo che, per ipotesi, $|x| < 1$. Inoltre, per k abbastanza grandi, $1 - \frac{|x|}{k} > \frac{1}{2}$. Dunque possiamo stimare il modulo come

$$\frac{1}{|x|}\sum_{k=1}^{+\infty} \frac{|x|^2}{k^2\left(1-\frac{|x|}{k}\right)} \leq \frac{2}{|x|}\sum_{k=1}^{+\infty} \frac{1}{k^2}$$

Che notoriamente converge, se si fissa x. Ora, tornando allo sviluppo come somma di serie di potenze, il fatto che la serie sia assolutamente convergente permette di riordinare la sommatoria, raccogliendo i termini con lo stesso esponente:

$$\frac{1}{x}\sum_{k=1}^{+\infty}\sum_{n=2}^{+\infty} \frac{(-1)^n x^n}{k^n} = \frac{1}{x}\sum_{n=2}^{+\infty} x^n \sum_{k=1}^{+\infty} \frac{(-1)^n}{k^n} = \sum_{n=1}^{+\infty} (-1)^{n+1} x^n \sum_{k=1}^{+\infty} \frac{1}{k^{n+1}}$$

Dunque la funzione si scrive come serie di potenze in $(-1, 1)$. Ora, dato che g risolve l'equazione funzionale $g(x) = g(x-1) + \frac{1}{x}$ osserviamo che, se in $x_0 - 1$ naturale la funzione si sviluppa come serie di potenze, ciò accade anche in x_0, in quanto anche la funzione $\frac{1}{x}$ è analitica in ogni punto naturale, con raggio di convergenza almeno 1. Infatti, è possibile svilupparla come serie di potenze:

$$\frac{1}{x_0 + \delta} = \frac{1}{x_0}\frac{1}{1 + \frac{\delta}{x_0}} = \frac{1}{x_0}\sum_{k=0}^{+\infty} \frac{(-1)^k \delta^k}{x_0^k}$$

Dunque, procedendo induttivamente, otteniamo che la funzione g è analitica su tutto il dominio.

[30]Si dovrebbe mostrare che la concavità passa al limite, ma abbiamo già osservato in preccedenza che le proprietà puntuali passano sempre al limite. Non è invece detto che la concavità rimanga stretta.

[31]La dimostrazione è un po' lunga e viene omessa, quando avrò tempo la aggiungerò.

20.3 Primitiva di g e costante di Eulero-Mascheroni

Cerchiamo ora una primitiva $\Phi(x)$ della funzione g, con la proprietà di annullarsi in $x = 0$ (sempre facendo finta di non sapere che $g = F'/F$).
Consideriamo allora la successione $a_n(x) = \frac{x}{n} - \log\left(1 + \frac{x}{n}\right)$ e sia $\Phi(x) = \sum_{n=0}^{+\infty} a_n(x)$. La derivata del singolo termine è $a_n'(x) = \frac{1}{n} - \frac{1}{n+x}$, cioè è il termine generale della serie di g, e da questo vorremmo dedurre che

$$g(x) = \sum_{n=0}^{+\infty} \frac{1}{n} - \frac{1}{n+x} = \sum_{n=0}^{+\infty} a_n'(x) = \left(\sum_{n=0}^{+\infty} a_n(x)\right)' = (\Phi(x))'$$

cioè, vorremmo poter invertire la derivata e il limite. Questo, per un teorema sulle serie di funzioni,[32] è possibile se la serie delle derivate converge uniformemente al suo limite (in questo caso, $g(x)$) , e se esiste almeno un punto x_0 in cui la serie $\Phi(x)$ converge. Per quanto riguarda la prima condizione, sappiamo che la serie di g converge assolutamente su $(-1, +\infty)$. Questa condizione, nello spazio delle funzioni continue, è equivalente alla convergenza uniforme della serie, intesa come funzione limite delle somme parziali[33]. Dunque la serie delle derivate converge uniformemente a g.
Basta ora mostrare che esiste un punto in cui $\Phi(x)$ converge. Osserviamo che ciò accade in $x_0 = 0$, in quanto

$$\Phi(0) = \sum_{n=0}^{+\infty} a_n(0) = \sum_{n=0}^{+\infty} \frac{0}{n} - \log\left(1 + \frac{0}{n}\right) = 0$$

Dunque abbiamo determinato una funzione $\Phi(x)$ nulla in zero e la cui derivata è g. Per il teorema fondamentale del calcolo integrale vale allora che

$$\Phi(x) = \Phi(x) - \Phi(0) = \int_0^x g(t)\, dt$$

Un valore rilevante della funzione appena determinata è $\Phi(1) = \gamma \approx 0,577\ldots$, che prende il nome di costante di Eulero-Mascheroni.[34]

$$\gamma = \Phi(1) = \sum_{n=0}^{+\infty} \frac{1}{n} - \log\left(1 + \frac{1}{n}\right) = \lim_{N\to+\infty} \sum_{n=0}^{N} \frac{1}{n} - \sum_{n=0}^{N} \log\left(1 + \frac{1}{n}\right) =$$

$$= \lim_{N\to+\infty} \sum_{n=0}^{N} \frac{1}{n} - \log\left(\prod_{n=1}^{N} \frac{n+1}{n}\right) = \lim_{N\to+\infty} \sum_{n=0}^{N} \frac{1}{n} - \log(N+1)$$

Poiché $\log(N + 1)$ e $\log(N)$ sono asintotici, possiamo scrivere anche

$$\gamma = \lim_{N\to+\infty} \sum_{n=0}^{N} \frac{1}{n} - \log(N)$$

Dato che γ è una costante finita, abbiamo ottenuto che la serie armonica è asintotica al logaritmo naturale per $N \to +\infty$.
Mettendo assieme i pezzi, possiamo determinare un'equazione funzionale soddisfatta da $\Phi(x)$, nota la sua definizione come integrale:

$$\Phi(x) = \int_0^x g(t)\, dt = \int_0^1 g(t)\, dt + \int_1^x g(t)\, dt = \gamma + \int_1^x \frac{1}{t} + \int_1^x g(t-1)\, dt =$$

$$= \gamma + \log(x) + \int_0^{x-1} g(t)\, dt = \gamma + \log(x) + \Phi(x-1)$$

[32]Il lettore mi scuserà se d'ora in poi farò uso di qualche cannone che richiede conoscenze un po' più avanzate, ma non ho davvero voglia di fare tutti i conti manuali dove esistono dimostrazioni più eleganti.

[33]Questo criterio è noto come criterio di convergenza totale, o di Weierstrass, ed è equivalente alla completezza dello spazio vettoriale delle funzioni continue. Non ho tempo di dimostrare né l'una né l'altra cosa, anche questo è nella mia lista delle cose da fare.

[34]Ci sono diversi problemi aperti su questa costante. Ad esempio, si pensa che possa essere razionale, ma al momento non lo si è ancora dimostrato o confutato.

20.4 Teorema di Artin-Bohr-Mollerup

Esiste ed è unica una funzione $F : (-1, +\infty) \longrightarrow \mathbb{R}$[35] tale che

- $F(0) = 1$
- $F(x) = xF(x-1)$
- $\log F(x)$ è una funzione convessa.

Unicità Riformuliamo l'enunciato considerando la funzione $f = \log F$. Dobbiamo mostrare che esiste una sola f tale che

- $f(0) = 0$
- $f(x) = \log x + f(x-1)$
- $f(x)$ è una funzione convessa.

Dato che la funzione è convessa, essa ammette derivata destra crescente in ogni punto. Tramite la definizione di limite del rapporto incrementale da destra, si ottiene che

$$f'_+(x) = \frac{1}{x} + f'_+(x-1)$$

Ma allora f'_+ soddisfa la stessa equazione che caratterizza $g(x)$ ed è crescente. Allora, la funzione $f'_+(x) - f'_+(0)$, nulla in $x = 0$, deve coincidere con $g(x)$, per quanto visto in 20.2. Ora, la funzione $g(x)$ è continua, e dunque lo è $f'_+(x)$. Si può mostrare[36] che una funzione convessa f ammette derivata destra continua se e solo se è derivabile. Allora f è derivabile e vale che

$$f'(x) = f'(0) + g(x)$$

Integrando tra zero e x otteniamo, per teorema fondamentale del calcolo integrale, che

$$f(x) = f(x) - f(0) = \int_0^x f'(t)\,dt = \int_0^x f'(0) + g(t)\,dt = xf'(0) + \Phi(x)$$

Ora, determiniamo $f'(0)$ noto che dobbiamo avere

$$f(1) = \log 1 + f(0) = 0$$

$$0 = f(1) = f'(0) + \Phi(1) = f'(0) + \gamma \quad \longrightarrow \quad f'(0) = -\gamma$$

Allora f, se esiste, è univocamente determinata e vale

$$f(x) = -\gamma x + \Phi(x)$$

Da cui

$$F(x) = \exp\left(-\gamma x + \Phi(x)\right)$$

Esistenza Mostriamo che la funzione determinata al punto precedente esiste su $(-1, +\infty)$ e soddisfa le tre richieste.

- L'esistenza è assicurata dal fatto che l'esponenziale è un omeomorfismo e $\Phi(x)$ è definita su $(-1, +\infty)$.
- $F(0) = e^0 = 1$.

[35] Con un leggero abuso di notazione, useremo $F(x)$ per la funzione determinata da questo teorema, che a priori potrebbe non coincidere con l'omonima funzione identificata in 20.1. Verificheremo in seguito che le due funzioni coincidono, mostrando che il logaritmo della funzione già vista è convesso.

[36] Vedi 11.8.

- Mostriamo che F soddisfa l'equazione funzionale.

$$F(x) = \exp(-\gamma x + \Phi(x)) = \exp(-\gamma - \gamma(x-1) + \gamma + \log(x) + \Phi(x-1)) =$$
$$= x\exp(-\gamma(x-1) + \Phi(x-1)) = xF(x-1)$$

- Mostriamo che il suo logaritmo f è convesso. Per farlo, basta mostrare che lo è $\Phi(x)$, dato che la retta γx lo è sicuramente. Ricordando che

$$\Phi(x) = \sum_{k=1}^{+\infty} \frac{x}{k} - \log\left(1 + \frac{x}{k}\right)$$

  Osserviamo che i singoli termini della serie sono convessi, dato che il primo è lineare e il secondo è l'opposto di una funzione concava. Allora, per quanto già osservato la convessità passa al limite.

Concludiamo allora che

$$x! = \exp(-\gamma x + \Phi(x))$$

□

La funzione che si è determinata prende il nome di funzione fattoriale $x!$. La sua traslazione di 1 lungo l'asse x è detta funzione Gamma di Eulero:

$$\Gamma(x) = (x-1)!$$

La derivata logaritmica della funzione gamma (quella che finora abbiamo chiamato g) si indica in letteratura con $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ e prende il nome di funzione *digamma* (sia perché ci sono due gamma, sia perché ricorda nella "forma" la lettera F, il cui fonema era associato alla lettera greca arcaica *digamma*).

20.5 Rappresentazioni della funzione fattoriale

Formula di Eulero Mostriamo che la funzione fattoriale coincide con la funzione

$$F(x) = \int_0^{+\infty} t^x e^{-t}\, dt$$

vista in precedenza. Per farlo, sfruttiamo la caratterizzazione di $x!$ e mostriamo che F soddisfa le tre condizioni del teorema precedente. Abbiamo già osservato che $F(0) = 1$ e che la funzione soddisfa l'equazione funzionale. Ci chiediamo se il suo logaritmo sia convesso. Dato che F è di classe C^∞ è equivalente chiedersi se la derivata seconda di $\log F(x)$ sia positiva o nulla. Sviluppando la formula della derivata logaritmica otteniamo

$$(\log F)'' = \left(\frac{F'}{F}\right)' = \frac{F''F - (F')^2}{F^2} \geq 0$$

Basta verificare che

$$F''F \geq (F')^2$$
$$\left(\int_0^{+\infty} \log^2(t) t^x e^{-t}\, dt\right)\left(\int_0^{+\infty} t^x e^{-t}\, dt\right) \geq \left(\int_0^{+\infty} \log(t) t^x e^{-t}\, dt\right)^2$$

Ora, sullo spazio delle funzioni continue è possibile definire il seguente prodotto scalare, definito positivo, tramite l'integrale:

$$\langle f, g\rangle = \int_0^{+\infty} f(t)g(t)\, dt$$

Tale prodotto scalare induce una norma, detta norma 2, data da

$$\|f\|_2 = \sqrt{\langle f, f\rangle} = \sqrt{\int_0^{+\infty} f^2(t)\, dt}$$

Allora vale la disuguaglianza di Cauchy-Schwarz:[37]

$$\langle f, g\rangle^2 \leq \|f\|_2\|g\|_2$$

Scegliendo

$$f = \log(t)t^{x/2}e^{-t/2} \quad g = t^{x/2}e^{-t/2}$$

Otteniamo la disuguaglianza cercata, in quanto

$$\left(\int_0^{+\infty} \log(t)t^x e^{-t}\, dt\right)^2 \leq \left(\int_0^{+\infty} \log^2(t)t^x e^{-t}\, dt\right)\left(\int_0^{+\infty} t^x e^{-t}\, dt\right)$$

$\square$

Formula di Gauss Sviluppiamo la scrittura della funzione fattoriale.

$$x! = \exp(-\gamma x + \Phi(x)) = \exp\left(\lim_{n\to+\infty} x\log n - x\sum_{k=1}^{n}\frac{1}{k} + \sum_{k=1}^{n}\frac{x}{k} - \log\left(1+\frac{x}{k}\right)\right) =$$

$$= \exp\left(\lim_{n\to+\infty} x\log n - \sum_{k=1}^{n}\log\left(1+\frac{x}{k}\right)\right)$$

Dato che l'esponenziale è una funzione continua, il limite può essere portato dentro.

$$x! = \lim_{n\to+\infty}\exp\left(x\log n - \sum_{k=1}^{n}\log\left(1+\frac{x}{k}\right)\right) = \lim_{n\to+\infty}\frac{n^x}{\prod_{k=1}^{n} 1+\frac{x}{k}} =$$

$$= \lim_{n\to+\infty}\frac{n^x}{\prod_{k=1}^{n}\frac{x+k}{k}} = \lim_{n\to+\infty}\frac{n^x n!}{(x+1)\dots(x+n)}$$

Con un leggero abuso di notazione, possiamo scrivere

$$x! = \lim_{n\to+\infty}\frac{n^x}{\binom{x+n}{n}}$$

Si può mostrare che questa formula permette di estendere la funzione fattoriale anche a valori della x negativi, purché non interi (altrimenti, per induzione, si otterrebbe un valore finito in $x=-1$). In altri termini, la formula di Gauss è una continuazione analitica della funzione fattoriale oltre il suo dominio di definizione.

Formula di Weierstrass Sempre sviluppando la scrittura di $x!$ possiamo ottenere

$$x! = \exp(-\gamma x + \Phi(x)) = e^{-\gamma x}\prod_{k=1}^{+\infty}\frac{e^{\frac{x}{k}}}{1+\frac{x}{k}}$$

Anche questa scrittura si estende ai numeri negativi non interi.

20.6 Fattoriale di $-1/2$ e integrale della gaussiana.

In generale, i valori esatti di quasi tutti i fattoriali non interi non sono noti. Tuttavia, esistono alcuni risultati rilevanti. Ad esempio, calcoliamo, mediante la formula di Gauss, il fattoriale di $-1/2$.

$$(-\frac{1}{2})! = \lim_{n\to+\infty}\frac{n^{-1/2}n!}{(1-\frac{1}{2})\dots(n-\frac{1}{2})} = \lim_{n\to+\infty}\frac{n^{-1/2}n!2^n}{(2-1)\dots(2n-1)} = \lim_{n\to+\infty}\frac{n^{-1/2}n!2^n}{(2n-1)!!}$$

[37] Vedi 10.1

Moltiplichiamo sopra e sotto per $2n!! = 2^n n!$ per completare il bifattoriale

$$(-\frac{1}{2})! = \lim_{n\to+\infty} \frac{n^{-1/2} n! n! 4^n}{(2n)!} = \lim_{n\to+\infty} \frac{n^{-1/2} 4^n}{\binom{2n}{n}}$$

Ricordando lo sviluppo asintotico visto in 2.4, possiamo calcolare il valore del limite:

$$(-\frac{1}{2})! \lim_{n\to+\infty} \frac{n^{-1/2} 4^n}{(\pi n)^{-1/2} 4^n} = \sqrt{\pi}$$

Il risultato ottenuto permette di calcolare l'integrale di e^{-x^2}, ossia della distribuzione gaussiana di probabilità. Infatti, per la formula di Eulero abbiamo che

$$\sqrt{\pi} = (-\frac{1}{2})! = \int_0^{+\infty} t^{-1/2} e^{-t}\, dt$$

Cambiamo variabile, scegliendo $t = s^2$, da cui $dt = 2sds$. Otteniamo allora che

$$\sqrt{\pi} = 2\int_0^{+\infty} e^{-s^2}\, ds$$

Da cui, essendo la funzione pari, otteniamo

$$\sqrt{\pi} = \int_{-\infty}^{+\infty} e^{-s^2}\, ds$$

20.7 Sviluppi asintotici

- Prendendo il logaritmo del fattoriale otteniamo

  $$\log x! = -\gamma x + \Phi(x)$$

  Da cui

  $$\psi(x) = (\log x!)' = -\gamma + \phi(x)$$

  Ora, per come è definita la costante di Eulero-Mascheroni

  $$\gamma = \sum_{k=1}^{n} \frac{1}{k} - \log n + o(1) = \phi(n) - \log n + o(1)$$

  Sostituendo, abbiamo che la funzione *digamma* va come il logaritmo, in quanto

  $$\psi(x) = \log n + o(1)$$

- Cerchiamo ora una stima asintotica per il fattoriale di $(x + a)$. Mostreremo che

  $$(x + a)! \approx x! x^a$$

  Procedendo in modo simile all'induzione.

  - La tesi è vera per $a = 0$.
  - Se è vera per a, allora vale anche per $a + 1$, dato che il fattoriale si sviluppa come

    $$(x + a + 1)! = (x + a)!(x + a + 1) \approx x! x^a (x + a + 1) \approx x! x^{a+1}$$

  - Resta da mostrare la tesi per $0 < a < 1$. Osserviamo che è equivalente mostrare che il rapporto tra i due membri tende a 1, o ancora che il logaritmo del rapporto è infinitesimo. Allora

    $$\log\left(\frac{(x + a)!}{x! x^a}\right) = \log(x + a)! - \log x! - a \log x$$

Sia $h(a)$ la funzione alla RHS. Osserviamo che $h(0) = 0$. Allora, dividendo per a otteniamo il rapporto incrementale della funzione tra 0 e a. Possiamo dunque applicare il teorema del valor medio, nella sua forma più debole:

$$\frac{h(a) - h(0)}{a} \leq \sup_{0<t<a} \frac{d}{dt} (\log(x + t)! - \log x! - t \log x) = \sup_{0<t<a} (\log(x + t)!)' - \log x$$

Per quanto già osservato, la funzione digamma è asintotica al logaritmo, e dunque, dato che la crescita del logaritmo è infinitesima, otteniamo

$$h(a) \leq a \sup_{0<t<a} \log(x + t) - \log x = o(1)$$

Che è quanto volevamo mostrare.

□

20.8 Formula di Stirling

Vogliamo mostrare che lo sviluppo asintotico

$$x! = \sqrt{2\pi x} \left(\frac{x}{e}\right)^x (1 + o(1))$$

si estende a tutto il dominio della funzione. Per farlo mostreremo che

$$\log \frac{x^{x+1/2} e^{-x}}{x!} + \frac{1}{2} \log(2\pi) = o(1)$$

Sia g la funzione alla LHS. Dato che sappiamo già che lo sviluppo è corretto per x intero, ci basta verificare che la funzione non sbarella negli altri punti. A tal fine, possiamo mostrare che la derivata della funzione è infinitesima.

$$(\log x^{x+1/2} - x - \log(x!))' = \left(\log x \left(x + \frac{1}{2}\right) - x - \log(x!)\right)' =$$

$$= 1 + \frac{1}{2x} + \log x - 1 - (\log(x!))' \approx \frac{1}{2x} + \log x - \log x' = o(1)$$

□

20.9 Formule di moltiplicazione

La prima formula che vedremo è quella di duplicazione:

$$x! = \frac{2^x}{\sqrt{\pi}} \left(\frac{x}{2}\right)! \left(\frac{x-1}{2}\right)!$$

Sia h la funzione alla RHS. Useremo la caratterizzazione della funzione fattoriale per mostrare l'identità.

- $h(0) = \frac{1}{\sqrt{\pi}} \left(\frac{-1}{2}\right)! = 1$;
- $h(x) = xh(x - 1)$. Infatti

$$h(x) = \frac{2^x}{\sqrt{\pi}} \left(\frac{x}{2}\right)! \left(\frac{x-1}{2}\right)! = x\frac{2^{x-1}}{\sqrt{\pi}} \left(\frac{x-2}{2}\right)! \left(\frac{x-1}{2}\right)! = xh(x - 1)$$

- Mostriamo che il logaritmo della funzione h è convesso. Ciò è banale, in quanto prendendo il logaritmo si ottiene un pezzo lineare, ovviamente convesso, e due logaritmi di un fattoriale, che per caratterizzazione sono convessi.

□

Osserviamo che dall'identità precedente segue nuovamente lo sviluppo asintotico del binomiale centrale. Infatti, usando lo sviluppo visto in 20.7 $\left(\frac{x-1}{2}\right)! \approx \left(\frac{x}{2}\right)! x^{-1/2}$, otteniamo che

$$\binom{x}{x/2} \approx \frac{2^x}{\sqrt{\pi x}}$$

Ora, l'obiettivo sarebbe estendere la formula precedente al caso generale. Procedendo esattamente allo stesso modo, si mostra che per ogni naturale n esiste una costante C_n tale che

$$x! = C_n n^x \left(\frac{x}{n}\right)! \left(\frac{x-1}{n}\right)! \dots \left(\frac{x-n+1}{n}\right)!$$

Quanto vale C_n? Bella domanda. Cioè, è ovvio che debba essere scelta in modo che $0! = 1$, e dunque deve valere che

$$\frac{1}{C_n} = \prod_{k=1}^{n-1} \left(\frac{k}{n}\right)!$$

Dato che molti dei valori del fattoriale sui razionali non sono noti, questa espressione è praticamente irresolubile. Si può, invece, cercare di stimarne il valore tramite lo sviluppo di Stirling. Prendiamo i singoli termini della formula di moltiplicazione, estraiamone il logaritmo (in modo da avere solo delle somme) e dopo imponiamo l'uguaglianza asintotica. É più comodo scrivere la formula nel modo seguente:

$$nx! = C_n n^{nx} (x)! \left(x - \frac{1}{n}\right)! \dots \left(x - \frac{n+1}{n}\right)!$$

$$\log(nx!) = \log\left(\sqrt{2\pi} nx^{nx+1/2} e^{-nx}\right) = \frac{1}{2}\log(2\pi) + nx\log nx + \frac{1}{2}\log nx - nx + o(1) =$$

$$= \frac{1}{2}\log(2\pi) + nx\log n + nx\log x + \frac{1}{2}\log n + \frac{1}{2}\log x - nx + o(1)$$

D'ora in avanti, sia $H(x)$ una generica funzione che si scrive come somma di termini del tipo:

$$H(x) = ax\log x + b\log x - cx + o(1)$$

Osserviamo che per $x \to +\infty$ una funzione siffatta può solo tendere a infinito o a zero, dato che ogni termine è di orndine strettamente maggiore dei successivi. L'obiettivo del ragionamento è allora cercare di ottenere un'uguaglianza tra una costante e una funzione $H(x)$, cosa che porterà a dire che la costante deve essere nulla. Scriviamo allora

$$\log(nx!) = \frac{1}{2}\log(2\pi) + \frac{1}{2}\log n + H(x)$$

$$\log\left(x - \frac{k}{n}\right)! \approx \log x! + \log x^{-k/n} \approx \frac{1}{2}\log(2\pi) + H(x)$$

$$\log n^{nx} = H(x)$$

Allora otteniamo

$$\frac{1}{2}\log(2\pi) + \frac{1}{2}\log n = \log C + \frac{m}{2}\log(2\pi) + H(x)$$

Da questa uguaglianza, per quanto osservato, segue che

$$\frac{1}{2}\log(2\pi) + \frac{1}{2}\log n = \log C + \frac{m}{2}\log(2\pi)$$

Da cui, togliendo i logaritmi:

$$C = \frac{\sqrt{n}}{(2\pi)^{(m-1)/2}}$$

20.10 Formula di riflessione

La formula permette di calcolare il valore del fattoriale dell'opposto di un numero $x \in \mathbb{R} \setminus \mathbb{Z}$. Utilizzando la formula di Weierstrass si mostra che

$$x!(-x)! = e^{-\gamma x + \gamma x} \prod_{k \in \mathbb{N}^+} \frac{e^{\frac{x}{k} - \frac{x}{k}}}{\left(1 + \frac{x}{k}\right)\left(1 - \frac{x}{k}\right)} = \prod_{k \in \mathbb{N}^+} \frac{1}{1 - \frac{x^2}{k^2}}$$

Ora, ricollegandosi a quanto mostrato in 14.3, è possibile ricondursi allo sviluppo alternativo della funzione seno. Infatti:

$$\sin x = x \prod_{k \in \mathbb{N}^+} \left(1 - \frac{x^2}{\pi^2 k^2}\right) \quad \longrightarrow \quad x!(-x)! = \frac{1}{\prod_{k \in \mathbb{N}^+} \left(1 - \frac{x^2}{k^2}\right)} = \frac{\pi x}{\sin(\pi x)}$$

20.11 Funzione Beta

Prima di gingillarsi con la funzione Gamma, Eulero aveva pensato a un altro integrale divertente, cui aveva dato il nome di funzione Beta:

$$\beta(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt$$

Con un mucchio di conti che risparmio al lettore, si può mostrare che:

1. $\beta(x, y)$ è simmetrica rispetto allo scambio delle variabili ed assume valori finiti solo se $x > 0$ e $y > 0$;
2. $\beta(x, y)$ è di classe C^∞ come funzione in x, e la derivata passa dentro l'integrale;
3. Vale la seguente relazione con la funzione Gamma:

$$\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}$$

Si fa tutto esattamente come visto nei paragrafi precedenti, solo che lo si fa in una sola lezione. Si osservi che, come la funzione Gamma generalizza il fattoriale, la funzione Beta generalizza il coefficiente binomiale:

$$\binom{n}{k} = \frac{1}{(n+1)\beta(n+k-1,\, k+1)}$$

21 Successioni e serie di funzioni

21.1 Convergenza puntuale e uniforme

Data una successione di funzioni $f_n(x) : X \longrightarrow Y$, si dice che esse convergono puntualmente alla funzione $f(x)$ se, fissato comunque un punto $x_0 \in X$, $\lim_{n\to+\infty} f_n(x_0) = f(x_0)$. Si osserva che la convergenza può essere più o meno veloce a seconda di che punto si considera. In simboli

$$\forall x_0 \in X\, \forall \varepsilon\, \exists n(\varepsilon, x_0) \text{ tale che } \forall m \geq n(\varepsilon, x_0) \text{ si ha che } d(f_m(x_0), f(x_0)) < \varepsilon$$

Si dice che la successione converge uniformemente se definitivamente ogni funzione sta in una fascia larga ε attorno alla funzione limite. In simboli

$$\forall \varepsilon\, \exists n(\varepsilon) \text{ tale che } \forall m \geq n(\varepsilon)\, \forall x_0 \in X \text{ si ha che } d(f_m(x_0), f(x_0)) < \varepsilon$$

In altri termini, l'indice per cui la successione sta a meno di ε dal limite è indipedente dal punto scelto. La condizione di convergenza uniforme implica la convergenza puntuale ma è assai più potente.

21.2 Norma infinito

Dato uno spazio di funzioni reali limitate $f : X \longrightarrow \mathbb{R}$ si definisce la norma "infinito"[38] come

$$\|f\|_\infty = \sup_{x\in X} |f(x)|$$

Si verifica facilmente che la funzione definita sopra è una norma[39] e induce la seguente distanza, detta distanza del sup, distanza infinito o distanza uniforme:

$$d(f, g)_\infty = \sup_{x\in X} |f(x) - g(x)|$$

Maneggiando le definizioni, è facile vedere che una successione di funzioni converge assolutamente se e soltanto se la distanza infinito tra le funzioni tende a zero. Si dice che la distanza del sup "metrizza" la convergenza uniforme.

21.3 (*)Completezza delle funzioni limitate

Aggiungo questo paragrafo perché si capiscano meglio molte proprietà delle successioni di funzioni. Sia E uno spazio topologico. L'insieme $(B(E), d_\infty)$ delle funzioni limitate (bounded) da E in $\mathbb{R}$ è uno spazio metrico completo con la distanza uniforme. Ossia, data una successione di funzioni limitate f_n di Cauchy nella distanza uniforme, esiste una funzione limitata f cui la successione converge uniformemente.

Dimostrazione Sia f_n come sopra. La proprietà di Cauchy si esprime nel modo seguente:

$$\forall \varepsilon \exists n_o \text{ t.c. } \forall m, n \geq n_0, \quad d_\infty(f_n, f_m) = \sup_{x\in E} |f_n(x) - f_m(x)| < \varepsilon$$

Dunque, in particolare, se si fissa un x_0 la successione $f_n(x_0)$ è una successione di Cauchy di numeri reali, che dunque converge a un certo $f(x_0)$. In questo modo, si identifica una funzione $f(x)$ che è il limite puntuale della successione. Bisogna ora mostrare che è limitata e che la convergenza è uniforme. Per entrambe le richieste basta osservare che, definitivamente, $d_\infty(f_n, f) < \varepsilon$ essendo la successione uniformemente di Cauchy.

□

[38] Il motivo di questo nome, "di una certa vaghezza poetica" (cit.), viene dal fatto che la norma infinito è il limite per $p \to +\infty$ della norma p-esima, definita come $\|f\|_p = \sqrt[p]{\int_X |f(x)|\, dx}$, con $p \geq 1$. La dimostrazione nonè così semplice.

[39] Impropriamente, si può estendere il concetto di norma anche a funzioni non limitate, cosa che tuttavia crea dei problemi per quanto riguarda la convergenza.

21.4 Continuità del limite uniforme

Sia $f_n : (X, D) \longrightarrow (Y, d)$ una successione di funzioni che converge uniformemente a una funzione f. Se f_n sono continue in x_0, lo è anche il limite.

Dimostrazione Dato che la convergenza è uniforme, per ogni ε esiste un n tale che, per ogni $m \geq n$, $d_\infty(f, f_m) < \varepsilon$. Ora, possiamo scrivere per disuguaglianza triangolare che:

$$d(f(x_0), f(x)) \leq d(f(x_0), f_n(x_0)) + d(f_n(x_0), f_n(x)) + d(f_n(x), f(x))$$

Le due distanze laterali si stimano con la distanza uniforme. Il termine centrale, per $x \to x_0$, è infinitesimo, essendo f_n continua in x_0. Allora esiste un certo δ tale che

$$D(x, x_0) < \delta \quad \Longrightarrow \quad d(f(x_0), f(x)) \leq 2d_\infty(f, f_n) + d(f_n(x_0), f_n(x)) \leq 3\varepsilon$$

□

Da quanto appena visto, possiamo concludere anche che le funzioni continue sono un sottoinsieme chiuso delle funzioni limitate da (X, D) in $\mathbb{R}$, e dunque sono anch'esse un insieme completo rispetto alla distanza infinito.

21.5 Convergenza di funzioni equicontinue

Equicontinuità Delle funzioni si dicono equicontinue se sono uniformemente continue e hanno lo stesso modulo di continuità.

Rigidità delle funzioni equicontinue Se una successione di funzioni equicontinue converge puntualmente, il limite è equicontinuo alle funzioni.
Infatti, dato che vale $d(f_n(x), f_n(y)) \leq \omega(d(x, y))$, basta passare al limite su n e osservare che la distanza è una funzione continua in entrambe le variabili. Dunque

$$\lim_{n\to+\infty} d(f_n(x), f_n(y)) = d(\lim_{n\to+\infty} f_n(x), \lim_{n\to+\infty} f_n(y)) \leq \omega(d(x, y)) \Longrightarrow d(f(x), f(y)) \leq \omega(d(x, y))$$

L'implicazione segue dal teorema della permanenza del segno, e dal fatto che la successione è puntualmente convergente.

□

Abbiamo anche mostrato che il limite puntuale di una successione equicontinua è uniformemente continuo, dato che ammette un modulo di continuità.[40]

21.6 Teorema del ConDom

> Non dovrei fare sex jokes at school... but protected sex is okay.
>
> *Pietro Majer*

Sia $(a_{k,n})_{n\in\mathbb{N},k\in I}$ una famiglia di reali. Se vale che

1. Per $n \to +\infty$, $a_{k,n}$ tende a $\bar{a}_k$ reale;
2. $\forall k \in I\ \exists b_k$ tale che $\forall n \in \mathbb{N}$ vale $|a_{k,n}| \leq b_k$, ossia la successione $a_{k,n}$ è maggiorata in modulo dalla costante b_k;
3. $\sum_{k\in I} |b_k| < +\infty$, cioè i b_k sono assolutamente sommabili;

[40] In generale, tutte le proprietà che si esprimono in funzione di un numero finito di punti (in questo caso, due, ossia x e y) vengono conservate nel passaggio al limite puntuale. Non è sempre detto, invece, che si conservi la continuità, che è una proprietà di un intorno, ossia di un insieme non numerabile di punti.

Allora è possibile scambiare il limite su n e la somma su k. Brutalmente, è possibile fare il limite delle singole successioni e sommare i limiti, invece di dover fare il contrario.

$$\lim_{n\to+\infty}\sum_{k\in I} a_{k,n} = \sum_{k\in I}\lim_{n\to+\infty} a_{k,n} = \sum_{k\in I}\bar{a}_k$$

Dimostrazione Passando dalla scrittura come somma di parte positiva e negativa e utilizzando la disuguaglianza triangolare, si può facilmente mostrare che

$$\left|\sum_{k\in I} a_{k,n}\right| \leq \sum_{k\in I}|a_{k,n}|$$

Si fissa un $\varepsilon > 0$ e si mostra che, definitivamente, la differenza tra la somma dei termini n-esimi delle successioni e la somma dei limiti è minore di ε.
Dato un sottoinsieme $J \subseteq I$ finito e indicato con $J^C = I \setminus J$ il suo complementare, per proprietà associativa vale che

$$\left|\sum_{k\in I} a_{k,n} - \sum_{k\in I}\bar{a}_k\right| = \left|\sum_{k\in J} a_{k,n} - \sum_{k\in J}\bar{a}_k + \sum_{k\in J^C} a_{k,n} - \sum_{k\in J^C}\bar{a}_k\right|$$

Applicando la disuguaglianza triangolare

$$\left|\sum_{k\in I} a_{k,n} - \sum_{k\in I}\bar{a}_k\right| \leq \left|\sum_{k\in J} a_{k,n} - \sum_{k\in J}\bar{a}_k\right| + \left|\sum_{k\in J^C} a_{k,n}\right| + \left|\sum_{k\in J^C}\bar{a}_k\right|$$

I due termini a destra possono essere stimati con la somma dei b_k. Ciò è possibile perché stiamo stimando il singolo termine n-esimo di ogni successione, non l'intera successione.

$$\left|\sum_{k\in I} a_{k,n} - \sum_{k\in I}\bar{a}_k\right| \leq \left|\sum_{k\in J} a_{k,n} - \sum_{k\in J}\bar{a}_k\right| + 2\sum_{k\in J^C}|b_k| =$$

$$= \left|\sum_{k\in J} a_{k,n} - \sum_{k\in J}\bar{a}_k\right| + 2\left(\sum_{k\in I}|b_k| - \sum_{k\in J}|b_k|\right)$$

Ora, poiché la somma su famiglie di numeri positivi è definita come il sup delle somme su insiemi finiti, è possibile scegliere J_ε in modo che la differenza tra la somma totale dei $|b_k|$ e la somma parziale distino meno di $\varepsilon/3$. Allora si ha che

$$\left|\sum_{k\in I} a_{k,n} - \sum_{k\in I}\bar{a}_k\right| \leq \left|\sum_{k\in J_\varepsilon} a_{k,n} - \sum_{k\in J_\varepsilon}\bar{a}_k\right| + 2\varepsilon/3$$

Ora, poiché gli $a_{k,n}$ tendono ciascuno a $\bar{a}_k$, il limite della somma finita su J è la somma finita dei limiti. Dunque, definitivamente anche il termine di sinistra è minore di $\varepsilon/3$. Vale allora che, definitivamente,

$$\left|\sum_{k\in I} a_{k,n} - \sum_{k\in I}\bar{a}_k\right| \leq \varepsilon$$

Per definizione di limite segue la tesi.

$$\lim_{n\to+\infty}\sum_{k\in I} a_{k,n} = \sum_{k\in I}\bar{a}_k = \sum_{k\in I}\lim_{n\to+\infty} a_{k,n}$$

□

Il teorema si generalizza anche ai prodotti nella seguente forma:
Sia $(a_{k,n})_{n\in\mathbb{N},k\in I}$ una famiglia di reali positivi. Se vale che

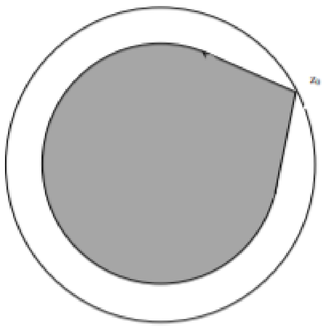
1. Per $n \to +\infty$, $a_{k,n}$ tende a $\bar{a}_k$ reale;
2. $\forall k \in I$, vale $|a_{k,n}| \leq b_k$, ossia la successione $a_{k,n}$ è maggiorata in modulo dalla costante b_k;
3. $\sum_{k\in I} |b_k| < +\infty$;

Allora vale che

$$\lim_{n\to+\infty} \prod_{k=0}^{+\infty}(1 + a_{k,n}) = \prod_{k=0}^{+\infty}(1 + \bar{a}_k)$$

21.7 Lemma di Abel per serie di potenze

Regione di Stolz Consideriamo la circonferenza nel piano complesso di raggio R centrata in zero. Sia z_0 tale che $|z_0| = R$ e sia $r < R$. Si definisce regione di Stolz di vertice z_0 e raggio r, e si indica con $S_{z_0,r}$, la regione del disco mostrata in figura, ottenuta tracciando le tangenti al cerchio di raggio r passanti per z_0.



Si osservi che il caso limite per $r = 0$ coincide con il segmento che collega l'origine e z_0

Lemma di Abel Sia $R \in (0, +\infty)$ il raggio di convergenza della serie di potenze $s(z) = \sum_{k=0}^{+\infty} a_k z^k$. Se la serie converge in un punto z_0 sul bordo del disco di convergenza, allora converge uniformemente su ogni regione di Stolz di vertice z_0.

Dimostrazione Il cambio di variabile $z = z_0 w$ trasforma la serie in $\sum_{k=0}^{+\infty} a_k z_0^k w^k$, che converge in $w = 1$ e ha raggio di convergenza 1. Dunque possiamo ridurci a considerare il caso $z_0 = R = 1$. Inoltre, a meno di modificare il coefficiente a_0, possiamo supporre $s(z_0) = s(1) = \sum_{k=0}^{+\infty} a_k = 0$. Consideriamo la regione $S_{1,r}$. Sia inoltre $E = B(0,1) \cup \{1\}$. Sia $s_n(z)$ la somma parziale della serie e sia $A_n = s_n(1) = \sum_{k=0}^{n} a_k$ la serie dei coefficienti. Applicando la sommazione per parti di Abel otteniamo, $\forall z \in E$, che

$$s_n(z) = \sum_{k=0}^{n} a_k z^k = A_n z^n + \sum_{k=0}^{n-1} A_k(z^k - z^{k+1})$$

Da cui, tenendo conto del fatto che, nelle nostre ipotesi, $|z| \leq 1$ in E e $A_n \to 0$, otteniamo

$$\sum_{k=0}^{+\infty} A_k(z^k - z^{k+1}) = \lim_{n\to+\infty} s_n(z) - A_n z^n = s(z)$$

Dunque vale che

$$|s(z) - s_n(z)| = \left| \sum_{k=0}^{+\infty} A_k(z^k - z^{k+1}) - A_n z^n - \sum_{k=0}^{n-1} A_k(z^k - z^{k+1}) \right| =$$

$$= \left| \sum_{k=n}^{+\infty} A_k(z^k - z^{k+1}) - A_n z^n \right| \leq \sum_{k=n}^{+\infty} |A_k||z^k - z^{k+1}| + |A_n|$$

Ora, scegliamo n sufficientemente grande che, per ogni $k \geq n$, $|A_n| < \varepsilon$. Possiamo allora scrivere

$$|s(z) - s_n(z)| \leq \varepsilon \left(1 + \sum_{k=n}^{+\infty} |z^k - z^{k+1}| \right) \leq \varepsilon \left(1 + \sum_{k=0}^{+\infty} |z^k - z^{k+1}| \right)$$

Dunque basta mostrare che $\sum_{k=0}^{+\infty} |z^k - z^{k+1}|$ è limitato uniformemente in $z \in S_{1,r}$. Osserviamo intanto che

$$\sum_{k=0}^{+\infty} |z^k - z^{k+1}| = |1 - z| \sum_{k=0}^{+\infty} |z^k| = \begin{cases} 0 & \text{se } z = 1 \\ \dfrac{|1-z|}{1-|z|} & \text{se } z \neq 1 \end{cases}$$

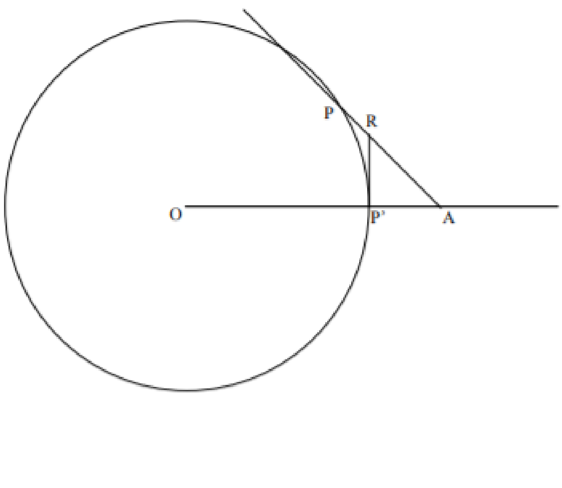
Dobbiamo mostrare che

$$\sup_{z \in S_{1,r} \setminus \{1\}} \frac{|1-z|}{1-|z|} < +\infty$$

Suddividiamo la dimostrazione in due parti. Per i punti interni al disco $D_{0,r}$ di centro l'origine e raggio r abbiamo che

$$\frac{|1-z|}{1-|z|} \leq \frac{2}{r}$$

Consideriamo ora i punti $z \in F = S_{1,r} \setminus D_{0,r}$. Identifichiamo geometricamente 1 con il punto A e z con il punto P. Tracciamo la circonferenza di centro l'origine e raggio $\overline{OP} = |z|$, che interseca l'asse reale nel punto $P' = 1 - |z|$. Tracciamo la perpendicolare all'asse reale per P', che interseca la semiretta AP in R. La situazione è riportata in figura.



Si ha $\overline{PR} \leq \overline{P'R}$ per teorema della secante. Detto $\alpha = P\hat{A}P'$ abbiamo

$$\overline{AP} = \overline{AR} + \overline{PR} \leq \overline{AR} + \overline{P'R} = \frac{\overline{AP'}}{\cos\alpha} + \overline{AP'} \tan\alpha \leq \frac{2\overline{AP'}}{\cos\alpha}$$

Dato che il punto $P = z$ è interno al cerchio, $\cos\alpha > 0$. Abbiamo allora

$$\frac{|1-z|}{1-|z|} = \frac{\overline{AP}}{\overline{AP'}} \leq \frac{2}{\cos\alpha} < +\infty$$

Come volevasi dimostrare.

$\square$

Il lemma di Abel serve a calcolare il valore di alcune serie di cui è nota la convergenza ma non il valore. Consideriamo ad esempio la serie di potenze $\sum_{k=1}^{+\infty} \frac{x^k}{k}$, che per $|x| < 1$ si identifica con la funzione $-\log(1-x)$. Il suo raggio di convergenza è 1, e per criterio di Leibniz converge nel punto $x = -1$ del bordo. Il lemma di Abel afferma che nella regione di Stolz (degenere) $[-1, 0]$ la serie converge uniformemente. Questo significa in particolare che la serie è continua in -1, dato che lo sono le somme parziali. Dunque possiamo calcolare il valore della serie armonica a segni alterni:

$$\sum_{k=1}^{+\infty} \frac{(-1)^k}{k} = \lim_{x \to -1} \sum_{k=1}^{+\infty} \frac{x^k}{k} = \lim_{x \to (-1)^-} -\log(1-x) = -\log 2$$

Si noti che non si può dire nulla sui punti del bordo in cui non è noto se la serie converga o meno. Il lemma di Abel dice solo che, se la serie converge, il limite della serie è la serie del limite.

22 Equazioni differenziali ordinarie

22.1 Riscaldamento

Dato che l'argomento è abbastanza noto da altri corsi[41] non ci dilungheremo troppo sulla teoria, limitandoci a evidenziale qualche trucchetto utile.
Iniziamo con un po' di casi semplici.

- Consideriamo un'equazione lineare di grado 1, del tipo

  $$y' = ay$$

  Possiamo procedere in due modi. Il primo è dividere per y, osservare che la quantità alla LHS diventa una derivata logaritmica e passare agli esponenti. Questo metodo, col senno di poi, è corretto, ma a priori non possiamo essere sicuri che non si stia dividendo per una cosa nulla. Allora, è meglio portare a destra y e moltiplicare da entrambe le parti per e^{-ax}. In questo modo si completa la derivata di un prodotto, e ci riconduce a un'equazione differenziale banalissima. Questo metodo prende il nome di metodo del "fattore integrante"

  $$y'e^{-ax} - aye^{-ax} = (ye^{-ax})' = 0$$

  Da cui

  $$ye^{-ax} = \text{cost} \quad \Longrightarrow \quad y = ce^{ax}$$

  Peraltro, abbiamo mostrato che l'unica soluzione possibile a questo sistema è della forma trovata, perché non abbiamo fatto alcuna assunzione sulla forma che questa dovesse avere (cosa che, invece, capita spesso).

- Passiamo alle equazioni di secondo grado. La più elementare è

  $$y'' = ay$$

  In questo caso, è comodo supporre l'esistenza di una soluzione del tipo e^{kx} per un certo k e provare a vedere se risolve l'equazione. Otteniamo

  $$k^2 e^{kx} = ae^{kt}$$

  Da cui

  $$k = \pm\sqrt{a}$$

  Dato che l'equazione è lineare, qualsiasi combinazione lineare di due soluzioni è una soluzione. Allora, una possibile soluzione è del tipo

  $$y = Ae^{kx} + Be^{-kx}$$

  Non sempre l'esponente è reale, caso in cui basta maneggiare i coefficienti A, B ed esprimere la soluzione come somma di seni e coseni:

  $$y = Ae^{ikx} + Be^{-ikx} = (A+B)\frac{e^{ikx} + e^{-ikx}}{2} + (A-B)\frac{e^{ikx} - e^{-ikx}}{2i} = C\cos(kx) + D\sin(kx)$$

  Questo metodo non assicura che non esistano altre soluzioni, cosa che si dovrà verificare a parte in altri modi.

- Il caso precedente può essere risolto anche come un sistema lineare in due incognite, introducendo una nuova funzione e riducendo tutto allo studio di funzioni lineari.

  $$\begin{cases} y' = z \\ z' = y'' = ay \end{cases}$$

  In generale, il metodo appena mostrato permette di ricondurre una qualsiasi equazione lineare a un sistema di equazioni di grado 1.

[41] Maledetti fisici.

22.2 Equazioni lineari omogenee a coefficienti costanti

Vogliamo ora risolvere un'equazione differenziale ordinaria lineare a coefficienti costanti (d'ora in avanti, EDOLCC) di grado n, priva di termine noto. In altri termini, l'equazione è del tipo

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0 y = 0$$

Osserviamo innanzitutto che la soluzione deve essere una funzione di classe C^∞. Infatti, sicuramente y è di classe C^n. Ora, esplicitando $y^{(n)}$, osserviamo che il resto della somma è derivabile almeno un'altra volta, e dunque otteniamo che anche $y^{(n)}$ è ancora derivabile con continuità. La tesi segue procedendo induttivamente.
Per risolvere l'equazione affrontiamo il problema dal punto di vista geometrico. Nello spazio vettoriale delle funzioni C^∞ su $\mathbb{C}$, la funzione "derivata" $D(y) = y'$ è un'applicazione lineare. Allora, risolvere l'equazione iniziale significa determinare il nucleo dell'applicazione lineare

$$p(D) = D^{(n)} + a_{n-1}D^{(n-1)} + \ldots + a_0 Id$$

Dove $p(z) \in \mathbb{C}[z]$ è un polinomio a coefficienti complessi[42], e con la convenzione che $D^0 = Id$. Per teorema di decomposizione primaria, se $p(z) = \prod_{k=1}^{r}(z - \lambda_k)^{n_k}$, il nucleo del polinomio calcolato in D si spezza come

$$Ker(p(D)) = \bigoplus Ker(D - \lambda_k Id)^{n_k}$$

Ciò vuol dire che la soluzione dell'equazione lineare è una combinazione lineare di soluzioni dell'equazione $(D - \lambda Id)^n y = 0$. Ora, mostriamo che l'applicazione $D - \lambda Id$ è simile all'applicazione D stessa. Infatti, vale che

$$(D - \lambda Id)y = y' - \lambda y = e^{\lambda x}(ye^{-\lambda x})'$$

Detta $E_\lambda(y) = ye^{\lambda x}$, che è ovviamente un'applicazione lineare, otteniamo

$$(D - \lambda Id) = E_\lambda D E_\lambda^{-1}$$

Da cui, elevando alla n,

$$(D - \lambda Id)^n = E_\lambda D^n E_\lambda^{-1}$$

Ora, come è noto se due applicazioni sono simili i relativi nuclei sono isomorfi tramite l'applicazione che li coniuga. Otteniamo dunque che

$$Ker(D - \lambda Id)^n = E_\lambda(KerD^n)$$

Dato che le uniche funzioni la cui derivata n-esima è nulla sono i polinomi di grado strettamente minore di n, la cui base canonica sono le potenze di x dal grado nullo al grado $n-1$, possiamo concludere che le soluzioni della sotto-equazione sono tutte e sole le funzioni della forma

$$y = p(x)e^{\lambda x} = e^{\lambda x}\sum_{i=0}^{n-1}\alpha_i x^i$$

Dunque, le soluzioni dell'equazione originaria sono della forma

$$y = \sum_{k=1}^{r} e^{\lambda_k x}\sum_{i=0}^{n_k-1}\alpha_i x^i$$

Si osservi che ci sono in tutto n parametri liberi, ossia lo spazio delle soluzioni ha dimensione pari al grado dell'equazione. Risulta facile, ma abbastanza contoso, mostrare che è possibile determinare univocamente una soluzione se si conoscono i valori delle derivate in un punto x_0, dal grado 0 al grado $n-1$.

[42] D'ora in avanti, assumeremo sempre che il polinomio sia monico.

22.3 EDOLCC non omogenee

Supponiamo ora di voler risolvere l'equazione non omogenea

$$P(D)u = f$$

Dove f è una funzione continua. Innanzitutto, l'operatore derivata D è surgettivo da C^{n+1} a C^n, mentre l'operatore M_λ di moltiplicazione per $e^{\lambda x}$ è ovviamente invertibile. Dunque, $P(D) = \prod_k (D - \lambda_k)^{n_k} = \prod_k M_{\lambda_k} D^{n_k} M_{-\lambda_k}$ è un operatore surgettivo da C^n in C^0, dove n è il grado del polinomio, perché composizione di funzioni surgettive. Quindi, esiste sempre una controimmagine di una funzione continua f, ossia il sistema è sempre risolubile.
Ora, essendo l'equazione lineare, lo spazio delle soluzioni è un traslato dello spazio delle soluzioni dell'omogenea, ossia una qualsiasi soluzione è la somma di una soluzione particolare e della soluzione generale dell'omogenea. Il procedimento appena illustrato mostra un modo "costruttivo" per ottenere una soluzione particolare: partendo da f, si continua a moltiplicare per $e^{-\lambda_k}$e a integrare un numero n_k di volte. Tuttavia, sebbene il metodo funzioni, è spesso più semplice trovare qualche trucco che bypassi la mole enorme di conti.

22.4 Problema di Cauchy (indovinate un po'? Per le EDOLCC)

La soluzione di un'equazione lineare di grado n esiste ed è unica se sono assegnati i valori delle derivate[43] della funzione in un punto x_0, dal grado 0, ossia $u(x_0)$, al grado $n-1$. In altri termini, il sistema

$$\begin{cases} u(x_0) = a_0 \\ \dots \\ u^{(n-1)}(x_0) = a_{n-1} \\ P(D)u = f \end{cases}$$

ammette una e una sola soluzione in $C^n(\mathbb{C})$.

Unicità Date due soluzioni y_1, y_2, la loro differenza $u = y_1 - y_2$ soddisfa il sistema

$$\begin{cases} u(x_0) = 0 \\ \dots \\ u^{(n-1)}(x_0) = 0 \\ P(D)u = 0 \end{cases}$$

Dunque è sufficiente mostrare che questo sistema ammette come unica soluzione la funzione identicamente nulla. Procediamo per induzione sul grado del polinomio.

- Per il caso $n = 1$ non c'è molto da dire. WLOG, sia $x_0 = 0$. Abbiamo che

$$\begin{cases} u(0) = 0 \\ u' = \lambda u \end{cases}$$

  Che diventa

$$\begin{cases} u(0) = 0 \\ u = ce^{\lambda x} \end{cases}$$

  Da cui banalmente $u \equiv 0$.

- Scomponiamo il polinomio $P(D)$ in un fattore lineare per un polinomio di grado $n-1$, su cui faremo scattare l'induzione. Otteniamo

$$\begin{cases} u(0) = 0 \\ \dots \\ u^{(n-1)}(0) = 0 \\ P(D)u = Q(D)(D - \lambda)u = Q(D)(u' - \lambda u) = 0 \end{cases}$$

[43] D'ora innanzi, si intenderà che "le prime n derivate" di una funzione in x_0 sono i valori $u(x_0), \dots, u^{(n-1)}(x_0)$.

Detta $v = u' - \lambda u$, il sistema si riscrive come

$$\begin{cases} v(0) = 0 \\ \dots \\ v^{(n-2)}(0) = 0 \\ Q(D)v = 0 \end{cases}$$

Per ipotesi induttiva, se la soluzione esiste è unica ed è $v \equiv 0$. Allora otteniamo nuovamente il caso $n = 1$, che si è già risolto.

$$\begin{cases} u(0) = 0 \\ v = u' - \lambda u = 0 \end{cases}$$

Da cui $u \equiv 0$.

Esistenza Consideriamo l'operatore che associa a ogni funzione $h \in Ker(P(D))$ il vettore $\big(h(x_0), \dots, h^{(n-1)}(x_0)\big) \in \mathbb{R}^n$ delle sue prime n derivate in x_0. Tale operatore è lineare, poiché lo è la derivata. Inoltre, per quanto dimostrato nel punto precedente, è iniettivo, dato che se due funzioni h_1 e h_2 hanno le stesse derivate la loro differenza è identicamente nulla. Dato che lo spazio di partenza e di arrivo hanno la stessa dimensione, l'operatore è anche suriettivo, ossia ogni n-upla di derivate ammette una controimmagine. In altri termini, il problema di Cauchy di un'equazione omogenea ammette una soluzione quali che siano i dati iniziali.
Per quanto riguarda la soluzione dell'equazione non omogenea, sia $g(x)$ una soluzione particolare, la cui esistenza è stata provata nel paragrafo precedente. Il discorso appena fatto induce una bigezione naturale tra $Ker(P(D)) + g(x)$ e $\mathbb{R}^n$, quella che associa a ogni $h(x)$ il vettore $\big(h(x_0) + g(x_0), \dots, h^{(n-1)}(x_0) + g^{n-1}(x_0)\big)$. Basta allora scegliere h in modo che il vettore di arrivo siano i dati iniziali del problema di Cauchy.

$\square$

22.5 Principio di Duhamel

Sia $g : I \longrightarrow \mathbb{R}$ l'unica funzione che risolve il sistema

$$\begin{cases} g(x_0) = 0 \\ \dots \\ g^{(n-2)}(x_0) = 0 \\ g^{(n-1)}(x_0) = 1 \\ P(D)g = 0 \end{cases}$$

g esiste ed è unica per quanto osservato in precedenza. Si dimostra che, scelta comunque una funzione continua $f \in C^0(I)$, il sistema

$$\begin{cases} u(x_0) = 0 \\ \dots \\ u^{(n-1)}(x_0) = 0 \\ P(D)u = f \end{cases}$$

Ammette come unica soluzione

$$u(x) = \int_{x_0}^{x} g(x + x_0 - t) f(t) dt$$

Dimostrazione WLOG, studieremo il caso $x_0 = 0$. Premettiamo il seguente lemma, di cui faremo uso in seguito.

Lemma Sia $H(x,t)$ una funzione derivabile. Si ha che

$$\frac{d}{dx}\int_0^x H(x,t)dt = H(x,x) + \int_0^x \frac{\partial}{\partial x}H(x,t)dt$$

Ossia, la derivata di un integrale che dipende da x sia nell'integrando che negli estremi di integrazione si ottiene sommando l'integrando calcolato in $t = x$, come se l'integrando non dipendesse da x, e l'integrale della derivata parziale, come se gli estremi non dipendessero da x.
Dimostriamo il lemma in un caso particolare, sufficiente a permetterci di applicarlo al teorema. Mostriamo innanzitutto che vale per $H(x,t) = a(x)b(t)$. Infatti, dato che a x fissato la funzione $a(x)$ è un parametro, che si può portare fuori e dentro dall'integrale senza problemi, otteniamo

$$\frac{d}{dx}\int_0^x a(x)b(t)dt = \frac{d}{dx}\left(a(x)\int_0^x b(t)dt\right) = a'(x)\int_0^x b(t)dt + a(x)b(x) = \int_0^x a'(x)b(t)dt + a(x)b(x)$$

Ora, dato che la derivata è una relazione lineare, la formula vale per ogni combinazione lineare del tipo $H(x,t) = \sum_i a_i(x)b_i(t)$. Questo è sufficiente, perché nel caso che tratteremo ora l'integrando è della forma

$$H(x,t) = g(x-t)f(t) = f(t)\sum_{k=0}^{n-1} p_k(x-t)e^{\lambda_k(x-t)}$$

Per certi polinomi $p_k(x-t)$. Ciò segue dal fatto che g risolve un sistema lineare omogeneo, e dunque è una combinazione lineare di polinomi per esponenziali. Si osserva facilmente che la scrittura sopra si decompone in una somma finita di prodotti di termini in cui compare solo x o solo t, ed è dunque possibile applicare il lemma.
Ora, per far vedere che la soluzione è della forma cercata imponiamo $u(x) = \int_0^x g(x-t)f(t)dt$ e verifichiamo che risolve il sistema. Innanzitutto calcoliamo le derivate di u fino al grado n.

$$\frac{du}{dx} = \frac{d}{dx}\int_0^x g(x-t)f(t)dt$$

Per il lemma precedente, la derivata si spezza come

$$\frac{du}{dx} = \int_0^x g'(x-t)f(t)dt + g(0)f(x)$$

Il secondo termine è nullo, dato che lo è $g(0)$ per ipotesi. Allora la derivata diventa

$$\frac{du}{dx} = \int_0^x g'(x-t)f(t)dt$$

Similmente, le derivate fino alla $n-1$-esima si esprimono come

$$\frac{d^k u}{dx^k} = \int_0^x g^{(k)}(x-t)f(t)dt$$

L'unico caso in cui il secondo termine non si annulla è quello in cui $k = n$, dato che $g^{(n)}(0) = 1$ per ipotesi. Abbiamo allora che

$$\frac{d^n u}{dx^n} = \int_0^x g^{(n)}(x-t)f(t)dt + f(x)$$

Iniziamo a osservare che le prime n derivate della funzione u sono nulle in $x = 0$, dato che si sta integrando da 0 a 0, e che quindi la funzione u rispetta le condizioni del problema di Cauchy. Ora, sommiamo le derivate di u in modo da comporre il polinomio $P(D)u = u^{(n)}\sum_{k=0}^{n-1} a_k u^{(k)}$. Otteniamo allora che

$$P(D)u = \int_0^x \left(g^{(n)} + \sum_{k=0}^{n-1} a_k g^{(k)}\right)(x-t)f(t)dt + f(x) = \int_0^x (P(D)g)(x-t)f(t)dt + f(x)$$

Per ipotesi, $P(D)g \equiv 0$. Allora otteniamo esattamente l'equazione iniziale

$$P(D)u = f(x)$$

$\square$

22.6 Equazioni lineari alle differenze finite

Consideriamo lo spazio vettoriale delle successioni bilatere $V = \mathbb{C}^{\mathbb{Z}}$. Supponiamo di voler esplicitare analiticamente una successione definita per ricorrenza dalla formula

$$x(n+m) + \alpha_{m-1}x(n+m-1) + \ldots + \alpha_0 x(n) = 0$$

L'equazione sopra si dice equazione alle differenze finite. In particolare è lineare, omogenea e a coefficienti costanti. Ci aspettiamo dunque che il metodo risolutivo ricalchi quello mostrato in 22.2 per le equazioni differenziali dello stesso tipo, dove l'operatore derivata è sostituito dallo "shift".

$$\begin{aligned} S: & V && \longrightarrow V \\ & x(n) && \longrightarrow x(n+1) \end{aligned}$$

L'operatore associa a una successione la successione "scalata" di 1 all'indietro. Possiamo allora rileggere l'equazione come

$$p(S)x(n) = 0$$

Per un certo polinomio $p = \prod_{k=1}^{r}(z - \lambda_k)^{n_k} \in \mathbb{C}[x]$. Per teorema di decomposizione primaria, l'insieme delle soluzioni, ossia il nucleo dell'applicazione lineare, si spezza come

$$Kerp(S) = Ker \prod_{k=1}^{r}(S - \lambda_k Id)^{n_k} = \bigoplus Ker(S - \lambda_k Id)^{n_k}$$

Ora, una possibile soluzione dell'equazione $(S - \lambda Id)^m x_n$ è data dalle successioni $x_n = n^j \lambda^n$, per j che varia tra 0 e $m-1$.[44] Infatti, sia $n^j = p_j(n)$, inteso come polinomio in n di grado j. Se applichiamo l'operatore alla soluzione otteniamo

$$(S - \lambda Id)x_n = p_j(n+1)\lambda^{n+1} - p_j(n)\lambda^{n+1} = \lambda^{n+1}p_{j-1}(n)$$

Dove $p_{j-1}(n)$ è un polinomio. Mostriamo che esso deve avere grado inferiore a j, cosa che giustifica la notazione $p_{j-1}(n)$. Ora, qualsiasi sia $p_j(n) = a_n n^j + \ldots$, la differenza tra i valori che assume in due interi successivi vale: $p_{j-1}(n) = a_n(n+1)^j + \ldots - a_n n^j + \ldots = a_n\left(\sum_{k=0}^{j-1}(n+1)^k n^{j-k-1}\right) + \ldots$. Dunque, procedendo induttivamente, si scopre che applicando $j+1$ volte l'operatore la successione si annulla, per motivi di grado. Allora, dato che $m \geq j+1$, la successione costruita è soluzione. Osserviamo ora che le soluzioni trovate sono tra loro linearmente indipendenti. Per farlo, basta far vedere che sono indipendenti le successioni $y_j(n) = n^j$ al variare di $j \in \mathbb{N}$. Per assurdo, non lo siano. Allora esiste una combinazione lineare a coefficienti non nulli tale che

$$a_n n^k + \ldots + a_0 = 0$$

Abbiamo dunque un polinomio in n non identicamente nullo, che si annulla però su ogni $n \in \mathbb{N}$, cosa ovviamente impossibile.
Quanto visto permette di concludere, per dimensionalità, che tutte e sole le soluzioni sono generate da combinazioni lineari dei termini dati.

22.7 Equazioni lineari del primo ordine

Studiamo ora un'equazione del tipo

$$y'(x) = a(x)y(x) + b(x)$$

dove $a(x)$, $b(x)$ sono generiche funzioni continue. Possiamo procedere in due modi

[44] Mi scuso se sembra calata dal cielo, ma di fatto è così che matematici molto più irreprensibili del sottoscritto presentano l'argomento.

Soluzione dell'omogenea e variazione delle costanti. Risolviamo prima l'equazione omogenea associata

$$y'(x) = a(x)y(x)$$

Separiamo le variabili, cioè scriviamo $y' = dy/dx$[45] e integriamo da ambo le parti.

$$\frac{dy}{y} = a(x)dx$$

$$\log |y| = C + \int a(x)dx$$

$$y = ke^{\int a(x)dx}$$

Per una certa costante k.
Adottiamo ora il cosiddetto "metodo di variazione delle costanti". Supponiamo, cioè, che $k = k(x)$ sia una funzione da determinare, e sostituiamo la soluzione trovata nell'equazione originaria.

$$y'(x) = k(x)a(x)e^{\int a(x)dx} + k'(x)e^{\int a(x)dx} = k(x)a(x)e^{\int a(x)dx} + b(x)$$

Da cui

$$k(x) = C + \int e^{-\int a(x)dx}b(x)dx$$

Dunque

$$y = e^{\int a(x)dx}\left(C + \int e^{-\int a(x)dx}b(x)dx\right)$$

Metodo del fattore integrante Riscriviamo l'equazione come

$$y'(x) - a(x)y(x) = b(x)$$

Moltiplichiamo da ambo le parti per $e^{-\int a(x)dx}$, in modo da completare la derivata del prodotto a sinistra.

$$\left(y(x)e^{-\int a(x)dx}\right)' = e^{-\int a(x)dx}b(x)$$

Da cui

$$y(x) = e^{\int a(x)dx}\left(C + \int e^{-\int a(x)dx}b(x)\right)$$

Che fortunatamente è la stessa soluzione di prima.

22.8 (*)Equazione di Riccati

Consideriamo un'equazione differenziale del primo ordine del tipo

$$u' = a(t) + b(t)u + c(t)u^2$$

Per determinare una soluzione, supponiamo che la coppia $(x(t), y(t))$ risolva il sistema

$$\begin{cases} x' = a(t)x + b(t)y \\ y' = c(t)x + d(t)y \end{cases}$$

Una semplice verifica permette di dire che la funzione $u = x/y$ soddisfa l'equazione di Riccati

$$u' = b(t) + (a(t) - d(t))u - c(t)u^2$$

In realtà questo sposta solo il problema, dato che determinare le funzioni x e y è generalmente poco facile. Infatti, il vettore (x, y) è soluzione del sistema differenziale vettoriale

$$\mathbf{x}' = A(t)\mathbf{x}$$

[45] Sembra una porcata, ma ha un fondamento matematico serio. Tuttavia, sfogliare la dimostrazione di quest'ultimo è quasi più doloroso di vedere qualche dx a spasso per il foglio, dunque mi permetto di non darla.

Dove $A(t)$ è una matrice 2×2. Per risolvere questo sistema si dovrebbe parlare di esponenziazione di matrici, e l'autore decide di fermarsi qui.

Molti sono gli ostacoli che impediscono di sapere, sia l'oscurità dell'argomento sia la brevità della vita umana.

Protagora sulle EDO.

22.9 Come sfondare le equazioni differenziali in 16 minuti

Ragazzi, sono veramente euforico.

Io quando risolvo una EDO a coefficienti non costanti.

In generale, le equazioni differenziali non si sanno risolvere. Esistono allora alcuni trucchi sempreverdi cui aggrapparsi come ultima speranza.

- Si può provare a supporre che l'equazione sia una serie di potenze, e verificare poi che il raggio di convergenza sia non nullo. Consideriamo ad esempio il seguente problema di Cauchy:

$$\begin{cases} u'' = tu \\ u(0) = 1 \\ u'(0) = 0 \end{cases}$$

Supponiamo che la soluzione sia della forma $u(t) = \sum_{k=0}^{+\infty} a_k t^k$. Sostituendo e derivando termine a termine (supponendo ovviamente che il punto t sia interno al raggio di convergenza della serie) otteniamo

$$u'' = \sum_{k=2}^{+\infty} k(k-1)a_k t^{k-2} = tu = \sum_{k=0}^{+\infty} a_k t^{k+1}$$

Eguagliando termine a termine i coefficienti si ottiene

$$\begin{cases} a_0 = u(0) = 1 \\ a_1 = u'(0) = 0 \\ a_2 = 0 \\ a_{k+3} = \frac{a_k}{(k+2)(k+3)} \end{cases}$$

La serie che ne deriva ha sicuramente raggio di convergenza non nullo, dato che i coefficienti sono minori di $1/k$ e la serie $\sum x^k/k$ ha raggio di convergenza 1.

- Se siete skillati, potete provare a chiedere a Wolfram Mathematica o simili di risolvere numericamente l'equazione. Cosa che è un po' barare, ma tutto fa brodo.

23 (*)Esercizi e altre amenità irrilevanti

23.1 Formica su una banda elastica

Una formica si muove su una banda elastica di lunghezza iniziale 1 m, a velocità costante. Ogni secondo, la banda si allunga di 1 m. Dimostrare che la formica riesce, in un tempo finito, a uscire dalla banda.

Dimostrazione Quando la banda si allunga, la frazione del cammino già percorso rimane la stessa. L'obiettivo della formica è fare sì che la somma di tali frazioni sia prima o poi maggiore di 1.
Al tempo t=n, la formica percorre una frazione della banda pari a $1/n$, perhcé la banda è lunga n metri e ne percorre 1. Allora la frazione totale della banda percorsa è $\sum_n 1/n$, che diverge.

□

23.2 $\prod_{\text{p primo}} \left(1 - \frac{1}{p^k}\right)^{-1}$

Dato un numero reale positivo k, vale che

$$\prod_{\text{p primo}} \frac{1}{1 - \frac{1}{p^k}} = \sum_{n \in \mathbb{N} \setminus \{0\}} \frac{1}{n^k}$$

Dimostrazione Consideriamo $\frac{1}{1-\frac{1}{p^k}}$ come la somma di una serie geometrica di ragione $\frac{1}{p^k}$.
Otteniamo allora che

$$\prod_{\text{p primo}} \frac{1}{1 - \frac{1}{p^k}} = \prod_{p \text{ primo}} \sum_{n \in \mathbb{N}} \frac{1}{p^k}$$

Ora, eseguendo i prodotti tra i termini delle sommatorie si ottiene una e una sola volta il reciproco di ogni numero naturale non nullo, scegliendo i termini della sua fattorizzazione in irriducibili. Segue la tesi.

□

23.3 Postulato di Bertrand

$\forall n \in \mathbb{N}$, esiste un numero primo compreso tra n e 2n.

Dimostrazione La prima dimostrazione di questo enunciato si deve a Chebyshev. Daremo qui la dimostrazione di Erdös.
Consideriamo la fattorizzazione in primi del binomiale centrale $\binom{2n}{n}$. Raccogliamo un po' di fatti sulla sua possibile fattorizzazione in primi:

1. É noto che la somma dei binomiali con lo stesso apice k fa 2^k. Allora, con $k = 2n$, otteniamo

   $$4^n = \sum_{i=0}^{2n} \binom{2n}{i}$$

   Dato che il termine più grande della sommatoria è il termine centrale, se $n > 0$ possiamo scrivere

   $$4^n = \sum_{i=0}^{2n} \binom{2n}{i} \leq 2n \binom{2n}{n}$$

   Si è preso il binomiale centrale $2n$ volte, e non $2n + 1$ come si sarebbe dovuto, perché tanto $\binom{2n}{0} = \binom{2n}{2n} = 1$ e il termine centrale è sicuramente maggiore di 2. Perciò vale

   $$\frac{4^n}{2n} \leq \binom{2n}{n}$$

2. Per quanto mostrato in 2.3, se un primo divide il binomiale centrale vale che

$$p^k \mid \binom{2n}{n} \implies p^k \leq 2n$$

3. Come caso particolare del fatto precedente, se $p > \sqrt{2n}$, allora p^2 non divide $\binom{2n}{n}$. Ossia, se ci sono fattori primi oltre $\sqrt{2n}$, essi devono avere molteplicità singola.

4. Non ci sono fattori primi nella fascia $2n/3 < p \leq n$. Infatti, se si esplicita il binomiale col teorema dei tre fattoriali, si ottiene:

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2}$$

   Il fattore p può comparire al massimo due volte al numeratore, dato che $2n < 3p$. Inoltre, sicuramente compare due volte al denominatore, essendo $p \leq n$. Dunque si cancella e non può dividere il binomiale.

5. La funzione primoriale di n, definita come

$$n\# = \prod_{p \leq n \text{ primo}} p$$

   è maggiorata da 4^n
   Per dimostrarlo, consideriamo il prodotto tra i primi compresi tra un certo $k/2$ non incluso e k incluso, per k intero. Questo prodotto divide il binomiale $\binom{k}{k/2}$ (o eventualmente $\binom{k}{(k-1)/2}$ se k è dispari). Infatti, ogni primo compare solo al numeratore e non al denominatore. Allora

$$\prod_{k/2<p\leq k,\ p \text{ primo}} p \mid \binom{k}{k/2} \leq \sum_{i=0}^{k} \binom{k}{i} = 2^k$$

   Ora, fissato n, si può spezzare il prodotto dei primi mnori di lui in una quantità finita di sottoprodotti della forma

$$n\# = \prod_{p\leq n \text{ primo}} p = \left(\prod_{n/2<p\leq n,\ p \text{ primo}} p \right) \left(\prod_{n/4<p\leq n/2,\ p \text{ primo}} p \right) \dots$$

   Maggiorando ciascun prodotto parziale otteniamo

$$n\# \leq 2^{n+n/2+\dots} = 2^{2n} = 4^n$$

Veniamo al sodo. Mostreremo che, se per assurdo non esistessero primi tra n e $2n$ non basterebbe il prodotto di tutti i primi inferiori per raggiungere il binomiale $\binom{2n}{n}$. É tautologico dire che

$$\binom{2n}{n} = \prod_{p \mid \binom{2n}{n}} p^r$$

Si osserva che, per il fatto 2, $p^r \leq 2n$. Ora, spezziamo il prodotto in quattro prodotti parziali.

$$\binom{2n}{n} = \left(\prod_{p\leq\sqrt{2n}} p^r \right) \left(\prod_{\sqrt{2n}<p\leq 2n/3} p^r \right) \left(\prod_{2n/3<p\leq n} p^r \right) \left(\prod_{n<p\leq 2n} p^r \right)$$

Il prodotto dei primi minori o uguali di $\sqrt{2n}$ è al massimo $\prod_{k\leq\sqrt{2n}} 2n = (2n)^{\sqrt{2n}}$, anche supponendo che ogni naturale minore di $\sqrt{2n}$ sia primo. Il prodotto dei primi tra $\sqrt{2n}$ escluso e $2n/3$ è composto solo da termini di molteplicità 1 per il fatto 4. Allora è maggiorato dal primoriale di $2n/3$, a sua volta maggiorato da $4^{2n/3}$ per il fatto 5. Per il fatto 3, non vi è nessun primo tra $2n/3$ e n, e per assurdo non c'è nessun primo tra n e $2n$. Allora:

$$\binom{2n}{n} \leq (2n)^{\sqrt{2n}} 4^{2n/3}$$

Per fatto 1, sappiamo che

$$\frac{4^n}{2n} \leq \binom{2n}{n}$$

Troveremo l'assurdo nel supporre che sia sempre vero che

$$\frac{4^n}{2n} \leq (2n)^{\sqrt{2n}} 4^{2n/3}$$

Passando ai logaritmi, troviamo che

$$n \log 4 - \log 2n \leq \sqrt{2n} \log 2n + \frac{2n}{3} \log 4$$

$$\frac{n \log 4}{3} - \log 2n(1 + \sqrt{2n}) \leq 0$$

Consideriamo la funzione

$$f(x) = \frac{x \log 4}{3} - \log 2x(1 + \sqrt{2x})$$

E chiediamoci quando sia negativa. Osserviamo che è convessa, perché $y = x$ è convessa e $y = -\log 2x(1 + \sqrt{2x})$ è convessa essendo l'opposta di una funzione concava. Ciò significa che, se da un certo punto in poi è positiva, lo è definitivamente, perché la sua derivata è crescente. Con un po' di sani contazzi si riesce a mostrare che l'ultimo zero della funzione si trova tra 467 e 468. Dunque, se n è maggiore di 468 si cade in un assurdo, dato che la funzione è positiva. Per $n < 468$, il postulato di Bertrand può essere mostrato a mano: basta osservare che ogni primo minore di 468 è minore del doppio del precedente, cosa che è equivalente alla tesi che tra n e $2n$ ci sia un primo.[46]

□

23.4 Algoritmo di Erone per la radice quadrata

Dato un numero reale positivo $a > 0$, si definisce per ricorsione la successione

$$\begin{cases} x_0 = 1 \\ x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \end{cases}$$

Si dimostra che la successione converge a $\sqrt{a}$, cosa che può essere considerata una dimostrazione dell'esistenza della radice n-esima di un numero reale in $\mathbb{R}$.

Dimostrazione Mostriamo innanzitutto che, da $n = 1$ in poi, $\sqrt{a} \leq x_n \leq x_{n+1}$. Si procede per induzione su n.

- Per $n = 1$, $x_1 = \frac{a+1}{2}$. Per disuguaglianza delle medie si ha che $x_1 > \sqrt{a}$.
- Se $x_n \geq \sqrt{a}$, allora

  $$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \geq \sqrt{a}$$

  Che conduce a

  $$x_n{}^2 + a - 2\sqrt{a}x_n = (x_n - \sqrt{a})^2 \geq 0$$

[46] Il lettore a questo punto manderà a fanculo l'autore, dicendo che se avesse saputo di doversi fare i casi a mano avrebbe di certo scelto un modo più intelligente di passare il suo tempo. A questa manifestazione di rabbia l'autore risponde facendo spallucce.

Tale disequazione è verificata per ogni $x \neq \sqrt{a}$. Per l'altra disuguaglianza si osserva che

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \leq x_n \quad \longrightarrow \quad {x_n}^2 + a \leq 2{x_n}^2$$

Che è vero perché $x_n \geq \sqrt{a}$.

Ora, la monotonia della successione assicura che il limite esiste in $\mathbb{R}$. Resta da mostrare che tale limite è esattamente $\sqrt{a}$. Per fare ciò, consideriamo la funzione $f(x) = \frac{1}{2}\left(x + \frac{a}{x}\right)$. Mostriamo che tale funzione è una contrazione nell'intervallo $[\sqrt{a}, x_1]$. Infatti, se ne consideri la derivata $f'(x) = \frac{1}{2}\left(1 - \frac{a}{x^2}\right)$. Essa è strettamente crescente per $x > 0$ e tende a $1/2$ all'infinito. Inoltre, in $x = \sqrt{a}$ la derivata è nulla. Ciò significa che la derivata è minore di 1 in ogni punto dell'intervallo $[\sqrt{a}, x_1]$, e dunque la funzione è una contrazione. Per teorema delle contrazioni, il limite della successione deve essere un punto fisso della funzione $f(x)$. Dunque, detto L il limite, si ha che $L = f(L) = \frac{1}{2}\left(L + \frac{a}{L}\right)$, e si è già visto che tale equazione ha soluzione per $L = \sqrt{a}$.

□

23.5 Algoritmo di Erone generalizzato

Procedendo in maniera analoga al caso $n = 2$ è possibile mostrare che la seguente successione tende a $\sqrt[n]{a}$:

$$\begin{cases} x_0 > a \\ x_{k+1} = \frac{n-1}{n}\left(x_k + \frac{a}{(n-1)x_k^{n-1}}\right) \end{cases}$$

Dimostrazione

- Osserviamo innanzitutto che la successione è sempre maggiore di $\sqrt[n]{a}$. Procedendo per induzione, se $x_k > \sqrt[n]{a}$ allora

  $$x_{k+1} = \frac{n-1}{n}\left(x_k + \frac{a}{(n-1)x_k^{n-1}}\right) = \frac{1}{n}\left(x_k + \ldots + x_k + \frac{a}{x_n^{n-1}}\right) \geq \sqrt[n]{a}$$

  Dove l'ultima disuguaglianza è la disuguaglianza delle medie.

- Mostriamo che x_k è decrescente. Infatti

  $$x_{k+1} = \frac{n-1}{n}\left(x_k + \frac{a}{(n-1)x_k^{n-1}}\right) < x_k$$

  Moltiplicando per x_k^{n-1}, si ottiene

  $$\frac{a}{n} < \frac{1}{n}x_k^n$$

  Che è vero essendo $x_k > \sqrt[n]{a}$. Dunque il limite della successione esiste in $\mathbb{R}$.

- Per teorema delle contrazioni, reiterando le osservazioni già fatte per il caso $n = 2$, segue che il limite deve essere un punto fisso della funzione $f(x) = \frac{n-1}{n}\left(x + \frac{a}{(n-1)x^{n-1}}\right)$, che si verifica essere il punto $\bar{x} = \sqrt[n]{a}$.

□

23.6 Teorema di Borel

Per ora questo teorema è stato solo citato, ma sai mai.
Data una successione di reali $(a_n)_{n\in\mathbb{N}}$, esiste una funzione di classe C^∞ la cui derivata n-esima in zero vale a_n.

23.7 Sviluppo asintotico di $c = x + \log x$

Cerchiamo di risolvere l'equazione $c = x + \log x$. Osserviamo fin da subito che, essendo entrambe le funzioni strettamente crescenti e continue, il valore c verrà assunto una e una sola volta. Consideriamo il caso in cui $c \gg 1$. Sicuramente si dovrà avere $x \geq 1$, e dunque il logaritmo sarà positivo. Ciò implica che $c \geq x$, e per monotonia del logaritmo $\log c \geq \log x$. Dunque si giunge a

$$c - \log c \leq x - \log x = x \leq c \quad \longrightarrow \quad x = c + \mathcal{O}(\log c)$$

Dove $\mathcal{O}$ è la "O grande" di Landau. Inseriamo questa prima soluzione nell'equazione originaria e continuiamo a sviluppare.

$$x = c - \log x = c - \log(c + \mathcal{O}(\log c)) = c - \log\left(c\left(1 + \mathcal{O}\left(\frac{\log c}{c}\right)\right)\right) = c - \log c - \log\left(1 + \mathcal{O}\left(\frac{\log c}{c}\right)\right)$$

Essendo $c \gg \log c$ asintoticamente, l'argomento del logaritmo tende a 1, e dunque va come $\mathcal{O}\left(\frac{\log c}{c}\right) + o\left(\mathcal{O}\left(\frac{\log c}{c}\right)\right) = \mathcal{O}\left(\frac{\log c}{c}\right)$. Allora si ha che

$$x = c - \log c - \log\left(1 + \mathcal{O}\left(\frac{\log c}{c}\right)\right) = c - \log c - \mathcal{O}\left(\frac{\log c}{c}\right)$$

Inseriamo nuovamente la soluzione nell'equazione e continuiamo ad affinare questo conto infinito.

$$x = c - \log x = c - \log\left(c - \log c - \mathcal{O}\left(\frac{\log c}{c}\right)\right) = c - \log c - \log\left(1 - \frac{\log c}{c} - \mathcal{O}\left(\frac{\log c}{c^2}\right)\right)$$

Come prima, usiamo che $\log(1 + x) = x + o(x)$, ottenendo

$$x = c - \log c - \log\left(1 - \frac{\log c}{c} - \mathcal{O}\left(\frac{\log c}{c^2}\right)\right) = c - \log c + \frac{\log c}{c} + o\left(\frac{\log c}{c}\right)$$

E ci fermiamo qui, per carità.

23.8 Formula di inversione di Moebius

Se una funzione $f(n)$ sui naturali è espressa come somma sui divisori di n, vale la formula seguente:

$$f(n) = \sum_{d|n} g(d) \quad \iff \quad g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) \tag{26}$$

Dove $\mu(d)$ è la funzione di Moebius, definita come

$$\mu(d) = \begin{cases} (-1)^n & \text{se } d \text{ è il prodotto di } n \text{ primi distinti;} \\ 0 & \text{altrimenti} \end{cases}$$

23.9 Funzione a spuntoni

Consideriamo la funzione

$$f_0(x) = \min\{2x, 2 - 2x\} \quad \text{per } 0 \leq x \leq 1$$

Tale funzione è uno "spuntone" che ha un picco in $1/2$. Componendo la funzione con se stessa, si ottiene la successione di funzioni così definita:

$$\begin{cases} f_0(x) = \min\{2x, 2 - 2x\} \quad \text{per } 0 \leq x \leq 1 \\ f_{n+1} = f(f_n) = \begin{cases} f_n(2x) & sex \leq 1/2 \\ f_n(2 - 2x) & sex > 1/2 \end{cases} \end{cases}$$

f_n è lineare in ciascun intervallo del tipo $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$, in cui assume tutti i valori tra 0 e 1 per teorema dei valori intermedi. Perciò, la funzione f_n ha 2^n punti fissi, cioè la funzione originaria ammette 2^n

punti periodici di periodo $d|n$. In altri termini, esistono orbite periodiche della funzione, e dunque il trucchetto di reiterare f per trovare il limite di una successione può portare a successioni prive di limite. Questo perché la nostra funzione non è né monotona strettamente crescente, né una contrazione.
Ci si chiede ora quanti siano i punti di periodo esattamente n. Sia $P(n)$ tale numero. I punti fissi di f_n sono in totale 2^n, e ognuno di essi ha periodo $d|n$ per qualche d. Vale allora che

$$2^n = \sum_{d|n} P(d)$$

Applicando l'inversione di Moebius si ottiene

$$P(n) = \sum_{d|n} \mu(d) 2^{n/d}$$

Dividendo per n si ottiene il numero di orbite di periodo n, in quanto in ogni orbita ci sono n punti di periodo n.

23.10 Curva del Pudding

Si mostra che esiste una curva continua su [0,1] ma non derivabile in nessun punto e di lunghezza infinita in ogni sottointervallo. Tale curva è nota come "blancmange curve" per la sua somiglianza con l'omonimo pudding.

Definizione Definiamo per ricorrenza la successione di funzioni

$$\begin{cases} u_1 = d(x, \mathbb{Z}) \text{ (è la distanza di } x \text{ dagli interi);} \\ u_{k+1} = \frac{u_1(2^k x)}{2^k} \end{cases}$$

La funzione u_{k+1} è una contrazione di u_1 in scala $1 : 2^k$. Inoltre, dato che u_1 è 1-Lipschitziana e la composizione di funzioni lipschitziane è lipschitziana di costante il prodotto delle costanti, anche u_k è 1-Lipschitziana, dato che $2^k x$ e $\frac{x}{2^k}$ sono lipschitziane di costante l'una l'inversa dell'altra.
Definiamo la seguente funzione frattale:

$$f_n(x) = \sum_{k=1}^{n} u_k(x), \quad x \in [0,1]$$

$$f(x) = \lim_{n\to+\infty} f_n(x) = \sum_{k=0}^{+\infty} \frac{u_1(2^k x)}{2^k}, \quad x \in [0,1]$$

In altri termini, f risolve l'equazione funzionale[47]

$$f(x) = u(x) + \frac{1}{2} f(2x)$$

Osserviamo subito che la funzione è ben definita, perché dato che $|u_k(x)| \leq 2^{-k+1}$, il valore assunto dalla funzione è al più $\sum_{k=0}^{+\infty} 2^{-k} = 2$.
La funzione è inoltre continua, perché limite uniforme di funzioni continue. Consideriamo ora la derivata della funzione u_k nei punti in cui è definita. Vale che

$$u_k'(x) = (-1)^{\lfloor 2^k x \rfloor} \text{ se } x \neq n2^k, \quad n \in \mathbb{Z}$$

Si osserva che $u_k'(x) = 1$ se $x \neq n2^{-k}$ e nella scrittura in base 2 di x la cifra k-esima dopo la virgola è 0; viceversa, $u_k'(x) = -1$ se $x \neq n2^{-k}$ e nella scrittura in base 2 di x la cifra k-esima dopo la virgola è 1. Osserviamo inoltre che, nei punti in cui è derivabile, f_n ha derivata della stessa parità di n, dato che è la somma di un numero n di 1 e -1 Ci si chiede ora su quanti intervalli

[47]Peraltro, è l'unica soluzione limitata.

la somma parziale f_{2n} sia piatta, ossia dove la somma delle derivate delle prime $2n$ funzioni si annulla. Si osserva che ogni tratto piatto è univocamente identificato da una $2n$-upla di 1 e -1 in pari numero, che indicano i valori delle derivate delle singole funzioni u_k, nell'ordine. Il numero di tali $2n$-uple è dato dalle permutazioni di $2n$ oggetti, di cui n e n ripetuti, ossia da $\binom{2n}{n}$. Ogni tratto ha ampiezza $L = 2^{-2n}$, dato che ci sono 2^{2n} punti in cui la funzione u_k ha derivata non definita. Ora, definiamo $P_k = \{n2^{-k}\}_{0\leq n\leq 2^k}$. Dato che la curva è continua, la sua lunghezza può essere calcolata come limite delle lunghezze delle spezzate di vertici P_k, essendo i numeri della forma $n2^{-k}$ densi in $[0, 1]$. Consideriamo di quanto aumenta la lunghezza della curva (che in $\mathbb{R}$ coincide con la somma in modulo delle variazioni della curva, cioè di quanto sale e scende) passando da f_{2m} a f_{2m+1}. Nei tratti in cui f_{2m} era crescente, la derivata era almeno 2, dovendo avere la stessa parità di $2m$. Allora f_{2m+1} è ancora crescente in quei tratti, dato che la derivata decresce al massimo di 1. Stessa cosa per i tratti in cui la funzione era decrescente. Ciò significa che nei tratti in cui f_{2m} era monotona la variazione totale della curva rimane invariata, e dunque non vi è alcun aumento della lunghezza. Invece, in ogni tratto piatto la lunghezza aumenta della stessa dimensione $L = 2^{-2m}$ del tratto, dato che esso viene sostituito da una cuspide di altezza $L/2$. Dunque, l'aumento totale della lunghezza è pari al numero di tratti piatti per la loro lunghezza.

$$\Delta L = \binom{2m}{m} 2^{-2m}$$

Sviluppando il binomiale con l'approssimazione di cui alla sezione 2.4, otteniamo

$$\Delta L \geq \frac{2^{2m}}{\sqrt{\pi m(1 + 1/2)}} 2^{-2m} = \frac{1}{\sqrt{\pi m(1 + 1/2)}}$$

Che è il termine generale di una serie divergente. Si può inoltre mostrare che la lunghezza rimane invariata passando da f_{2m+1} a f_{2m+2}, sempre per motivi di parità della derivata (la funzione rimane crescente, in senso debole, dov'era crescente).
Abbiamo allora trovato che $\lim_{n\to+\infty} L(f_n, [0, 1]) = +\infty$. In generale ciò non basta ancora per concludere che anche la funzione limite ha lunghezza infinita. Tuttavia, per definizione di lunghezza, $L(f, [0, 1]) \geq L(f, P_m) = \lim_{n\to+\infty} L(f_n, P_m)$. Ora, se si fissa m, dopo che n ha raggiunto m la lunghezza della curva su P_m rimane costante, dato che da lì in poi la funzione f_n è sempre nulla sui punti di P_m per costruzione, e dunque le immagini di tali punti rimangono fisse. Dunque $\lim_{n\to+\infty} L(f_n, P_m) = L(f_m, P_m) = L(f_m, [0, 1])$ essendo f_m lineare a tratti. Poiché vale che $L(f, [0, 1]) \geq L(f_m, [0, 1])$ $\forall m \in \mathbb{N}$, la curva ha lunghezza infinita.
Ora, si può facilmente mostrare che la curva contiene una copia omotetica, ossia "riscalata", di se stessa in ogni sottointervallo. Cio segue dal fatto che, ogniqualvolta la funzione ha una tratto piatto, si genera una copia della funzione stessa, e i tratti piatti sono evidentemente densi in $[0, 1]$. Ciò significa che in ogni sottointervallo la curva ha lunghezza infinita. Da questa considerazione segue che la curva non è lipschitziana in nessun punto, dato che non ha lunghezza finita.

□

23.11 Randomiche osservazioni sulla funzione precedente

[48]Per mostrare che la funzione non è derivabile in nessun punto, si osserva che, se si considera la scrittura in base 2 di un punto x, a essa è moralmente associata una somma infinita di 1 e -1, che corrisponde al limite destro della derivata della funzione in x. Tale somma può solo divergere o oscillare, e in entrambi i casi la derivata destra non è finita.
Se si vuole calcolare il massimo della funzione, basta osservare che il massimo della somma parziale f_{2m} si ottiene nei tratti piatti ottenuti alternando derivate positive e negative. Infatti, per $m = 1$ la tesi è ovvia. Per il passo induttivo si osserva che, se il massimo della funzione precedente si ottiene sul tratto piatto ottenuto alternando derivate negative, il massimo di f_{2m+2} si ottiene alzando dal tratto piatto della funzione precedente un nuovo tratto piatto, ottenuto con una derivata positiva e una negativa. Allora, il massimo di f si ottiene nei punti irrazionali che si ottengono alternando

[48]Da qui in poi i concetti saranno un po' vaghi, e me ne scuso. Non ho trovato modo migliore di esprimermi, e poi dopo tutte queste pagine spero che il lettore mi concederà tregua, *n.d.A*

uno "spostamento a destra" di lunghezza 2^{-k} con uno a sinistra della metà. Dunque, dato che ogni volta che si crea un nuovo tratto piatto l'altezza della curva cresce di 2^{-2m-1}, il massimo della funzione sarà:

$$\max f = \sum_{m\in\mathbb{N}} 2^{-2m-1} = \frac{1}{2}\sum_{m\in\mathbb{N}} 4^{-m} = \frac{1}{2}\frac{1}{1-1/4} = \frac{2}{3}$$

Un'altra osservazione possibile è il fatto che la funzione è integrabile, in quanto limite uniforme di una successione di funzioni integrabili. Calcoliamo il suo integrale, noto che $f(x) = u(x) + \frac{1}{2}f(2x)$

$$\int_0^1 f(x)\,dx = \int_0^1 u(x)\,dx + \frac{1}{2}\int_0^1 f(2x)\,dx = \frac{1}{4} + \frac{1}{2}\int_0^1 f(2x)\,dx$$

Cambiando variabile otteniamo

$$\int_0^1 f(x)\,dx = \frac{1}{4} + \frac{1}{4}\int_0^2 f(y)\,dy = \frac{1}{4} + \frac{1}{2}\int_0^1 f(y)\,dy$$

L'ultima disuguaglianza segue dal fatto che la funzione è periodica di periodo 1. Allora isolando l'integrale troviamo

$$\int_0^1 f(x)\,dx = \frac{1}{2}$$

23.12 Variante troll della funzione del pudding

E se considerassimo una funzione pressoché identica alla precedente, ma in cui si sostituisce il 2 al denominatore con un 4?

$$f(x) = \sum_{k=0}^{+\infty} \frac{u_1(2^k x)}{4^k}$$

Fun fact: quello che viene fuori è la parabola $y = 2x - 2x^2$. Per mostrarlo, consideriamo i numeri reali $x \in [0,1]$ che non sono multipli di una potenza di 2. In altri termini, la loro scrittura in base 2 è del tipo:

$$x = a_1 2^{-1} + a_2 2^{-2} \dots \quad \text{con } a_i \in \{0,1\}$$

E tale scrittura è frequentemente non costante. Si osserva che $a_0 = 0$ dato che tutti i punti sono minori di 1. Ora, questi punti costituiscono un insieme, sia esso X, denso in $[0,1]$. nel quale la derivata della funzione $u_1(2^k x)$ è ben definita $\forall k$, in quanto $x \in X$ non può mai essere un picco. Per costruzione vale che

$$u_1'(2^k x) = \begin{cases} 1 & \text{se } a_{k+1} = 0 \\ 0 & \text{se } a_{k+1} = 1 \end{cases} = 1 - 2a_{k+1}$$

Questo perché il coefficiente $a_k + 1$ ci dice se il nostro punto si trova a destra o a sinistra di un picco della funzione $u_1(2^k x)$. Ora, mostriamo che la derivata della funzione è esattamente la derivata della parabola, ossia che nell'insieme degli x considerati la nostra funzione risolve il problema di Cauchy

$$\begin{cases} f'(x) = 2 - 4x \\ f(0) = 0 \end{cases}$$

Se ciò è vero, allora la funzione deve coincidere con la parabola, essendo unica la soluzione integrale. Calcoliamo allora la derivata della funzione, assumendo[49] che la derivata della serie sia la serie delle derivate.

$$f'(x) = \sum_{k=0}^{+\infty} \frac{1 - 2a_{k+1}}{2^k} \stackrel{?}{=} 2 - 4x = 2\left(1 - 2\sum_{k=0}^{+\infty} \frac{a_k}{2^k}\right)$$

[49] É lecito farlo dato che la successione delle derivate converge uniformemente, e si può dimostrare che questa condizione è sufficiente.

La serie alla RHS può essere fatta partire da $k = 1$ dato che $a_0 = 0$. Analogamente, possiamo riscrivere la serie alla LHS partendo da $k = 1$ e sostituendo k con $k - 1$. Portando a sinistra il 2 che moltiplica la RHS riotteniamo 2^k al denominatore.

$$f'(x) = \sum_{k=1}^{+\infty} \frac{1 - 2a_k}{2^k} \stackrel{?}{=} \left(1 - 2\sum_{k=1}^{+\infty} \frac{a_k}{2^k}\right)$$

Osservando che $1 = \sum_{k=1}^{+\infty} \frac{1}{2^k}$ otteniamo:

$$\sum_{k=1}^{+\infty} \frac{1 - 2a_k}{2^k} = \sum_{k=1}^{+\infty} \frac{1 - 2a_k}{2^k}$$

Che è quanto volevamo mostrare. Abbiamo allora che $f(x) = 2x - 2x^2 \ \forall x \in X$. Basta ora estendere la funzione a tutto $[0, 1]$. Tuttavia è noto che la funzione $f(x)$ è continua, in quanto è il limite uniforme di una successione di somme parziali continue, ossia $\sum_{k=0}^{n} \frac{u_1(2^k x)}{4^k}$. L'uniforme convergenza segue dal fatto che la coda della serie convergente $\sum_{k=0}^{+\infty} \frac{1}{4^k}$ è infinitesima, e dunque la somma parziale è vicina a piacere alla funzione limite. Ora, dato che la funzione è definita su un insieme denso, ogni punto $y \in [0, 1]$ ammette una successione di punti $x_n \in X$ che gli converge. Per continuità, si deve avere che $f(y) = \lim_{n\to+\infty} f(x_n) = \lim_{x_n\to y} 2x_n - 2x_n^2 = 2y - 2y^2$.[50]

□

23.13 (*)Feynman's Trick e applicazione alla gaussiana

Si dice che, all'università, Feynman avesse sviluppato un metodo alternativo per cannonare integrali altrimenti impossibili, tramite l'aggiunta di parametri arbitrari e tramite la differenziazione rispetto a quelli. Ho deciso di mostrarlo in questa sezione perché lo trovo elegante e troppo poco conosciuto. Iniziamo con un teorema che permette di derivare sotto il segno di integrale[51].

Teorema Sia $f(x, y) : X \times Y \longrightarrow \mathbb{R}$ una funzione continua in due variabili, dove $X = [a, b]$ e Y è un compatto di $\mathbb{R}$. Supponiamo che esista la funzione derivata parziale rispetto a x, sia essa $\frac{\partial f(x,y)}{\partial x}$, e che sia continua in entrambe le variabili. Allora la funzione

$$g(x) = \int_Y f(x, y)\, dy$$

è di classe C^1 e la sua derivata rispetto a x è la derivata parziale dell'argomento. In simboli

$$\frac{dg(x)}{dx} = \int_Y \partial_x f(x, y)\, dy$$

Dimostrazione La continuità è ovvia e segue banalmente dall'uniforme continuità di f. Per quanto riguarda la derivabilità, consideriamo

$$\frac{dg}{dx} = \lim_{h\to 0} \frac{\int_Y f(x + h, y)\, dy - \int_Y f(x + h, y)\, dy}{h} = \lim_{h\to 0} \int_Y \frac{f(x + h, y) - f(x, y)\, dy}{h}$$

Essendo f continua e derivabile in x, per teorema di Lagrange esiste $c \in (x, x + h)$ tale che

$$\frac{dg(x)}{dx} = \lim_{h\to 0} \int_Y \partial_x f(c, y)\, dy$$

[50]Quanto detto è una semplificazione di un teorema che assicura l'esistenza e l'unicità del prolungamento di una funzione. Si può dimostrare che, se una funzione f è uniformemente continua da un insieme X a un insieme Y completo, allora esiste ed è unica una funzione g che estende f alla chiusura di X, in modo uniformemente continuo. Le ipotesi sono verificate nel caso in questione, dato che $\mathbb{R}$ è completo, f è uniformemente continua perché continua su un compatto e la chiusura di X è tutto $[0, 1]$, essendo X denso.

[51]Tra i fisici è noto come "Derivating Under The Integral Sign", ossia "DUTIS", dato che per loro è un dovere morale far commutare le derivate e gli integrali. Tanto, le ipotesi sono sempre verificate in fisica, per assioma.

Ora, per $h \to 0$ c tende a x e la convergenza è uniforme in y per uniforme continuità della derivata parziale. Allora, per teorema di passaggio al limite sotto il segno di integrale, si conclude che

$$\frac{dg(x)}{dx} = \int_Y \partial_x f(x,y)\, dy$$

Da cui la tesi.

$\square$

Con un po' più conti, è possibile generalizzare il teorema al caso in cui Y è un generico intervallo. Ora, il trucco di Feynman consiste nell'aggiungere parametri a un integrale, derivare rispetto a quelli utilizzando il teorema precedente e solo alla fine ricondursi all'integrale originale.
Ad esempio, si voglia calcolare l'integrale della Gaussiana $\int_0^{+\infty} e^{-x^2}\, dx$. Definiamo

$$I(b) = \int_0^{+\infty} \frac{e^{-x^2}\, dx}{1+(x/b)^2} = \int_0^{+\infty} \frac{e^{-b^2y^2}}{1+y^2} b\, dy$$

L'integrale dipende dal parametro b, e il nostro obiettivo è calcolare il limite dell'integrale per b all'infinito. Non facciamolo subito: cerchiamo prima una relazione che lega gli $I(b)$ tra loro. Proviamo, ad esempio, a derivare rispetto a b la seguente funzione

$$\frac{1}{b}e^{-b^2}I(b) = \frac{1}{b}e^{-b^2}\int_0^{+\infty} \frac{e^{-b^2y^2}}{1+y^2} b\, dy$$

Dato che b è un parametro, posso portarlo dentro e fuori dall'integrale senza problemi. Allora derivando e applicando il teorema precedente si ottiene

$$\frac{d}{db}\frac{1}{b}e^{-b^2}I(b) = \frac{d}{db}\int_0^{+\infty} \frac{e^{-b^2(1+y^2)}}{1+y^2}\, dy = \int_0^{+\infty} \frac{\partial}{\partial b}\left(\frac{e^{-b^2(1+y^2)}}{1+y^2}\right) dy = \int_0^{+\infty} -2be^{-b^2(1+y^2)}\, dy$$

Portiamo fuori nuovamente tutto ciò che dipende da b e otteniamo

$$\frac{d}{db}\frac{1}{b}e^{-b^2}I(b) = -2be^{-b^2}\int_0^{+\infty} e^{-b^2y^2}\, dy = -2e^{-b^2}\int_0^{+\infty} e^{-x^2}\, dx = -2e^{-b^2}I(\infty)$$

Ora, integriamo a destra e a sinistra in db tra zero e infinito. Osserviamo che $\int_0^{+\infty} e^{-b^2} db = I(\infty)$. Otteniamo allora

$$\frac{1}{b}e^{-b^2}I(b) \restriction_0^{+\infty} = -2I(\infty)^2$$

Calcoliamo il termine di sinistra. All'infinito è nullo, dato che sicuramente $I(+\infty)$è finito. Per quanto riguarda il valore che assume in zero, osserviamo che

$$\frac{1}{b}e^{-b^2}I(b) \restriction_0 = \int_0^{+\infty} \frac{e^{-b^2(1+y^2)}}{1+y^2}\, dy \restriction_0 = \int_0^{+\infty} \frac{1}{1+y^2}\, dy$$

L'integrale ottenuto è quell' dell'arcotangente, che fa $\pi/2$. Allora

$$-2I(\infty)^2 = -\pi/2 \quad \Longrightarrow \quad I(\infty) = \sqrt{\pi}/2$$

$\square$

23.14 Insieme di Cantor

Definiamo la seguente successione di insiemi chiusi, incapsulati uno dentro l'altro.

$$\begin{cases} C_0 = [0,1] \\ C_{n+1} = \frac{1}{3}C_n \cup \left(\frac{1}{3}C_n + \frac{2}{3}\right) \end{cases}$$

É semplice verificare che ogni insieme è composto da 2^n intervalli disgiunti, di ampiezza 3^{-n} ciascuno, e che quindi l'ampiezza totale tenda a zero essendo data da $\frac{2^n}{3^n}$. Si può pensare che a ogni passo si sottragga "il terzo centrale" da ogni intervallo dell'insieme precedente.
Ora, consideriamo lo sviluppo in base 3 dei numeri reali in $[0,1]$, ossia scriviamo $x = \sum_{n=1}^{+\infty} a_n 3^{-n}$ con $a_n \in \{0,1,2\}$. Per eliminare le ambiguità insite nella scrittura periodica (ad esempio 0.1 si scrive anche come $0.0\bar{2}$) considereremo sempre la scrittura la cui prima cifra differente dall'altra è la minore, ossia approssimeremo sempre dal basso. L'insieme C_n contiene tutti e soli i numeri reali in $[0,1]$ le cui prime n cifre dopo la virgola sono diverse da 1. Infatti, a ogni passo si tronca la parte centrale di ogni intervallo, quella i cui elementi hanno come n-esima cifra 1. Segue allora facilmente che $C = \bigcap C_n$, noto come insieme di Cantor, è l'insieme dei numeri reali in $[0,1]$ nella cui scrittura decimale non compaiono 1. Osserviamo che ha la cardinalità del continuo, perché ha la stessa cardinalità delle successioni a valori in $\{0,1,2\}$. Inoltre, è trascurabile, dato che $C \subset C_n$, ogni C_n è un'unione finita di intervalli e l'ampiezza $|C_n|$ è infinitesima. Infine, osserviamo che il complementare dell'insieme di Cantor è denso in $[0,1]$, perché tra due numeri ne esiste sempre uno in cui almeno una cifra vale 1.

23.15 Funzione di Cantor-Vitali

Mostriamo che esiste ed è unica una funzione continua tale che

$$f(x) = \begin{cases} \dfrac{f(3x)}{2} & \text{se } 0 \le x \le 1/3 \\ \dfrac{1}{2} & \text{se } 1/3 < x \le 2/3 \\ \dfrac{f(3x-2)}{2} + \dfrac{1}{2} & \text{se } 2/3 < x \le 1 \end{cases}$$

Tale funzione prende il nome di "scala del Diavolo", o funzione di Cantor-Vitali.

Esistenza Consideriamo la seguente successione di funzioni continue:

$$\begin{cases} f_0(x) = x \\ f_{n+1}(x) = \begin{cases} \dfrac{f_n(3x)}{2} & \text{se } 0 \le x \le 1/3 \\ \dfrac{1}{2} & \text{se } 1/3 < x \le 2/3 \\ \dfrac{f_n(3x-2)}{2} + \dfrac{1}{2} & \text{se } 2/3 < x \le 1 \end{cases} \end{cases}$$

Ciascuna funzione è costante a tratti sull'insieme C_n definito sopra, e nel resto del dominio raccorda in modo lineare i pezzi costanti. Mostriamo che la successione di funzioni è di Cauchy uniformemente nella distanza del sup, cosa che permetterà di dire che converge uniformemente a una certa funzione f.[52] Osserviamo che ogni funzione coincide con la precedente, eccetto nei punti in cui una retta obliqua si spezza in due rette e un segmento costante al centro. In tal caso, con banali considerazioni geometriche[53], si può dire che la distanza tra due funzioni successive vale al massimo $\frac{1}{3}\frac{1}{2^n}$. Allora, la distanza tra due funzioni generiche si stima come

$$d_\infty(f_n, f_m) \le \frac{1}{3}\sum_{k=n}^{m} \frac{1}{2^k} \le \frac{1}{3}\sum_{k=n}^{+\infty} \frac{1}{2^k}$$

Essendo la coda della serie geometrica infinitesima, segue la tesi. Otteniamo allora che la successione converge uniformemente a una funzione che rispetta la seguente equazione funzionale:

$$f(x) = \lim_{n\to+\infty} f_{n+1}(x) = \begin{cases} \dfrac{\lim_{n\to+\infty} f_n(3x)}{2} = \dfrac{f(3x)}{2} & \text{se } 0 \le x \le 1/3 \\ \dfrac{1}{2} & \text{se } 1/3 < x \le 2/3 \\ \dfrac{\lim_{n\to+\infty} f_n(3x-2)}{2} + \dfrac{1}{2} = \dfrac{f(3x-2)}{2} + \dfrac{1}{2} & \text{se } 2/3 < x \le 1 \end{cases}$$

[52]Vedasi 21.3
[53]Fatevi un disegnino e convincetevene.

Inoltre, la funzione è continua dato che è limite uniforme di funzioni continue.

Unicità Basta mostrare che, se esiste una funzione continua che rispetta l'equazione funzionale, essa è univocamente determinata sul complementare dell'insieme di Cantor, che è denso. I restanti valori verranno univocamente determinati per continuità.
Sia allora $x \notin C$. Dunque nel suo sviluppo decimale compare una cifra 1. Ciò significa che, reiterando la funzione e cioè continuando a "scalare" le cifre, a un certo punto si otterrà $f(0,1\ldots) = 1/2$. Più formalmente, sia n la posizione del primo 1. Allora otteniamo, continuando a esplicitare $f(3x)$ n volte, che $f(x)$ si esprime come una somma finita di potenze di $1/2$, cosa che la determina univocamente.

□

Proprietà

- Essendo limite di funzioni crescenti, la funzione è debolmente crescente. Tuttavia ha derivata nulla quasi ovunque in quanto non è derivabile sull'insieme di Cantor, che ha misura nulla, mentre è localmente costante sul suo complementare. Ciò fa sì che la funzione del demonio sia annoverata tra le cosiddette "funzioni patologiche".[54]

- La restrizione della funzione all'insieme di Cantor è ancora suriettiva in $[0, 1]$. Infatti, dato che il complementare dell'insieme di Cantor è denso in $[0, 1]$, per continuità l'immagine di un punto dell'insieme è il limite delle immagini di una successione di punti nel complementare. Ora, visto che anche le immagini dei punti del complementare, ossia i punti a sviluppo binario finito, sono dense in $[0, 1]$, qualsiasi valore $y \in [0, 1]$ è limite di una successione di punti a sviluppo finito, e dunque immagine di un punto dell'insieme di Cantor.
  Quanto appena visto mostra nuovamente che l'insieme di Cantor non è numerabile.

- La funzione è H́olderiana di costante $\log_2 3$, e dunque uniformemente continua. Mostriamo infatti che vale l'implicazione

  $$|x - y| < 3^{-n} \quad \Longrightarrow \quad |f(x) - f(y)| < 2^{-n}$$

  Da cui, esplicitando $n = -\log_3(|x - y|)$, otterremo che

  $$|f(x) - f(y)| < 2^{-n} = 2^{\log_3(|x-y|)} = 2^{\frac{\log_2(|x-y|)}{\log_2 3}} = |x - y|^{\frac{1}{\log_2 3}} = |x - y|^{\log_3 2}$$

  Ora, consideriamo due punti a distanza al più 3^{-n}. Dato che la funzione è crescente, il massimo dislivello tra di loro è quello tra due tratti costanti successivi a distanza 3^{-n}. Questa osservazione permette di ridurre lo studio alla funzione f_n, cioè quella che contiene tutti i tratti costanti di lunghezza almeno 3^{-n}. Infatti, comunque si scelgano due tratti rettilinei a distanza al più 3^{-n}, essi saranno compresi (o uguali) tra due tratti costanti della funzione f_n adiacenti.
  Detto ciò, ci basta calcolare di quanto cresce la funzione f_n nei tratti obliqui. Dato che deve attraversare $[0, 1]$ ed è costante su un ampiezza totale di $1/3 \sum_{k=0}^{n-1} (2/3)^k$, la pendenza dei tratti obliqui vale

  $$m = \frac{1}{1 - 1/3 \sum_{k=0}^{n-1} (2/3)^k} = \frac{3^n}{2^n}$$

  Moltiplicando per 3^{-n} otteniamo che il massimo dislivello vale 2^{-n}, come volevasi dimostrare.

  □

- La lunghezza del grafico vale 2. Per mostrarlo, osserviamo innanzitutto che esiste un insieme di spezzate con vertici sul grafico il cui sup della lunghezza è 2. Tali spezzate sono esattamente

[54] Una versione "smussata" della funzione di Cantor-Vitali è l'ancor più enigmatica funzione $?(x)$ di Minkowski. Per i curiosi, rimando a `https://en.wikipedia.org/wiki/Minkowski%27s_question-mark_function`.

le funzioni f_n, costituite da dei pezzi costanti di ampiezza totale $1/3\sum_{k=0}^{n-1}(2/3)^k$ e dei segmenti obliqui, la cui lunghezza totale, per teorema di Pitagora, vale

$$\sqrt{1+\left(1-1/3\sum_{k=0}^{n-1}(2/3)^k\right)^2}=\sqrt{1+\frac{4^n}{9^n}}$$

Consideriamo l'estremo superiore delle lunghezze delle f_n, che evidentemente coincide con il loro limite dato che a ogni passo si sostituiscono segmenti con spezzate. Otteniamo allora che la lunghezza della funzione deve essere almeno

$$\mathcal{L}(f,[0,1])\geq\lim_{n\to+\infty}\sqrt{1+\frac{4^n}{9^n}}+\frac{1}{3}\sum_{k=0}^{n-1}\left(\frac{2}{3}\right)^k=2$$

Ora, mostriamo che, comunque si scelga una partizione $x_{k\,k=1\ldots h}$, la sua lunghezza è al più 2. Basta infatti sostituire ogni pezzo rettilineo con un pezzo verticale e uno orizzontale, la cui somma delle lunghezze è maggiore per disuguaglianza triangolare, e osservare che un insieme di segmenti che collega $(0,0)$ e $(1,1)$ con tratti solo orizzontali e verticali, percorsi sempre dal basso verso l'alto e da sinistra verso destra, ha lunghezza totale 2.

23.16 Funzione analitica modellabile a piacere

Sia $\phi:\ [0,+\infty[\longrightarrow[0,+\infty[$ crescente. Si vuole determinare una funzione analitica a valori complessi (anche detta funzione intera) tale che $f(x)\geq\phi(x)$ per ogni x reale non negativo.
Cerchiamo una funzione del tipo

$$f(x)=C+\sum_{k=1}^{+\infty}\left(\frac{x}{k}\right)^{n_k}$$

Per una certa successione $n_k\in\mathbb{N}$ strettamente crescente e per una certa costante C. Osserviamo che f è crescente e che il suo minimo su $\mathbb{R}^+$ è C, dato che la serie è a termini positivi. Imponiamo che $C=\phi(2)$, cosicché $f(x)\geq\phi(2)\geq\phi(x)$ per $0\leq x\leq 2$, e definiamo induttivamente la successione n_k in modo da soddisfare la richiesta nell'intervallo $k+1\leq x\leq k+2$. Imponiamo

$$f(x)\geq f(k+1)\geq\left(\frac{k+1}{k}\right)^{n_k}\geq\phi(k+2)$$

Da cui, le condizioni minime da imporre su n_k sono

$$\begin{cases}n_k\geq\frac{\log\phi(k+2)}{\log\left(1+\frac{1}{k}\right)}\\ n_k>n_{k-1}\end{cases}$$

Il risultato ottenuto è un caso particolare di un teorema strettamente più forte

Teorema Date due funzioni $\phi(x)<\psi(x)$ continue reali, esiste una funzione analitica f tale che $\phi(x)<f(x)<\psi(x)$.

Corollario Le funzioni analitiche sono dense nelle funzioni continue nella norma del sup. Basta infatti considerare $\psi(x)=\phi(x)-\varepsilon$ per trovare una funzione analitica a meno di ε da ϕ.