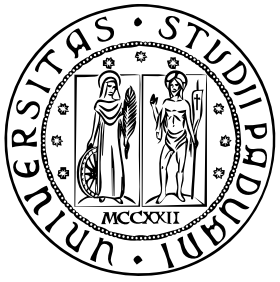

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea in Matematica

Teoria dei numeri trascendenti: dal settimo problema di Hilbert al Teorema di Baker

Relatrice:
Prof.ssa Luisa Fiorot

Laureando: Aurora Tibaldi
Matricola: 2012695

Anno Accademico 2025/2026

data 17/07/2026

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea in Matematica

Teoria dei numeri trascendenti: dal settimo problema di Hilbert al Teorema di Baker

Relatrice:
Prof.ssa Luisa Fiorot

Laureando: Aurora Tibaldi
Matricola: 2012695

Anno Accademico 2025/2026

data 17/07/2026

Indice

Introduzione

Lo studio dei numeri trascendenti si è sviluppato in una teoria fertile ed estesa che arricchisce vasti rami della matematica. Si tratta dell'analisi dei numeri complessi che non sono radici di alcun polinomio non nullo a coefficienti interi. Sebbene l'esistenza di tali numeri sia stata ipotizzata per secoli, la loro identità è rimasta a lungo congetturale dando origine a numerosi quesiti. Il presente lavoro si propone di ripercorrere la storia e i progressi di tale disciplina focalizzandosi sul Teorema di Alan Baker che ha segnato una svolta nel campo delle forme lineari nei logaritmi. Il testo si articola in tre capitoli tracciando una linea di continuità che va dai risultati classici ai metodi moderni.

Il Capitolo 1 introduce la teoria dei numeri trascendenti così come si presentava all'incirca all'inizio del Novecento. La sua evoluzione storica parte dalla dimostrazione dell'esistenza dei numeri trascendenti da parte di Joseph Liouville. Attraverso il teorema che porta il suo nome, Liouville stabilì un limite all'approssimazione razionale dei numeri algebrici permettendo di identificare i primi esempi di numeri trascendenti. La trattazione prosegue con i contributi di molti matematici come Axel Thue, Klaus Roth, Ferdinand von Lindemann e in particolare Charles Hermite che fu in grado di dimostrare la trascendenza del numero e. Le argomentazioni di Hermite condussero alla teoria degli approssimanti di Padé e permisero a Lindemann di risolvere il problema della quadratura del cerchio, determinando la trascendenza di π. In seguito viene approfondito il settimo problema di Hilbert con la soluzione di Aleksandr Osipovič Gel'fond e Theodor Schneider: per α e β algebrici con $\alpha \neq 0, 1$ e β irrazionale, il numero α^β è trascendente. Un esempio è $2^{\sqrt{2}}$, un altro meno ovvio è e^π. Le dimostrazioni di Gel'fond e Schneider arrivarono dopo lo studio delle funzioni intere a valori interi utilizzando formule di interpolazione che risalgono a Hermite. Tuttavia, nonostante la profondità di tali risultati, i metodi di Gel'fond e Schneider presentavano una limitazione in quanto erano applicabili esclusivamente a relazioni coinvolgenti due soli logaritmi, lasciando il caso generale di n variabili come uno dei problemi più ardui della teoria. La generalizzazione di questo risultato fu formulata da Alan Baker. Il teorema afferma che i logaritmi di numeri algebrici linearmente indipendenti sui razionali rimangono linearmente indipendenti anche quando si estendono i coefficienti al campo dei numeri algebrici. Il lavoro di Baker diede un forte impulso alla materia con le sue applicazioni alla risoluzione di equazioni diofantee e alla determinazione dei limiti inferiori per le forme lineari nei logaritmi. Gli anni successivi hanno visto infatti una rinascita di questo campo. Il capitolo si conclude

con l'introduzione degli strumenti preliminari algebrici necessari per il resto della trattazione.

L'obiettivo dei capitoli successivi è quello di addentrarsi nella complessa struttura formale del Teorema di Baker nella sua versione originale [1]. Come riportato nelle note di James Harper [3], sembra essere un'opinione comunemente condivisa che la dimostrazione sia piuttosto misteriosa e controintuitiva, cercheremo dunque di svelare questo mistero rendendo l'argomentazione il più possibile immediata. Per comprendere il filo conduttore che unisce il Capitolo 2 e il Capitolo 3 presentiamo l'intuizione di Baker delineando la strategia che guida l'intera dimostrazione.

L'ipotesi di partenza presuppone l'esistenza di una relazione di dipendenza lineare non banale tra i logaritmi di numeri algebrici, con coefficienti anch'essi algebrici. L'idea chiave consiste nell'intrappolare questa relazione all'interno di una funzione ausiliaria costruita come una combinazione lineare di funzioni esponenziali in più variabili complesse. Essa unisce le proprietà di rigidità delle funzioni analitiche complesse a quelle aritmetiche dei numeri algebrici.
Il primo passo strategico sfrutta il metodo di Siegel per garantire l'esistenza di coefficienti interi non nulli della funzione ausiliaria in modo tale che essa, insieme a un numero elevato di sue derivate parziali si annulli in un prefissato insieme di punti interi. Questo costringe la funzione ad essere "piatta", pur non essendo identicamente nulla, in una specifica regione dello spazio.
A questo punto si attiva un innovativo processo induttivo basato sulla tecnica dell'estrapolazione. Si dimostra che se la funzione ausiliaria possiede una molteplicità di zeri sufficientemente alta vicino all'origine, in virtù del fatto che è olomorfa essa deve rimanere quasi nulla anche a distanze considerevoli usando il Teorema di Schwarz e il Principio del Massimo Modulo. Se valutiamo la funzione in questi nuovi punti il valore ottenuto è, per costruzione, un numero algebrico. Per le proprietà intrinseche dei numeri algebrici, essi non possono essere arbitrariamente vicini a zero senza essere esattamente zero. Questo meccanismo obbliga la funzione ausiliaria ad annullarsi anche nel nuovo insieme allargato di punti interi. L'acquisizione di questi nuovi zeri aumenta l'ordine di annullamento globale della funzione permettendo di ripetere il processo di estrapolazione su un raggio ancora più ampio.
La contraddizione finale emerge quando il raggio di estrapolazione diventa sufficientemente grande: il numero di zeri cresce al punto da forzare le derivate della funzione nell'origine a essere estremamente piccole. Quindi da un lato, la natura discreta dei coefficienti interi impone loro una grandezza minima non nulla, dall'altro le stime analitiche derivanti dall'estrapolazione li spingono a diventare arbitrariamente vicini allo zero. L'intersezione di queste due stime inconciliabili produce l'assurdo. L'ipotesi iniziale di dipendenza lineare viene quindi invalidata dimostrando il Teorema di Baker.

Un altro obiettivo è descrivere alcuni dei principali strumenti disponibili. Il Capitolo 2 è dunque dedicato all'esposizione dei risultati tecnici necessari per affrontare la dimostrazione del Teorema di Baker tra cui la costruzione della funzione ausiliaria Φ e l'algoritmo di estrapolazione.

Il Capitolo 3 presenta la dimostrazione formale del teorema ed espone i principali

corollari che ne derivano, come la trascendenza di combinazioni lineari di logaritmi e di espressioni esponenziali complesse. Un altro aspetto trattato è l'effettività del metodo di Baker che ha permesso di calcolare limiti inferiori precisi per le forme lineari nei logaritmi. Vengono inoltre accennate strategie alternative per dimostrare il teorema principale. Tra queste vi è l'approccio con i determinanti di interpolazione di Laurent che supera la necessità di ricorrere al principio dei cassetti di Dirichlet.
Il lavoro si chiude con uno sguardo agli sviluppi successivi e alle congetture ancora aperte come la Congettura di Schanuel che rappresenta oggi una delle sfide più ambiziose nel campo della trascendenza.

Capitolo 1

Evoluzione storica e formalismo algebrico

In questo primo capitolo delineiamo il percorso storico della teoria dei numeri trascendenti ponendo l'enfasi sui risultati che hanno prodotto la più ampia applicazione. In particolare ci soffermeremo sulla soluzione del settimo problema di Hilbert e il Teorema di Alan Baker.
Nell'ultima parte del capitolo si presentano gli strumenti preliminari indispensabili per lo sviluppo delle dimostrazioni successive. Vengono formalizzate le definizioni di numero e intero algebrico, di polinomio minimo, e i concetti di coniugati e norma. A ciò si affiancano la nozione di altezza e la stima sul modulo dei coniugati. Inoltre, viene introdotta sia la proposizione sulla norma di un intero algebrico non nullo, la quale stabilisce che il suo modulo sia maggiore o uguale a 1, sia la stima induttiva sulla crescita dei coefficienti del polinomio minimo.

1.1 Dalle origini al Teorema di Baker

Lo studio dei numeri trascendenti ha origine da antichi problemi greci come la quadratura del cerchio, ma si afferma come disciplina indipendente solo a partire dal XIX secolo.
Nel 1844 Joseph Liouville dimostrò per la prima volta l'esistenza dei trascendenti, ovvero numeri che non sono algebrici e che quindi non sono radici di alcun polinomio a coefficienti interi. Liouville, inoltre, appurò che la distanza tra un numero algebrico non razionale e un numero razionale non può essere inferiore a un certo valore teorico. Esiste, dunque, un limite sull'accuratezza con cui possono essere approssimati questi numeri. Grazie a questa osservazione si riuscì a formulare il primo criterio pratico con cui costruire i trascendenti.

Teorema 1.1 (Liouville)**.** *Sia α un numero algebrico di grado $n > 1$, allora esiste $c(\alpha) > 0$ una costante tale che, per ogni numero razionale p/q $(q > 0)$,*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^n}.$$

Di conseguenza i numeri che violano il limite del Teorema 1.1 sono definiti trascendenti. Tra questi troviamo i decimali non periodici con lunghe sequenze di zeri e le frazioni continue con quozienti parziali la cui crescita è estremamente rapida.
I numeri reali ξ con una successione di approssimazioni razionali distinte p_n/q_n $(n = 1, 2, \dots)$ tali che $|\xi - p_n/q_n| < 1/q_n^{w_n}$, dove $\limsup w_n = \infty$, sono noti come "numeri di Liouville" e sono pertanto trascendenti.
Sfruttando il Teorema 1.1, uno dei primi numeri di cui è stata dimostrata la trascendenza è

$$\xi = \sum_{n=1}^{\infty} 10^{-n!}.$$

In generale dimostrare la trascendenza di un determinato numero rappresenta ancora un problema complicato.

Nel tempo il Teorema di Liouville è stato migliorato diverse volte.

Teorema 1.2 (Thue)**.** *Sia α un numero algebrico di grado $n > 1$, p, q interi razionali $(q > 0)$ allora*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha, k)}{q^k}$$

e $c(\alpha, \kappa) > 0$ costante esiste per $\kappa > \frac{1}{2}n + 1$.

Il Teorema di Roth rappresenta la miglior stima possibile.

Teorema 1.3 (Roth)**.** *Sia α un numero algebrico, p, q interi razionali $(q > 0)$ allora*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha, k)}{q^k}$$

e $c(\alpha, \kappa) > 0$ costante esiste per qualsiasi $\kappa > 2$.

Nel 1874 Georg Cantor introdusse il concetto di numerabilità da cui si concluse che "quasi tutti" i numeri devono essere trascendenti. Un anno prima nel 1873 Charles Hermite dimostrò la trascendenza di e, la base naturale dei logaritmi, dando vita ad una nuova era nella teoria dei trascendenti. La sua dimostrazione si basava su quelle che oggi sono note come approssimazioni di Padé per $e^x, \dots, e^{nx}$. Circa un decennio dopo, nel 1882 Ferdinand von Lindemann, generalizzando i metodi usati da Hermite a $e^{\alpha_1 x}, \dots, e^{\alpha_n x}$, riuscì a stabilire che anche π è trascendente. Questo risultato fu molto importante perché risolse definitivamente l'antico problema della quadratura del cerchio.

I Greci avevano cercato di costruire, utilizzando solo riga e compasso, un quadrato di area uguale a quella di un cerchio dato. Se consideriamo un cerchio di raggio unitario la sua area è π. Di conseguenza il quadrato equivalente deve avere il lato di lunghezza $\sqrt{\pi}$. Il problema si traduce nel definire due punti nel piano a una distanza $\sqrt{\pi}$ l'uno dall'altro. Tutti i punti che si possono tracciare con riga e compasso sono ottenuti dall'intersezione di linee rette e cerchi, quindi dalla soluzione di equazioni di primo e secondo grado, segue che le loro coordinate, in un opportuno sistema

di riferimento, sono date da numeri algebrici. La trascendenza di π implica che la quadratura del cerchio è impossibile.

Successivamente Lindemann formulò ulteriori teoremi che vedono la trascendenza di e e π come loro casi particolari. Uno di questi è il Teorema di Hermite-Lindemann.

Teorema 1.4 (Hermite-Lindemann)**.** *Sia β un numero complesso non nullo. Allora almeno uno dei due numeri β, e^β è trascendente.*

Per cui, se β è algebrico, e^β deve essere trascendente.
Un altro suo importante risultato che fu poi provato da Karl Weierstrass è il seguente.

Teorema 1.5 (Lindemann-Weierstrass)**.** *Siano $\alpha_1, \ldots, \alpha_n$ numeri algebrici distinti. Allora per ogni $\beta_1, \ldots, \beta_n$ algebrici non nulli vale*

$$\beta_1 e^{\alpha_1} + \cdots + \beta_n e^{\alpha_n} \neq 0.$$

Il Teorema 1.5 prova automaticamente la trascendenza di e e π: la trascendenza di e si ottiene ponendo $n = 2$, $\alpha_1 = 0$, $\alpha_2 = 1$ e $\beta_2 = -1$, mentre la trascendenza di π deriva dall'equazione di Eulero $e^{i\pi} = -1$. Un enunciato equivalente è conosciuto come Teorema di Lindemann.

Teorema 1.6 (Lindemann)**.** *Se $\alpha_1, \ldots, \alpha_n$ sono numeri algebrici linearmente indipendenti su $\mathbb{Q}$, allora gli n numeri $e^{\alpha_1}, \ldots, e^{\alpha_n}$ sono algebricamente indipendenti.*

Questo è uno dei pochissimi risultati sull'indipendenza algebrica dei numeri collegati alla funzione esponenziale.
Come corollari del Teorema 1.5 abbiamo anche la trascendenza delle funzioni trigonometriche $\cos(\alpha), \sin(\alpha), \tan(\alpha)$ per ogni algebrico $\alpha \neq 0$, e di $\log(\alpha)$ per ogni α algebrico diverso da 0 o 1.

Gli studi derivanti dagli scritti di Hermite-Lindemann hanno prodotto misure di irrazionalità e di trascendenza per e, π. Kurt Mahler ottenne $|\pi - p/q| > 1/q^{42}$ valido per tutti i razionali p/q (con $q > 1$). Per quanto riguarda e si ha $|e - p/q| > c/q^{2+\varepsilon}$ per ogni $\varepsilon > 0$, dove $c = c(\varepsilon)$ è una costante positiva effettiva. Più in generale abbiamo la misura di trascendenza $|e - \alpha| > c/H^{n+1+\varepsilon}$ valida per tutti gli algebrici α con grado n, altezza H, definita come il massimo dei moduli dei coefficienti del polinomio minimo di α, e c è una costante positiva dipendente da n e da ε.
Dopo questi primi contributi, sono stati rilevanti i lavori forniti dai matematici Carl Ludwig Siegel, Aleksandr Osipovič Gel'fond e Theodor Schneider che permisero di risolvere il settimo problema di Hilbert.

Nel 1900 al Congresso Internazionale dei Matematici tenutosi a Parigi David Hilbert sollevò come settimo punto della sua celebre lista di 23 problemi, la questione se un logaritmo irrazionale di un numero algebrico in una base algebrica sia trascendente. Una delle formulazioni del problema proposte da Hilbert riporta questa forma:

Settimo Problema di Hilbert. *Si può affermare che se $\alpha \neq 0, 1$ è un numero algebrico e β è un irrazionale algebrico, l'espressione α^β rappresenta necessariamente un trascendente?*

Nonostante lo scetticismo iniziale di Hilbert sulla fattibilità della soluzione, un punto di svolta arrivò nel 1929. Gel'fond, impiegando tecniche di interpolazione, mostrò che il logaritmo di un numero algebrico in una base algebrica non può essere un irrazionale quadratico immaginario, ossia un numero complesso, soluzione di un'equazione di secondo grado a coefficienti interi il cui discriminante è negativo. Questo significa che α^β è trascendente per ogni algebrico $\alpha \neq 0, 1$ e per ogni irrazionale quadratico immaginario β. In particolare, ciò implica che $e^\pi = (-1)^{-i}$ è trascendente.
Il risultato fu esteso agli irrazionali quadratici reali β da Rodion Osievich Kuzmin nel 1930. Solo nel 1934 il problema fu definitivamente risolto da Gel'fond e Schneider, indipendentemente e simultaneamente. Gli argomenti da loro scoperti sono applicabili a ogni irrazionale β ed entrambi dipendono dalla costruzione di una funzione ausiliaria che si annulla in certi punti selezionati. Una tecnica simile era stata utilizzata pochi anni prima da Siegel. La soluzione del problema è racchiusa nel Teorema di Gel'fond-Schneider.

Teorema 1.7 (Gel'fond-Schneider)**.** *Per ogni insieme di numeri algebrici non nulli* $\alpha_1, \alpha_2, \beta_1, \beta_2$ *con* $\log \alpha_1, \log \alpha_2$ *linearmente indipendenti sui razionali, si ha*

$$\beta_1 \log \alpha_1 + \beta_2 \log \alpha_2 \neq 0.$$

Il legame con il problema di Hilbert si vede più chiaramente enunciando il Teorema 1.7 come segue:

Se α *e* β *sono due numeri complessi con* $\alpha \notin \{0, 1\}$ *e* $\beta \notin \mathbb{Q}$*, allora almeno uno dei tre numeri* α*,* β *o* α^β *(inteso come* $e^{\beta \log \alpha}$*) è trascendente.*

Segue che se α è un algebrico diverso da 0 e 1, $\log \alpha$ un suo logaritmo complesso non nullo, e β un algebrico irrazionale, allora $\alpha^\beta = e^{\beta \log \alpha}$ è un trascendente. Si evince che $2^{\sqrt{2}}$ è un numero trascendente. La trascendenza di e^π si ottiene, ad esempio, scegliendo la base algebrica $\alpha = 1$, il suo logaritmo complesso $\log \alpha = 2i\pi$ e l'esponente algebrico irrazionale $\beta = -i/2$.
Lo stesso Gel'fond sottolineò l'importanza di ottenere una generalizzazione di questo enunciato a più di due logaritmi. Questo traguardo, di fondamentale importanza per la teoria della trascendenza, si raggiunse nel 1966 con il matematico inglese Alan Baker.

Baker dimostrò un enunciato riguardante le forme lineari nei logaritmi di numeri algebrici che generalizza il Teorema 1.7 a n logaritmi algebrici.

Teorema 1.8 (Baker)**.** *Siano* $\alpha_1, \ldots, \alpha_n$ *numeri algebrici non nulli tali che* $\log \alpha_1, \ldots, \log \alpha_n$ *siano linearmente indipendenti sui razionali.*
Allora $1, \log \alpha_1, \ldots, \log \alpha_n$ *sono linearmente indipendenti sul campo di tutti i numeri algebrici.*

Qui $\log \alpha_1, \ldots, \log \alpha_n$ sono determinazioni fisse qualsiasi dei logaritmi.
Questa scoperta contribuì a far ottenere a Baker la Medaglia Fields nel 1970.
La dimostrazione del teorema dipende dalla costruzione di una funzione ausiliaria di

più variabili complesse che generalizza la funzione di una singola variabile impiegata originariamente da Gel'fond e supera le limitazioni tecniche precedenti attraverso una procedura di estrapolazione sofisticata, in cui l'intervallo di interpolazione viene esteso, mentre l'ordine delle derivate viene ridotto. L'analisi dettagliata di questa dimostrazione sarà l'oggetto dei capitoli successivi.

1.2 Nozioni preliminari

Dopo aver delineato il quadro storico e l'importanza del Teorema 1.8, introduciamo ora il formalismo algebrico necessario per analizzarne la complessa struttura dimostrativa.

Definizione 1.9. Un numero complesso $\alpha \in \mathbb{C}$ si dice algebrico se esiste un polinomio non nullo $P(X) \in \mathbb{Z}[X]$ tale che $P(\alpha) = 0$. In caso contrario, α è detto trascendente. Se $P(X)$ è monico allora α è detto un intero algebrico. L'insieme di tutti i numeri algebrici forma un campo, indicato con $\overline{\mathbb{Q}}$, mentre gli interi algebrici costituiscono un sottoanello di quest'ultimo.

Definizione 1.10. Per ogni $\alpha \in \overline{\mathbb{Q}}$, esiste un polinomio a coefficienti interi, con coefficiente principale positivo, di grado minimo $d \geq 1$ che ammette α come radice. Tale polinomio è detto polinomio minimo di α e diciamo che α ha grado d. Le d radici del polinomio minimo sono denominate coniugati di α e si indicano con $\alpha^{(1)} = \alpha, \alpha^{(2)}, \ldots, \alpha^{(d)}$.

Definizione 1.11. La norma di $\alpha \in \overline{\mathbb{Q}}$, indicata con $\mathcal{N}(\alpha)$, è definita come il prodotto dei suoi coniugati

$$\mathcal{N}(\alpha) = \alpha^{(1)} \ldots \alpha^{(d)}.$$

La norma di un intero algebrico soddisfa una proprietà essenziale:

Proposizione 1.12. *Sia α un intero algebrico non nullo. Allora $|\mathcal{N}(\alpha)| \geq 1$.*

Dimostrazione. Sia $P(X) = X^d + A_1 X^{d-1} + \cdots + A_d$ il polinomio minimo di α. Per definizione di intero algebrico, $P(X) \in \mathbb{Z}[X]$ è monico, dunque il coefficiente principale è $A_0 = 1$. Dalle formule di Viète, il prodotto delle radici soddisfa:

$$\mathcal{N}(\alpha) = (-1)^d A_d.$$

Poiché $\alpha \neq 0$ e il polinomio è irriducibile in $\mathbb{Z}[X]$, il termine noto A_d deve essere un intero non nullo e pertanto:

$$|\mathcal{N}(\alpha)| = |(-1)^d A_d| \geq 1.$$

□

Definizione 1.13. L'altezza di un polinomio $P(X) \in \mathbb{Q}[X]$, indicata come $H(P)$, è il massimo dei valori assoluti dei suoi coefficienti. L'altezza di un numero algebrico α, indicata come $h(\alpha)$, è l'altezza del suo polinomio minimo in $\mathbb{Z}[X]$. Con $\overline{|\alpha|}$ indichiamo il massimo dei valori assoluti dei coniugati di α (rispetto a $\mathbb{Q}$).

Osserviamo che ogni numero algebrico α può essere ricondotto a un intero algebrico tramite un'opportuna normalizzazione. Se $P(X) = A_0X^d + A_1X^{d-1} + \cdots + A_d$ è il polinomio minimo di α e A_0 è il suo coefficiente principale, è sempre possibile definire $\gamma = A_0\alpha$ affinché γ sia un intero algebrico: $P(\alpha) = A_0\alpha^d + A_1\alpha^{d-1} + \cdots + A_d = 0$, sostituendo $\alpha = \gamma/A_0$ e moltiplicando per A_0^{d-1} si ottiene un polinomio monico a coefficienti interi. Il minimo intero positivo ν tale che $\nu\alpha$ sia un intero algebrico si dice denominatore di α.
Sfruttando la struttura del polinomio minimo, è possibile determinare esplicitamente l'espressione delle potenze successive dell'intero algebrico $A_0\alpha$. Il risultato è formalizzato nella proposizione seguente.

Proposizione 1.14. *Sia $\alpha \in \overline{\mathbb{Q}}$ di grado $d \geq 1$ tale che $P(\alpha) = A_0\alpha^d + A_1\alpha^{d-1} + \cdots + A_d = 0$ con $A_0, \ldots, A_d$ interi razionali. Per ogni intero non negativo j, l'espressione $(A_0\alpha)^j$ può essere scritta come combinazione lineare delle potenze inferiori di α:*

$$(A_0\alpha)^j = A_0^{(j)} + A_1^{(j)}\alpha + \cdots + A_{d-1}^{(j)}\alpha^{d-1}$$

con $A_m^{(j)}$ interi razionali che soddisfano

$$A_m^{(j)} = A_0A_{m-1}^{(j-1)} - A_{d-m}A_{d-1}^{(j-1)} \quad (per\ 0 \leq m < d, j \geq d) \qquad (1.1)$$
$$e\ A_{-1}^{(j-1)} = 0.$$

Dimostrazione. Si procede per induzione su j.

1. Caso base ($0 \leq j < d$): per queste potenze $(A_0\alpha)^j = A_0^j\alpha^j$. Poiché j è minore del grado d, questa è già una combinazione lineare di $1, \alpha, \ldots, \alpha^{d-1}$ dove l'unico coefficiente non nullo è $A_j^{(j)} = A_0^j$.

2. Passo induttivo ($j \geq d$): supponiamo che la tesi sia valida per $j-1$:

$$(A_0\alpha)^{j-1} = \sum_{n=0}^{d-1} A_n^{(j-1)}\alpha^n.$$

Moltiplicando entrambi i membri per $A_0\alpha$ e sostituendo il nuovo termine α^d con l'espressione $A_0\alpha^d = -A_1\alpha^{d-1} - \cdots - A_d$ ottenuta da $P(\alpha) = 0$ risulta che $(A_0\alpha)^j$ è una combinazione lineare delle potenze inferiori di α. Raccogliendo i termini per ogni potenza α^m con $0 \leq m < d$, si trova la relazione ricorsiva sui coefficienti $A_m^{(j)}$.

□

Considerando i coefficienti $A_m^{(j)}$ definiti nella Proposizione 1.14 si dimostra un ulteriore risultato.

Proposizione 1.15. *Se $A_0, \ldots, A_d$ hanno valore assoluto al più A allora $|A_m^{(j)}| \leq (2A)^j \quad \forall j \geq 0$ e $0 \leq m < d$.*

Dimostrazione. Si procede per induzione su j.

1. Caso base ($0 \leq j < d$): risulta $A_j^{(j)} = A_0^j$ allora $|A_j^{(j)}| = |A_0|^j \leq A^j \leq (2A)^j$.
2. Passo induttivo ($j \geq d$): supponiamo che la proprietà sia vera per $j-1$, ovvero che per ogni $m \in \{0, \ldots, d-1\}$ valga $|A_m^{(j-1)}| \leq (2A)^{j-1}$. Utilizzando la relazione ricorsiva (1.1) e la disuguaglianza triangolare:

$$\begin{aligned} |A_m^{(j)}| &\leq |A_0||A_{m-1}^{(j-1)}| + |A_{d-m}||A_{d-1}^{(j-1)}| \\ &\leq A \cdot (2A)^{j-1} + A \cdot (2A)^{j-1} = (2A)^j. \end{aligned}$$

□

Dalla definizione di altezza deduciamo che la Proposizione 1.15 è verificata per $A = h(\alpha)$.
Chiudiamo la sezione con una stima fondamentale che permette di limitare la crescita di un numero algebrico tramite i parametri del suo polinomio minimo. Questa stima risulta quindi valida anche per ciascuno dei suoi coniugati.

Proposizione 1.16. *Sia α un numero algebrico di grado $d \geq 1$ e altezza $h(\alpha)$. Allora vale la seguente stima*

$$|\alpha| \leq d \cdot h(\alpha).$$

Dimostrazione. Sia $P(X) = A_0X^d + A_1X^{d-1} + \cdots + A_d \in \mathbb{Z}[X]$ il polinomio minimo di α. Per definizione di altezza i coefficienti soddisfano $|A_j| \leq h(\alpha)$ per ogni $j = 0, \ldots, d$ e $A_0 \geq 1$. Se $|\alpha| < 1$ la stima è banalmente verificata.
Supponiamo $|\alpha| \geq 1$, da cui $\alpha^{-k} \leq 1$ per ogni $k \geq 1$. Isolando il termine con la potenza più alta e dividendo per α^{d-1} in $P(\alpha) = 0$ si ottiene

$$A_0\alpha = -(A_1 + A_2\alpha^{-1} + \cdots + A_d\alpha^{-d+1}).$$

Prendendo il valore assoluto di entrambi i membri si ricava

$$|\alpha| \leq |A_0\alpha| = |A_1 + A_2\alpha^{-1} + \cdots + A_d\alpha^{-d+1}| \leq |A_1| + |A_2| + \cdots + |A_d| \leq d \cdot h(\alpha).$$

□

Capitolo 2

Strumenti tecnici e funzione ausiliaria

La struttura della dimostrazione del Teorema di Baker 1.8 è molto articolata, per questo motivo si ricorre all'utilizzo di alcuni risultati intermedi che permettono di renderla più lineare. Tuttavia, per favorire un confronto immediato con la dimostrazione di Alan Baker [1], si è scelto di mantenere la numerazione dei lemmi cardine preceduta dalla lettera B (es. Lemma B1 per il Lemma di Siegel) che corrispondono alla numerazione originale di Baker, riservando la numerazione decimale (es. Lemma 2.1) ai soli contributi tecnici aggiuntivi. Prima di esporre questi enunciati anticipiamo una loro breve presentazione al fine di semplificarne la lettura e chiarire le relazioni tra di essi e il teorema principale.
Il Lemma B1 ovvero il Lemma di Siegel è un risultato preliminare sulla risoluzione di sistemi di equazioni lineari con coefficienti interi. Si basa sul principio dei cassetti di Dirichlet ed è necessario per garantire l'esistenza dei coefficienti della funzione ausiliaria definita in seguito, senza che essi siano tutti nulli o eccessivamente grandi.
Con il Lemma B2 si costruisce, generalizzando il metodo di Gel'fond, la funzione ausiliaria Φ di più variabili complesse affinché essa e le sue derivate si annullino in determinati punti.
I tre lemmi successivi analizzano la funzione ausiliaria e implementano la procedura di estrapolazione. Il Lemma B3 stabilisce le stime analitiche di base per Φ, fondamentali per supportare l'algoritmo di estrapolazione definito nel Lemma B4. Quest'ultimo rappresenta il passo induttivo della dimostrazione che permette di estendere l'intervallo di interpolazione riducendo contemporaneamente l'ordine delle derivate. Il Lemma B5 conclude tale processo fornendo una stima estremamente piccola per le derivate di Φ nell'origine. Queste disuguaglianze sono quelle che porteranno alla contraddizione finale nella dimostrazione del Teorema di Baker.
Gli ultimi due lemmi costituiscono degli strumenti supplementari necessari per la dimostrazione del teorema. Il Lemma B6 presenta un limite inferiore per una forma lineare di logaritmi, mentre il Lemma B7 introduce un polinomio di interpolazione.

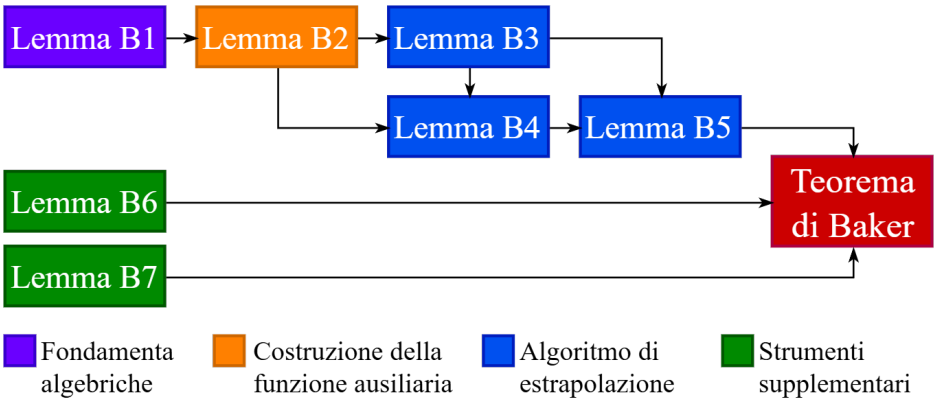

Figura 2.1: Mappa logica della dimostrazione del Teorema di Baker.

2.1 Notazioni

Fissiamo in questa sezione alcune notazioni e le ipotesi generali che si intenderanno valide per il resto del capitolo. Nel seguito assumeremo $\alpha_1, \dots, \alpha_n$ numeri algebrici non nulli i cui logaritmi $\log \alpha_1, \dots, \log \alpha_n$ sono linearmente indipendenti su $\mathbb{Q}$ e $\beta_0, \dots, \beta_n$ numeri algebrici non tutti nulli tali che

$$\beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n = 0.$$

Senza perdere di generalità supponiamo che β_n sia diverso da 0. Poiché la relazione precedente rimane valida sostituendo β_j con $-\beta_j/\beta_n$ è lecito considerare $\beta_n = -1$. In forma esponenziale si ottiene l'equazione

$$e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}} = \alpha_n. \tag{2.1}$$

Sia d il massimo dei gradi di $\alpha_1, \dots, \alpha_n, \beta_0, \dots, \beta_{n-1}$ e siano $a_1, \dots, a_n, b_0, \dots, b_{n-1}$ i coefficienti principali dei rispettivi polinomi minimi. Dalla Proposizione 1.14 si determinano le espressioni

$$(a_r \alpha_r)^j = \sum_{s_r=0}^{d-1} a_{r,s_r}^{(j)} \alpha_r^{s_r}, \quad (b_r \beta_r)^j = \sum_{t_r=0}^{d-1} b_{r,t_r}^{(j)} \beta_r^{t_r}, \tag{2.2}$$

dove $a_{r,s_r}^{(j)}, b_{r,t_r}^{(j)}$ sono interi razionali.
Definiamo $\mathcal{A} = \max_{\substack{1 \le i \le n \\ 0 \le j \le n-1}} \{h(\alpha_i), h(\beta_j)\}$. La Proposizione 1.15 implica che

$$|a_{r,s_r}^{(j)}|, |b_{r,t_r}^{(j)}| \le (2\mathcal{A})^j = c_1^j \quad \text{con } c_1 = 2\mathcal{A}. \tag{2.3}$$

Al fine di giustificare rigorosamente le stime asintotiche introduciamo un breve lemma di confronto.

Lemma 2.1. *Siano $f(h)$ e $g(h)$ due funzioni reali positive definite per interi $h > 0$. Se*

$$\lim_{h\to\infty} \frac{f(h)}{g(h)} = 0, \quad (f(h) = o(g(h)))$$

allora, per ogni $\epsilon > 0$, esiste una costante $C = C(\epsilon)$ tale che per ogni intero $h > C$ risulti:

$$f(h) < \epsilon \cdot g(h)$$

In particolare scegliendo $\epsilon = 1$ esiste una costante c tale che $f(h) < g(h)$ per ogni $h > c$.

Dimostrazione. Per ipotesi, $\lim_{h\to\infty} \frac{f(h)}{g(h)} = 0$. Fissato un $\epsilon > 0$, per la definizione di limite esiste un valore reale C tale che per ogni $h > C$ si ha:

$$\left| \frac{f(h)}{g(h)} - 0 \right| < \epsilon$$

Poiché abbiamo ipotizzato che $f(h)$ e $g(h)$ siano positive, la condizione si riduce a:

$$\frac{f(h)}{g(h)} < \epsilon \quad \Longrightarrow \quad f(h) < \epsilon \cdot g(h)$$

□

Indichiamo con h un valore intero positivo. D'ora in poi quando stabiliremo che il rapporto tra due funzioni tende a zero, $lim_{h\to\infty} f(h)/g(h) = 0$, scriveremo semplicemente che per h sufficientemente grande vale la disuguaglianza $f(h) < g(h)$. Tale espressione sottintende sempre l'esistenza di una costante $c > 0$ che dipende esclusivamente dai parametri fissi α_i, β_j e dai rispettivi logaritmi affinché la relazione sia verificata per qualsiasi $h > c$.
Con $c_1, c_2, \ldots$ denotiamo delle costanti positive di cui forniremo espressioni specifiche nel corso della trattazione e dipendenti, come c, solo dai valori α_i, β_j e dalle determinazioni originarie dei logaritmi.
Infine per brevità, poniamo

$$f_{m_0,\ldots,m_{n-1}}(z_0, \ldots, z_{n-1}) = \left(\frac{\partial}{\partial z_0}\right)^{m_0} \cdots \left(\frac{\partial}{\partial z_{n-1}}\right)^{m_{n-1}} f(z_0, \ldots, z_{n-1}),$$

dove f rappresenta una funzione intera e $m_0, \ldots, m_{n-1}$ sono interi non negativi.

2.2 Fondamenta algebriche

Lemma B1 (Siegel)**.** *Siano $M, N \in \mathbb{Z}$ con $N > M > 0$. Sia $A = (u_{i,j})$ una matrice $M \times N$ a coefficienti interi con $|u_{i,j}| \leq U$ (dove $U \geq 1$) per ogni $i = 1, \ldots, M$ e $j = 1, \ldots, N$. Allora esiste un vettore non nullo $v \in \mathbb{Z}^N$ tale che $Av = \vec{0}$ e ogni coordinata di v ha valore assoluto al più $(NU)^{M/(N-M)}$.*

Dimostrazione. Sia $B > 0$ un numero reale. Definiamo l'insieme

$$X_B := \{v = (x_1, \dots, x_N)^t \in \mathbb{Z}^N : 0 \le x_j \le \lfloor B \rfloor, \forall j = 1, \dots, N\}.$$

Ogni coordinata x_j può assumere esattamente $\lfloor B \rfloor + 1$ valori distinti, pertanto la cardinalità dell'insieme X_B è data da

$$|X_B| = (\lfloor B \rfloor + 1)^N.$$

Sia $f : \mathbb{Z}^N \to \mathbb{Z}^M$ l'applicazione lineare definita dalla moltiplicazione per la matrice A, $f(v) = Av$. Consideriamo l'immagine $Y_B = f(X_B)$. Per ogni $v \in X_B$ le componenti del vettore immagine $y = (y_1, \dots, y_M)^t$ sono espresse dalla sommatoria $y_i = \sum_{j=1}^{N} u_{i,j} x_j$. Per determinare l'intervallo di variazione di ciascuna componente y_i introduciamo le seguenti quantità:

$$W_i \coloneqq \sum_{j : u_{i,j} > 0} u_{i,j}, \quad -V_i \coloneqq \sum_{j : u_{i,j} < 0} u_{i,j}.$$

Dato il vincolo $0 \le x_j \le \lfloor B \rfloor$, il valore minimo di y_i si ottiene ponendo $x_j = \lfloor B \rfloor$ quando $u_{i,j} < 0$ e $x_j = 0$ altrove, per il valore massimo si procede in modo simmetrico. Risulta dunque

$$-V_i \lfloor B \rfloor \le y_i \le W_i \lfloor B \rfloor, \quad 1 \le i \le M.$$

Si osservi che $V_i + W_i = \sum_{j=1}^{N} |u_{i,j}| \le NU$ in virtù dell'ipotesi sui coefficienti della matrice A.
Definiamo un insieme $\overline{Y_B} \subseteq \mathbb{Z}^M$ che contenga l'immagine Y_B come segue

$$\overline{Y_B} := \{w = (y_1, \dots, y_M)^t \in \mathbb{Z}^M : \ -V_i \lfloor B \rfloor \le y_i \le W_i \lfloor B \rfloor, 1 \le i \le M\} \supseteq Y_B.$$

In tale intervallo ogni coordinata y_i può assumere $(W_i + V_i)\lfloor B \rfloor + 1 \le NU \lfloor B \rfloor + 1$ valori interi. Segue che la cardinalità di $\overline{Y_B}$ è limitata da

$$|\overline{Y_B}| \le (NU \lfloor B \rfloor + 1)^M.$$

Per il principio dei cassetti se $|X_B| > |\overline{Y_B}|$ esistono due vettori distinti $v', v'' \in X_B$ tali che $f(v') = f(v'')$ ovvero $Av' = Av''$. Definiamo il vettore differenza $v = v' - v''$. Poiché $v' \neq v''$, si ha $v \neq \vec{0}$ e $Av = \vec{0}$. Essendo $0 \le x'_j, x''_j \le B$ le coordinate x_j di v soddisfano $|x_j| \le B$.
La condizione $|X_B| > |\overline{Y_B}|$ è garantita se

$$(B + 1)^N > (NUB + 1)^M.$$

Utilizzando lo sviluppo binomiale analizziamo la differenza tra i due membri

$$\sum_{i=0}^{N} \binom{N}{i} B^{N-i} - \sum_{j=0}^{M} \binom{M}{j} (NUB)^{M-j} > 0.$$

Siccome $N > M$ possiamo spezzare la prima sommatoria in due parti

$$\sum_{i=0}^{M}\binom{N}{i}B^{N-i}+\sum_{i=M+1}^{N}\binom{N}{i}B^{N-i}-\sum_{i=0}^{M}\binom{M}{i}(NUB)^{M-i}>0$$
$$\sum_{i=0}^{M}\left[\binom{N}{i}B^{N-i}-\binom{M}{i}(NUB)^{M-i}\right]+\sum_{i=M+1}^{N}\binom{N}{i}B^{N-i}>0.$$

Affinché la disuguaglianza sia verificata è sufficiente che ogni componente della prima sommatoria sia non negativo.
Per $i = 0$ avendo tutti termini positivi e $N > M > 0$ richiediamo

$$B^N \geq (NUB)^M \implies B^{N-M} \geq (NU)^M \implies B \geq (NU)^{M/(N-M)}.$$

Questa condizione è sufficiente anche per il confronto con gli addendi successivi. Per $i = 1, \ldots, M$, $N - i > M - i$ e

$$B^N \geq (NUB)^M \implies B^{N-i} > (NUB)^{M-i}.$$

Dato che $\binom{N}{i} > \binom{M}{i}$ per $N > M$, si conclude che ogni addendo della prima sommatoria è strettamente positivo.
Fissando $B = (NU)^{M/(N-M)}$ si assicura l'esistenza della soluzione non banale richiesta. □

2.3 Costruzione della funzione ausiliaria

Lemma B2. *Esistono $p(\lambda_0, \ldots, \lambda_n)$ interi, non tutti nulli, con valore assoluto al più e^{h^3} tali che la funzione*

$$\Phi(z_0, \ldots, z_{n-1}) := \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) z_0^{\lambda_0} e^{\lambda_n \beta_0 z_0} \alpha_1^{\gamma_1 z_1} \ldots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}},$$

dove $\gamma_r = \lambda_r + \lambda_n \beta_r$ $(1 \leq r < n)$ e $L = \lfloor h^{2-1/4n} \rfloor$, soddisfi

$$\Phi_{m_0,\ldots,m_{n-1}}(l, \ldots, l) = 0 \tag{2.4}$$

per tutti gli interi l con $1 \leq l \leq h$ e tutti gli interi non negativi $m_0, \ldots, m_{n-1}$ con $m_0 + \cdots + m_{n-1} \leq h^2$.

Dimostrazione. Per l'ipotesi (2.1) si ha $\alpha_n = e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}}$. Sviluppiamo la derivata

$$\Phi_{m_0,\ldots,m_{n-1}}(z_0, \ldots, z_{n-1}) = \left(\frac{\partial}{\partial z_0}\right)^{m_0} \cdots \left(\frac{\partial}{\partial z_{n-1}}\right)^{m_{n-1}} \Phi(z_0, \ldots, z_{n-1})$$
$$= \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) \underbrace{\left(\frac{\partial}{\partial z_0}\right)^{m_0} (z_0^{\lambda_0} e^{\lambda_n \beta_0 z_0})}_{(A)} \underbrace{\prod_{r=1}^{n-1} \left(\frac{\partial}{\partial z_r}\right)^{m_r} (\alpha_r^{\gamma_r z_r})}_{(B)}. \tag{2.5}$$

Deriviamo la parte (A) applicando la regola di Leibniz per la derivata m_0-esima di un prodotto:

$$(A) = q(\lambda_0, \lambda_n, z_0) e^{\lambda_n \beta_0 z_0}$$

$$\text{con} \quad q(\lambda_0, \lambda_n, z_0) = \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1) z_0^{\lambda_0 - \mu_0} (\lambda_n \beta_0)^{m_0 - \mu_0}.$$

Calcoliamo la derivata della parte (B). Utilizzando la forma esponenziale $\alpha_r^{\gamma_r z_r} = e^{\gamma_r z_r \log \alpha_r}$ ricaviamo

$$\left(\frac{\partial}{\partial z_r}\right)^{m_r} (e^{\gamma_r z_r \log \alpha_r}) = (\gamma_r \log \alpha_r)^{m_r} \alpha_r^{\gamma_r z_r},$$

$$\text{per cui} \quad (B) = (\gamma_1 \log \alpha_1)^{m_1} \cdots (\gamma_{n-1} \log \alpha_{n-1})^{m_{n-1}} \alpha_1^{\gamma_1 z_1} \cdots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}}.$$

Valutiamo la derivata completa nel punto $(l, \ldots, l)$ sostituendo $\gamma_r = \lambda_r + \lambda_n \beta_r$ e l'espressione di α_n:

$$\begin{aligned}
\Phi_{m_0,\ldots,m_{n-1}}(l, \ldots, l) &= \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) q(\lambda_0, \lambda_n, l) e^{\lambda_n \beta_0 l} \alpha_1^{\gamma_1 l} \cdots \alpha_{n-1}^{\gamma_{n-1} l} \\
&\quad (\gamma_1 \log \alpha_1)^{m_1} \cdots (\gamma_{n-1} \log \alpha_{n-1})^{m_{n-1}} \\
&= \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) q(\lambda_0, \lambda_n, l) (e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}})^{\lambda_n l} \\
&\quad \alpha_1^{\lambda_1 l} \cdots \alpha_{n-1}^{\lambda_{n-1} l} (\gamma_1 \log \alpha_1)^{m_1} \cdots (\gamma_{n-1} \log \alpha_{n-1})^{m_{n-1}} \\
&= \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) q(\lambda_0, \lambda_n, l) \alpha_1^{\lambda_1 l} \cdots \alpha_n^{\lambda_n l} \\
&\quad \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} \log \alpha_1^{m_1} \cdots \log \alpha_{n-1}^{m_{n-1}}.
\end{aligned}$$

Poiché per ipotesi gli α_r sono numeri algebrici non nulli, i cui logaritmi sono linearmente indipendenti su $\mathbb{Q}$ e quindi non nulli, la condizione (2.4) risulta verificata se e solo se

$$\sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) q(\lambda_0, \lambda_n, l) \alpha_1^{\lambda_1 l} \cdots \alpha_n^{\lambda_n l} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} = 0. \tag{2.6}$$

Per trasformare l'equazione (2.6) in un sistema lineare a coefficienti interi (al fine di applicare il Lemma B1) introduciamo il fattore

$$P' = (a_1 \cdots a_n)^{Ll} b_0^{m_0} \cdots b_{n-1}^{m_{n-1}}. \tag{2.7}$$

Dalle proprietà dei numeri algebrici sappiamo che $(a_r \alpha_r)^j$ e $(b_r \beta_r)^j$ $(j \geq 0)$ sono interi algebrici. L'esponente Ll corrisponde al grado massimo di α_r nell'equazione,

mentre m_r compensa le potenze di β_r presenti in $q(\lambda_0, \lambda_n, l)$ e in γ_r. Moltiplicando l'equazione per P' si ottiene:

$$\sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) \underbrace{q(\lambda_0, \lambda_n, l) b_0^{m_0}}_{(C)} \underbrace{\alpha_1^{\lambda_1 l} a_1^{Ll} \cdots \alpha_n^{\lambda_n l} a_n^{Ll}}_{(D)} \cdot \underbrace{\gamma_1^{m_1} b_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} b_{n-1}^{m_{n-1}}}_{(E)} = 0. \tag{2.8}$$

Studiamo i singoli termini sostituendo a ciascun intero algebrico la relativa espressione definita in (2.2).

$$\begin{aligned}
(C) \quad &= \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1) l^{\lambda_0 - \mu_0} \lambda_n^{m_0 - \mu_0} (\beta_0 b_0)^{m_0 - \mu_0} b_0^{\mu_0} \\
&= \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1) l^{\lambda_0 - \mu_0} \lambda_n^{m_0 - \mu_0} \sum_{t_0=0}^{d-1} b_{0,t_0}^{(m_0 - \mu_0)} \beta_0^{t_0} b_0^{\mu_0} \\
&= \sum_{\mu_0=0}^{m_0} \sum_{t_0=0}^{d-1} q''' \beta_0^{t_0} \\
\text{con} \quad q''' &= \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1) l^{\lambda_0 - \mu_0} \lambda_n^{m_0 - \mu_0} b_0^{\mu_0} b_{0,t_0}^{(m_0 - \mu_0)}.
\end{aligned}$$

$$\begin{aligned}
(D) \quad &= (a_1 \alpha_1)^{\lambda_1 l} a_1^{(L - \lambda_1) l} \cdots (a_n \alpha_n)^{\lambda_n l} a_n^{(L - \lambda_n) l} \\
&= \sum_{s_1=0}^{d-1} a_{1,s_1}^{(\lambda_1 l)} \alpha_1^{s_1} a_1^{(L - \lambda_1) l} \cdots \sum_{s_n=0}^{d-1} a_{n,s_n}^{(\lambda_n l)} \alpha_n^{s_n} a_n^{(L - \lambda_n) l} \\
&= \sum_{s_1=0}^{d-1} \cdots \sum_{s_n=0}^{d-1} \alpha_1^{s_1} \cdots \alpha_n^{s_n} q' \\
\text{con} \quad q' &= \prod_{r=1}^{n} a_{r,s_r}^{(\lambda_r l)} a_r^{(L - \lambda_r) l}.
\end{aligned}$$

$$\begin{aligned}
(E) \quad &= \prod_{r=1}^{n-1} (\gamma_r b_r)^{m_r} = \prod_{r=1}^{n-1} (\lambda_r + \lambda_n \beta_r)^{m_r} b_r^{m_r} \quad \text{(dal Teorema binomiale)} \\
&= \prod_{r=1}^{n-1} \sum_{\mu_r=0}^{m_r} \binom{m_r}{\mu_r} \lambda_r^{m_r - \mu_r} (\lambda_n \beta_r)^{\mu_r} b_r^{m_r} \\
&= \prod_{r=1}^{n-1} \sum_{\mu_r=0}^{m_r} \binom{m_r}{\mu_r} \lambda_r^{m_r - \mu_r} \lambda_n^{\mu_r} \sum_{t_r=0}^{d-1} b_{r,t_r}^{(\mu_r)} \beta_r^{t_r} b_r^{m_r - \mu_r}
\end{aligned}$$

$$= \sum_{\mu_1=0}^{m_1} \cdots \sum_{\mu_{n-1}=0}^{m_{n-1}} \sum_{t_1=0}^{d-1} \cdots \sum_{t_{n-1}=0}^{d-1} \beta_1^{t_1} \cdots \beta_{n-1}^{t_{n-1}} q''$$

$$\text{con} \quad q'' = \prod_{r=1}^{n-1} \binom{m_r}{\mu_r} \lambda_r^{m_r-\mu_r} \lambda_n^{\mu_r} b_{r,t_r}^{(\mu_r)} b_r^{m_r-\mu_r}.$$

L'equazione (2.8) diventa

$$\sum_{s_1=0}^{d-1} \cdots \sum_{s_n=0}^{d-1} \sum_{t_0=0}^{d-1} \cdots \sum_{t_{n-1}=0}^{d-1} A(s,t) \alpha_1^{s_1} \cdots \alpha_n^{s_n} \beta_0^{t_0} \cdots \beta_{n-1}^{t_{n-1}} = 0$$

$$\text{dove} \quad A(s,t) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} \sum_{\mu_0=0}^{m_0} \cdots \sum_{\mu_{n-1}=0}^{m_{n-1}} p(\lambda_0, \ldots, \lambda_n) q' q'' q'''.$$

La condizione (2.4) è verificata se $A(s,t) = 0$ per ogni combinazione degli indici $s_1, \ldots, s_n, t_0, \ldots, t_{n-1}$. Tale espressione corrisponde a un sistema lineare in $p(\lambda_0, \ldots, \lambda_n)$ a coefficienti interi. I $2n$ indici s_r e t_r variando tra 0 e $d-1$ determinano d^{2n} equazioni lineari per ogni fissato insieme di parametri $(l, m_0, \ldots, m_{n-1})$. Individuiamo un maggiorante U per i coefficienti del sistema. Tenendo conto delle disuguaglianze $l \leq h$, $\binom{m_r}{\mu_r} \leq 2^{m_r}$ e delle stime $|a_{r,s_r}^{(j)}|, |b_{r,t_r}^{(j)}| \leq c_1^j$ fornite in (2.3) si ottiene:

$$|q'| = \left| \prod_{r=1}^{n} a_{r,s_r}^{(\lambda_r l)} a_r^{(L-\lambda_r)l} \right| \leq \prod_{r=1}^{n} c_1^{Lh} a_r^{Lh} \leq c_2^{Lh}$$

$$\text{con} \quad c_2 \geq \max_{1 \leq r \leq n} \{a_r\}^n c_1.$$

$$|q''| = \left| \prod_{r=1}^{n-1} \binom{m_r}{\mu_r} \lambda_r^{m_r-\mu_r} \lambda_n^{\mu_r} b_{r,t_r}^{(\mu_r)} b_r^{m_r-\mu_r} \right| \leq \prod_{r=1}^{n-1} 2^{m_r} (L b_r)^{m_r-\mu_r} L^{\mu_r} c_1^{\mu_r}$$

$$\leq \prod_{r=1}^{n-1} (2 b_r c_1 L)^{m_r} \leq \prod_{r=1}^{n-1} (c_3 L)^{m_r}$$

$$\text{con} \quad c_3 \geq 2 c_1 \max_{0 \leq r < n} \{b_r\}.$$

$$|q'''| = \left| \binom{m_0}{\mu_0} \lambda_0 (\lambda_0 - 1) \cdots (\lambda_0 - \mu_0 + 1) l^{\lambda_0-\mu_0} \lambda_n^{m_0-\mu_0} b_0^{\mu_0} b_{0,t_0}^{(m_0-\mu_0)} \right|$$

(se $\mu_0 \leq \lambda_0$)

$$\leq 2^{m_0} \lambda_0^{\mu_o} \lambda_n^{m_0-\mu_0} h^{\lambda_0-\mu_0} b_0^{\mu_0} c_1^{m_0-\mu_0}$$

$$\leq 2^{m_0} L^{\mu_o} L^{m_0-\mu_0} h^L b_0^{\mu_0} c_1^{m_0-\mu_0}$$

$$\leq (2 b_0 c_1 L)^{m_0} h^L \leq (c_3 L)^{m_0} h^L$$

con c_3 definito come sopra.

(se $\mu_0 > \lambda_0$)
$q''' = 0$.

Combinando questi contributi e ricordando $m_0 + \cdots + m_{n-1} \leq h^2$ segue che i coefficienti del sistema in $p(\lambda_0, \ldots, \lambda_n)$ hanno valore assoluto

$$\begin{aligned} U &\leq (m_0+1)\cdots(m_{n-1}+1)c_2^{Lh}(c_3L)^{m_0}h^L\prod_{r=1}^{n-1}(c_3L)^{m_r} \\ &\leq 2^{m_0}\cdots 2^{m_{n-1}}c_2^{Lh}h^L(c_3L)^{m_0+\cdots+m_{n-1}} \\ &\leq (2c_3L)^{h^2}c_2^{Lh}h^L \leq (2c_3L)^{h^2}c_4^{Lh} \\ \text{con} \quad & c_4 \geq c_2 h^{1/h}. \end{aligned}$$

Se $h \geq 3$ si ha $h^{1/h} \leq 1,45$ possiamo dunque assumere $c_4 = 2c_2$. Inoltre $1 \leq l \leq h$ e $0 \leq m_r \leq h^2$ con $r = 0, \ldots, n-1$ allora esistono al massimo $h(h^2+1)^n$ combinazioni distinte di indici $l, m_0, \ldots, m_{n-1}$. Il numero complessivo M di equazioni ad essi corrispondenti soddisfa $M \leq d^{2n}h(h^2+1)^n$. Il numero di incognite $p(\lambda_0, \ldots, \lambda_n)$ è dato da $N = (L+1)^{n+1}$.
Sostituendo la definizione di $L = \lfloor h^{2-\frac{1}{4n}} \rfloor$ si verifica $N > M$:

$$\begin{aligned} N = (L+1)^{n+1} &> (h^{2-\frac{1}{4n}})^{n+1} = h^{2n+\frac{7}{4}-\frac{1}{4n}} \geq h^{2n+\frac{3}{2}}, \quad (n>0) \\ 2M &\leq 2d^{2n}h(h^2+1)^n \end{aligned}$$

Considerando N e $2M$ come funzioni in h:

$$\lim_{h\to\infty}\frac{N}{2M} \geq \lim_{h\to\infty}\frac{h^{2n+\frac{3}{2}}}{2d^{2n}h(h^2+1)^n} = \infty.$$

Risulta $N > 2M$ per h sufficientemente grande per cui $M/(N-M) < 1$.
Il Lemma B1 assicura l'esistenza di una soluzione intera non banale tale che il valore assoluto dei coefficienti sia limitato da $(NU)^{M/(N-M)}$. La stima si semplifica in $|p(\lambda_0, \ldots, \lambda_n)| \leq NU$. Per completare la dimostrazione occorre mostrare che tale quantità è limitata superiormente da e^{h^3}. Per $h \to \infty$, $h^{2-\frac{1}{4n}} = o(h^2)$ dunque per h abbastanza grande:

$$L + 1 \leq h^{2-\frac{1}{4n}} + 1 \leq h^2.$$

Segue

$$\begin{aligned} NU &\leq (L+1)^{n+1}(2c_3L)^{h^2}c_4^{Lh} \leq h^{2n+2}(2c_3L)^{h^2}c_4^{Lh} \\ &\leq h^{2n+2}(2c_3h^{2-\frac{1}{4n}})^{h^2}c_4^{h^{3-\frac{1}{4n}}}. \end{aligned}$$

In forma logaritmica

$$\log(NU) \leq (2n+2)\log h + h^2\log(2c_3) + (2-\frac{1}{4n})h^2\log h + (h^{3-\frac{1}{4n}})\log c_4.$$

Per un valore di h opportunamente grande, la crescita di h^3 eccede strettamente la somma di questi termini. Si conclude che $NU \leq e^{h^3}$ come richiesto. □

2.4 Algoritmo di estrapolazione

Lemma B3. *Siano $m_0, \dots, m_{n-1}$ interi non negativi tali che $m_0 + \cdots + m_{n-1} \le h^2$ e sia $f(z) := \Phi_{m_0,\dots,m_{n-1}}(z, \dots, z)$. Allora per ogni $z \in \mathbb{C}$ si ha $|f(z)| \le c_5^{h^3+L|z|}$. Inoltre, per ogni intero positivo l o $f(l) = 0$ oppure $|f(l)| > c_6^{-h^3-Ll}$.*

Dimostrazione. La funzione $f(z)$ è definita nel Lemma B2 come

$$f(z) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \dots, \lambda_n) q(\lambda_0, \lambda_n, z) \alpha_1^{\lambda_1 z} \cdots \alpha_n^{\lambda_n z} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} P$$
$$\text{con} \quad P = (\log \alpha_1)^{m_1} \cdots (\log \alpha_{n-1})^{m_{n-1}}.$$

Per stabilire il $|f(z)|$ studiamo le singole componenti:

$|p(\lambda_0, \dots, \lambda_n)| \quad \le e^{h^3}$, dal Lemma B2.

$$|q(\lambda_0, \lambda_n, z)| \quad \le \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} |\lambda_0||\lambda_0 - 1| \cdots |\lambda_0 - \mu_0 + 1||z^{\lambda_0 - \mu_0}||\lambda_n \beta_0|^{m_0 - \mu_0}$$

(se $\mu_0 \le \lambda_0$)

$$\le \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0^{\mu_0} |z|^{\lambda_0} |\lambda_n \beta_0|^{m_0 - \mu_0}$$
$$\le \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} L^{\mu_0} L^{m_0 - \mu_0} |z|^L |\beta_0|^{m_0 - \mu_0}$$
$$\le (2L|\beta_0|)^{m_0} |z|^L \le (2c_7 L)^{m_0} |z|^L$$

con $c_7 \ge |\beta_0|$;

(se $\mu_0 > \lambda_0$) $|q(\lambda_0, \lambda_n, z)| = 0$.

Essendo $|\alpha_i^{\lambda_i z}| = |e^{\lambda_i z \log \alpha_i}| \le e^{\lambda_i |z| |\log \alpha_i|} \le e^{L|z||\log \alpha_i|}$ segue che

$$|\alpha_1^{\lambda_1 z} \cdots \alpha_n^{\lambda_n z}| \quad \le e^{L|z| \sum_{i=1}^n |\log \alpha_i|} \le c_8^{L|z|} \quad \text{con} \quad c_8 \ge e^{\sum_{i=1}^n |\log \alpha_i|}.$$

Dalla definizione di γ_r si ha $|\gamma_r| \le \lambda_r + \lambda_n |\beta_r| \le L(1 + |\beta_r|)$. Allora

$$|P \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}| = \prod_{r=1}^{n-1} |\gamma_r \log \alpha_r|^{m_r} \le \prod_{r=1}^{n-1} (L(1 + |\beta_r|)|\log \alpha_r|)^{m_r}$$
$$\le (c_9 L)^{m_1 + \cdots + m_{n-1}}$$

con $c_9 \ge \sup_{1 \le r < n} \{(1 + |\beta_r|)|\log \alpha_r|\}$.

Combinando le stime ottenute si ricava

$$|f(z)| \leq \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} e^{h^3} (2c_7 L)^{m_0} |z|^L c_8^{L|z|} (c_9 L)^{m_1+\cdots+m_{n-1}}. \tag{2.9}$$

Osserviamo che il numero di addendi nella sommatoria è $(L+1)^{n+1}$. Data la relazione $L \leq L+1 \leq h^2$ risulta

$$(L+1)^{n+1} \leq h^{2n+2} = e^{(2n+2)\log h}.$$

La disuguaglianza complessiva può essere riscritta in forma esponenziale per confrontare gli ordini di grandezza rispetto a h e $L|z|$. Utilizzando l'ipotesi $m_0+\cdots+m_{n-1} \leq h^2$:

$$\begin{aligned} |f(z)| &\leq e^{(2n+2)\log h} e^{h^3} (\sup\{2c_7, c_9\} h^2)^{h^2} |z|^L c_8^{L|z|} \\ &\leq e^{(2n+2)\log h + h^2 \log(\sup\{2c_7, c_9\}) + 2h^2 \log h + h^3} (ec_8)^{L|z|} \\ &\leq (ec_8)^{A(h)+h^3+L|z|} \end{aligned}$$

con $A(h) = [(2n+2)\log h + h^2 \log(\sup\{2c_7, c_9\}) + 2h^2 \log h]$. Per $h \to \infty$, $A(h) = o(h^3)$ dunque per h sufficientemente grande (dal Lemma 2.1) $A(h) \leq h^3$ e

$$|f(z)| \leq (ec_8)^{2h^3+L|z|} \leq c_5^{h^3+L|z|} \quad \text{con} \quad c_5 \geq (ec_8)^2.$$

Per dimostrare la seconda affermazione introduciamo $f' = (\frac{P'}{P}) f(l)$ con P' definito in (2.7). Dal Lemma B2 si evince che f' è un intero algebrico nel campo $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_{n-1})$ quindi

$$\deg_{\mathbb{Q}}(K) \leq \prod_{i=1}^{n} \prod_{j=0}^{n-1} \deg_{\mathbb{Q}}(\alpha_i) \deg_{\mathbb{Q}}(\beta_j) \leq d^{2n}.$$

I coniugati di f', indicati con $f'^{(k)}$ per $k = 1, \ldots, d^{2n}$ (dove $f'^{(1)} = f'$), si definiscono sostituendo a ciascun α_r e β_r i rispettivi coniugati arbitrari. Le stime precedenti rimangono valide per $f'^{(k)}$ sostituendo i moduli con il massimo dei valori assoluti di tutti i coniugati $|\overline{\alpha_r}|$ e $|\overline{\beta_r}|$. Si ottengono in questo modo le nuove costanti c_7', c_8'. Definiamo $c_9' = \max_{1 \leq r < n}\{1 + |\overline{\beta_r}|\}$ poiché il termine $|\log \alpha_r|$ non compare in f'. Ricordando la definizione di $\mathcal{A} = \max_{\substack{1 \leq i \leq n \\ 0 \leq j \leq n-1}} \{h(\alpha_i), h(\beta_j)\}$ si determina la seguente stima per P':

$$P' = (a_1 \cdots a_n)^{Ll} b_0^{m_0} \cdots b_{n-1}^{m_{n-1}} \leq \mathcal{A}^{nLl} \mathcal{A}^{m_0+\cdots+m_{n-1}} \leq \mathcal{A}^{nLl+h^2}.$$

Per ogni $k \in \{1, \ldots, d^{2n}\}$ risulta

$$|f'^{(k)}| \leq e^{(2n+2)\log h} e^{h^3} (2c_7' L)^{m_0} l^L c_8'^{Ll} (c_9' L)^{m_1+\cdots+m_{n-1}} \mathcal{A}^{nLl+h^2}$$

Procedendo in modo analogo a quanto visto sopra si conclude che per h sufficientemente grande (dal Lemma 2.1)

$$|f'^{(k)}| \leq (ec'_8\mathcal{A}^n)^{2h^3+Ll} \leq c_{10}^{h^3+Ll} \quad \text{con } c_{10} \geq (ec'_8\mathcal{A}^n)^2.$$

Essendo $f' = (\frac{P'}{P})f(l)$ anche $|P'/P| \leq c_{10}^{h^3+Ll}$. Dalle proprietà degli interi algebrici se $f' \neq 0$ la sua norma ha valore assoluto almeno pari a 1:

$$|\mathcal{N}(f')| = |f'| \prod_{k=2}^{d^{2n}} |f'^{(k)}| \geq 1$$

Ne consegue che

$$|f'| \geq \frac{1}{\prod_{k=2}^{d^{2n}} |f'^{(k)}|} \geq (c_{10}^{-h^3-Ll})^{d^{2n}-1} \geq c_{10}^{-d^{2n}(h^3+Ll)}.$$

Si arriva alla formulazione finale

$$|f(l)| = \left|\frac{P}{P'}\right| |f'| \geq c_{10}^{-h^3-Ll} c_{10}^{-d^{2n}(h^3+Ll)} = c_{10}^{-(d^{2n}+1)(h^3+Ll)} > c_6^{-h^3-Ll}$$

con $c_6 > c_{10}^{d^{2n}+1}$. □

Lemma B4. *Sia $J \in \mathbb{Z}$ con $0 \leq J \leq (8n)^2$.*
Allora l'equazione (2.4) $\Phi_{m_0,\ldots,m_{n-1}}(l,\ldots,l) = 0$ *è soddisfatta per tutti gli $l \in \mathbb{Z}$ con $1 \leq l \leq h^{1+\frac{J}{8n}}$ e per tutti gli $m_0,\ldots,m_{n-1} \in \mathbb{N}$ tali che $m_0 + \cdots + m_{n-1} \leq \frac{h^2}{2^J}$.*

Dimostrazione. Si procede per induzione su J. Definiamo

$$R_J = \lfloor h^{1+\frac{J}{8n}} \rfloor, \quad S_J = \left\lfloor \frac{h^2}{2^J} \right\rfloor.$$

Se $J = 0$ il risultato è dimostrato dal Lemma B2.
Sia $K \in \mathbb{Z}$ con $0 \leq K < (8n)^2$. Supponiamo che la tesi sia valida per $J = 0,\ldots,K$ e verifichiamola per $J = K+1$.
Per ipotesi induttiva $f(l) = \Phi_{m_0,\ldots,m_{n-1}}(l,\ldots,l) = 0$ con $1 \leq l \leq R_K$ e $m_0 + \cdots + m_{n-1} \leq S_K$. Essendo $S_{K+1} \leq S_K$ e $R_{K+1} \geq R_K$ per estendere la tesi a $K+1$ è sufficiente mostrare che $f(l) = 0$ nell'intervallo $R_K < l \leq R_{K+1}$ quando $m_0 + \cdots + m_{n-1} \leq S_{K+1} \leq S_K$.
Sia $r \in \{1,\ldots,R_K\}$ e $m \in \{0,\ldots,S_{K+1}\}$ allora

$$\begin{aligned} f_m(r) &= \left[\left(\frac{\partial}{\partial z_0} + \cdots + \frac{\partial}{\partial z_{n-1}}\right)^m \Phi_{m_0,\ldots,m_{n-1}}(z_0,\ldots,z_{n-1})\right]_{z_0=\cdots=z_{n-1}=r} \\ &= \sum_{j_0+\cdots+j_{n-1}=m} \frac{m!}{j_0!\cdots j_{n-1}!} \Phi_{m_0+j_0,\ldots,m_{n-1}+j_{n-1}}(r,\ldots,r) \end{aligned}$$

L'ordine totale di derivazione è $\sum m_i + \sum j_i \leq 2S_{K+1}$ inoltre

$$2S_{K+1} = 2\left\lfloor \frac{h^2}{2^{K+1}} \right\rfloor \leq \left\lfloor \frac{2h^2}{2^{K+1}} \right\rfloor = S_K.$$

Quindi dall'ipotesi induttiva $f_m(r) = 0$ per ogni $r \in \{1, \ldots, R_K\}$. Segue che $f(z)$ è una funzione olomorfa con zeri nei punti $1, \ldots, R_K$ di molteplicità almeno S_{K+1}. Una funzione olomorfa con queste stesse proprietà è data da

$$F(z) := \{(z-1)\ldots(z-R_K)\}^{S_{K+1}}.$$

La funzione quoziente $g(z) := \frac{f(z)}{F(z)}$ è per costruzione olomorfa su tutto $\mathbb{C}$ e quindi intera.
Il Principio del Massimo Modulo afferma che se $g : \overline{\Omega} \to \mathbb{C}$ è olomorfa con $\Omega \subset \mathbb{C}$ aperto connesso e limitato allora $\max_{z\in\overline{\Omega}} |g(z)| = \max_{z\in\partial\Omega} |g(z)|$. Analogo risultato vale per il minimo se la funzione non ha zeri all'interno del dominio.
Per ogni raggio $R > R_{K+1}$ applicando il Principio del Massimo Modulo sul disco B_R si ottiene

$$\max_{z\in B_R}\left|\frac{f(z)}{F(z)}\right| = \max_{z\in\partial B_R}\left|\frac{f(z)}{F(z)}\right| \leq \frac{\theta}{\Theta}$$

dove θ, Θ rappresentano rispettivamente l'estremo superiore di $|f(z)|$ e l'estremo inferiore di $|F(z)|$ con $|z| = R$. Per ogni $l \in \mathbb{N}$ tale che $1 \leq l \leq R$ si ricava la disuguaglianza fondamentale

$$\Theta \cdot |f(l)| \leq \theta \cdot |F(l)|. \tag{2.10}$$

Valutiamo le stime di Θ e θ. Per $|z| = R$ si ha

$$|z - r| \geq |z| - |r| = R - r \geq R - R_K$$

e poiché $R > R_{K+1}$, $R - R_K \geq R/2$ per un h opportuno. Segue che

$$\begin{aligned}\Theta &= \min_{|z|=R} |F(z)| = \min_{|z|=R} (|z-1|\cdots|z-R_K|)^{S_{K+1}} \\ &\geq (R - R_K)^{R_K S_{K+1}} \geq (R/2)^{R_K S_{K+1}}\end{aligned}$$

mentre dal Lemma B3

$$\theta = \max_{|z|=R} |f(z)| \leq c_5^{h^3+LR}.$$

Per $l \leq R_{K+1}$ si verifica

$$|F(l)| = |(l-1)\cdots(l-R_K)|^{S_{K+1}} \leq R_{K+1}^{R_K S_{K+1}}.$$

Supponiamo $f(l) \neq 0$ per qualche $l \in (R_K, R_{K+1}]$, il Lemma B3 implica

$$|f(l)| > c_6^{-h^3-LR}.$$

La disuguaglianza (2.10) diventa

$$c_6^{-h^3-LR}\left(\frac{R}{2}\right)^{R_K S_{K+1}} \leq c_5^{h^3+LR} R_{K+1}^{R_K S_{K+1}}$$
$$\implies \left(\frac{R}{2R_{K+1}}\right)^{R_K S_{K+1}} \leq (c_5 c_6)^{h^3+LR}. \tag{2.11}$$

Analizziamo gli ordini di grandezza ricordando l'ipotesi $K < (8n)^2$. Siccome $R > R_{K+1}$ possiamo parametrizzare il raggio come $R = R_{K+1}h^\delta$ con $\delta > 0$ e la disuguaglianza (2.11) si riscrive come

$$\left(\frac{h^\delta}{2}\right)^{R_K S_{K+1}} \leq (c_5 c_6)^{h^3+LR}.$$

In forma logaritmica la componente di sinistra risulta

$$R_K S_{K+1}(\delta \log h - \log 2).$$

Dalle definizioni

$$R_K S_{K+1} > (h^{1+K/(8n)} - 1)\left(\frac{h^2}{2^{K+1}} - 1\right) > h^{3+K/(8n)}\frac{1}{2^{K+1}} - O(h^2). \quad \text{Dunque}$$
$$\log\left(\frac{h^\delta}{2}\right)^{R_K S_{K+1}} > h^{3+K/(8n)}\frac{1}{2^{K+1}}\delta \log h - O(h^2).$$

L'ordine di grandezza di LR è dato da

$$LR \leq h^{2-\frac{1}{4n}} h^{1+\delta+\frac{K+1}{8n}} = h^{3+\delta+\frac{K-1}{8n}},$$

di conseguenza il termine dominante in $h^3 + LR$ è LR e

$$\log(c_5 c_6)^{h^3+LR} \leq h^{3+\delta+\frac{K-1}{8n}} \log(c_5 c_6) + O(h^3).$$

Confrontando gli esponenti dominanti della variabile h nelle due componenti

$$h^{3+K/(8n)} \log h \quad \text{contro} \quad h^{3+\delta+\frac{K-1}{8n}},$$

per $\delta = 1/(8n)$ si ottiene

$$h^{3+K/(8n)} \log h \geq h^{3+\delta+\frac{K-1}{8n}}.$$

Scegliendo $R = R_{K+1}h^{1/(8n)}$ si giunge a una contraddizione nella disuguaglianza (2.11) e dal Lemma B3 si conclude che $f(l) = 0$ per tutti i valori di l considerati nel passo induttivo. □

Lemma B5. *Scrivendo $\phi(z) = \Phi(z, \ldots, z)$ si ha $|\phi_j(0)| < e^{-h^{8n}}$ per ogni $0 \leq j \leq h^{8n}$.*

Dimostrazione. Definiamo i parametri

$$X = h^{8n}, \quad Y = \left\lfloor \frac{h^2}{2^{(8n)^2}} \right\rfloor.$$

Dal Lemma B4 si ha che $\Phi_{m_0,\dots,m_{n-1}}(l,\dots,l) = 0$ per tutti gli $l \in \mathbb{Z}$ con $1 \le l \le X$ e gli $m_0,\dots,m_{n-1} \in \mathbb{N}$ che soddisfano $m_0 + \cdots + m_{n-1} \le Y$. Inoltre, $\phi_m(r) = 0$ per tutti gli interi $r \in \{1,\dots,X\}$ e $m \in \{0,\dots,Y\}$. Introduciamo la funzione

$$E(z) = \{(z-1)\cdots(z-X)\}^Y.$$

Segue che $\phi(z)/E(z)$ è una funzione olomorfa sul disco B_R di raggio $R = Xh^{\frac{1}{8n}}$ centrato nell'origine.
Sia $w \in \mathbb{C}$ un punto tale che $|w| < X$. Per il Principio del Massimo Modulo applicato alla funzione $\phi(z)/E(z)$ sulla chiusura del disco B_R si deduce

$$|\phi(w)| \le \xi \cdot \Xi^{-1} \cdot |E(w)| \tag{2.12}$$

dove ξ e Ξ sono definiti come l'estremo superiore di $|\phi(z)|$ e quello inferiore di $|E(z)|$ con $|z| = R$.
Se $|w| < X$ si osserva che per ogni $i \in \{1,\dots,X\}$ vale $|w-i| \le |w| + i < 2X$. Per cui

$$|E(w)| = \prod_{i=1}^{X} |w-i|^Y \le (2X)^{XY}.$$

Sulla frontiera di B_R dove $|z| = R$, per ogni $i \in \{1,\dots,X\}$ e per h opportunamente grande risulta

$$|z-i| \ge |z| - X = X(h^{\frac{1}{8n}} - 1) \ge \frac{Xh^{\frac{1}{8n}}}{2} = \frac{R}{2}$$
$$\implies |\Xi| \ge \left(\frac{R}{2}\right)^{XY}.$$

Dal Lemma B3 si ricava $\xi \le c_5^{h^3+LR}$. Sostituendo tali stime nella disequazione (2.12) si ottiene

$$|\phi(w)| \le c_5^{h^3+LR} \left(\frac{R}{2}\right)^{-XY} (2X)^{XY} = c_5^{h^3+LR} \left(\frac{4}{h^{\frac{1}{8n}}}\right)^{XY}. \tag{2.13}$$

Il prodotto LR può essere maggiorato come segue:

$$LR \le h^{2-\frac{1}{4n}} h^{8n+\frac{1}{8n}} \le h^{8n+2} = X\frac{h^2}{2^{(8n)^2}} 2^{(8n)^2} \le XY 2^{(8n)^2+1}.$$

Di conseguenza, in termini logaritmici la relazione (2.13) assume la forma

$$\log|\phi(w)| \le h^3 \log(c_5) + XY 2^{(8n)^2+1} \log(c_5) - XY \log\left(h^{\frac{1}{8n}} \frac{1}{4}\right).$$

Per $h \to \infty$ essendo $XY = h^{8n}\left\lfloor \frac{h^2}{2^{(8n)^2}} \right\rfloor$ di ordine h^{8n+2} e $\log h \to \infty$ possiamo scrivere

$$\log|\phi(w)| \le -\frac{1}{8n} XY \log h + O(h^{8n+2}) \le -XY$$
$$\implies |\phi(w)| \le e^{-XY} \quad \text{per } h \text{ sufficientemente grande (dal Lemma 2.1).}$$

Applicando la formula integrale di Cauchy alla funzione olomorfa ϕ sul dominio $\Lambda = \{w \in \mathbb{C} : |w| = 1\}$, si ottiene

$$\phi_j(0) = \frac{j!}{2\pi i} \int_\Lambda \frac{\phi(w)}{w^{j+1}}\, dw.$$

Tramite la disuguaglianza di Darboux, posto $L(\Lambda) = 2\pi$ per la lunghezza della curva Λ, risulta

$$|\phi_j(0)| \le \frac{j!}{2\pi} \cdot L(\Lambda) \cdot \max_{|w|=1} \left| \frac{\phi(w)}{w^{j+1}} \right| \le j! e^{-XY} \le j^j e^{-XY}.$$

Per verificare la tesi ricordiamo che per ipotesi $j \le h^{8n}$ dunque

$$j^j e^{-XY} \le (h^{8n})^{h^{8n}} e^{-XY} = e^{8nh^{8n}\log(h) - XY}.$$

Poiché il termine dominante è XY di ordine h^{8n+2}, per h abbastanza grande si conclude $|\phi_j(0)| < e^{-h^{8n}+2} < e^{-h^{8n}}$. □

2.5 Strumenti supplementari

Lemma B6. *Per ogni insieme di interi $t_1, \dots, t_n$ non tutti nulli e con valore assoluto al più T si ha*

$$|t_1 \log \alpha_1 + \cdots + t_n \log \alpha_n| > c_{11}^{-T}.$$

Dimostrazione. Sia a_j con $1 \le j \le n$ il coefficiente principale del polinomio minimo di α_j se $t_j \ge 0$ o di α_j^{-1} se $t_j < 0$ e definiamo

$$w := a_1^{|t_1|} \dots a_n^{|t_n|} (\alpha_1^{t_1} \dots \alpha_n^{t_n} - 1).$$

Poiché ogni termine del tipo $a_j^{|t_j|}\alpha_j^{t_j}$ è un intero algebrico, w è esso stesso un intero algebrico in $K := \mathbb{Q}(\alpha_1, \dots, \alpha_n)$. Il grado di w è limitato dal grado dell'estensione $[K : \mathbb{Q}]$ che soddisfa

$$[K : \mathbb{Q}] \le \prod_{j=1}^{n} \deg_{\mathbb{Q}}(\alpha_j) \le d^n.$$

I coniugati di w, indicati con $w^{(k)}$ per $k = 1, \dots, d^n$ (con $w^{(1)} = w$), si definiscono sostituendo a ciascun α_r un suo coniugato arbitrario $\tilde{\alpha}_r$. Per ogni $k \in \{1, \dots, d^n\}$

risulta

$$\begin{aligned} |w^{(k)}| &= a_1^{|t_1|} \dots a_n^{|t_n|} |\tilde{\alpha_1}^{t_1} \dots \tilde{\alpha_n}^{t_n} - 1| \\ &\leq (a_1 \dots a_n)^T (|\tilde{\alpha_1}|^T \dots |\tilde{\alpha_n}|^T + 1) \leq c_{12}^T \end{aligned} \tag{2.14}$$

$$\text{con } c_{12} = a_1 \dots a_n \sup(|\tilde{\alpha_1}| \dots |\tilde{\alpha_n}| + 1).$$

Se $w = 0$ dalla definizione si evince che $\alpha_1^{t_1} \cdots \alpha_n^{t_n} = 1$ e in forma esponenziale $e^{t_1 \log \alpha_1 + \cdots + t_n \log \alpha_n} = e^0 = 1$ per cui

$$\Omega := t_1 \log \alpha_1 + \cdots + t_n \log \alpha_n = p(2\pi i), \quad p \in \mathbb{Z}.$$

I logaritmi $\log \alpha_1, \dots, \log \alpha_n$ sono per ipotesi linearmente indipendenti su $\mathbb{Q}$ e i coefficienti t_j non sono tutti nulli, di conseguenza $p \neq 0$ e $|\Omega| \geq 2\pi$. In questo caso la tesi $|\Omega| > c_{11}^{-T}$ è verificata per ogni $c_{11} \geq 1$ dato che $2\pi > 1 \geq c_{11}^{-T}$ $(T \geq 1)$.
Consideriamo $w \neq 0$, essendo w un intero algebrico non nullo, la sua norma $\mathcal{N}(w)$ soddisfa

$$|\mathcal{N}(w)| = |w| \prod_{i=2}^{d^n} |w^{(i)}| \geq 1.$$

Utilizzando la stima (2.14) si ricava

$$|w| \geq \frac{1}{\prod_{i=2}^{d^n} |w^{(i)}|} \geq c_{12}^{-T(d^n-1)} > c_{12}^{-Td^n}.$$

Per ogni $z \in \mathbb{C}$ tramite lo sviluppo in serie di Taylor vale la disuguaglianza

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

Esprimiamo w in funzione di Ω tramite la relazione $e^{\Omega} = \alpha_1^{t_1} \cdots \alpha_n^{t_n}$:

$$\begin{gathered} w = a_1^{|t_1|} \dots a_n^{|t_n|} (e^{\Omega} - 1) \quad \text{da cui ponendo } c_{13} = a_1 \cdots a_n \text{ segue} \\ |w| \leq (a_1 \dots a_n)^T |\Omega| e^{|\Omega|} \leq c_{13}^T |\Omega| e^{|\Omega|}. \end{gathered}$$

Se $|\Omega| \geq 1$ la tesi è banalmente vera per $c_{11} > 1$. Assumendo $|\Omega| < 1$ e confrontando le stime ottenute si ha

$$\begin{gathered} |w| > c_{12}^{-Td^n} \quad \text{e} \quad |w| \leq c_{13}^T |\Omega| e^{|\Omega|} \\ \implies |\Omega| > \frac{(c_{12}^{d^n} c_{13})^{-T}}{e^{|\Omega|}} > (c_{12}^{d^n} c_{13})^{-T} e^{-1} > (c_{12}^{d^n} c_{13} e)^{-T} > c_{11}^{-T} \end{gathered}$$

definendo $c_{11} = c_{12}^{d^n} c_{13} e$. □

Lemma B7. *Siano $R, S \in \mathbb{Z}$ positivi e $\sigma_0, \dots, \sigma_{R-1} \in \mathbb{C}$ distinti. Sia σ il massimo tra $1, |\sigma_0|, \dots, |\sigma_{R-1}|$ e ρ il minimo tra 1 e i valori $|\sigma_i - \sigma_j|$ con $0 \le i < j < R$. Allora per ogni coppia $r, s \in \mathbb{Z}$ con $0 \le r < R$, $0 \le s < S$ esistono $w_i \in \mathbb{C}$ con $0 \le i < RS$ e $|w_i| \le (8\sigma/\rho)^{RS}$ tali che il polinomio*

$$W(z) = \sum_{j=0}^{RS} w_j z^j \in \mathbb{C}(z)$$

soddisfi $W_j(\sigma_i) = 0$ per ogni coppia $(i,j) \neq (r,s)$ con $0 \le i < R$ e $0 \le j < S$ e $W_s(\sigma_r) = 1$.

Dimostrazione. Definiamo la funzione ausiliaria $U(z) \coloneqq \{(z-\sigma_0)\dots(z-\sigma_{R-1})\}^S$ di grado SR e il polinomio

$$W(z) \coloneqq \left(\frac{-1}{s!}\right) \frac{1}{2\pi i} \int_{C_r} \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta - z)U(\zeta)}\, d\zeta$$

dove C_r denota la circonferenza di centro σ_r e raggio sufficientemente piccolo, minore di ρ e di $|z-\sigma_r|$ per $z \neq \sigma_r$.
In generale dato un insieme di nodi $x_1, \dots, x_n \in \mathbb{C}$ con molteplicità m_i, la funzione $\Omega(z) \coloneqq \prod_{i=1}^n (z-x_i)^{m_i}$ ne racchiude la struttura differenziale. Per una generica funzione olomorfa $f(z)$ il polinomio di interpolazione di Hermite $P(z)$ di grado al più $(\sum_{i=1}^n m_i) - 1$ che coincide con $f(z)$ e le sue derivate fino all'ordine $m_i - 1$ nei punti assegnati ammette la seguente rappresentazione

$$P(z) = \frac{1}{2\pi i} \int_\Gamma \frac{\Omega(z) - \Omega(\xi)}{\Omega(z)} \frac{f(\xi)}{\xi - z}\, d\xi$$

dove Γ è un circuito chiuso semplice che contiene i nodi x_i e z.
$W(z)$ ha la forma di un polinomio di interpolazione di Hermite e nel seguito verificheremo che soddisfa le condizioni richieste dall'enunciato.
La dimostrazione si basa su due espressioni alternative per $W(z)$.
Il Teorema dei residui di Cauchy afferma che se $f : \Omega \to \mathbb{C}$ è olomorfa in $\Omega \subseteq \mathbb{C}$ aperto, semplicemente connesso ad eccezione di un numero finito di punti singolari isolati $a_1, \dots, a_k$ all'interno di una curva chiusa semplice γ allora

$$\int_\gamma f(z)\, dz = 2\pi i \sum_{j=1}^k Res(f, a_j).$$

Consideriamo la funzione

$$f(\zeta) \coloneqq \frac{(\zeta-\sigma_r)^s U(z)}{(\zeta - z)U(\zeta)}$$

che ha poli in $z, \sigma_0, \dots, \sigma_{R-1}$. Poiché gli ordini di grandezza rispetto a ζ del numeratore e denominatore sono rispettivamente ζ^s e ζ^{SR+1}, la condizione $RS > S > s$ assicura che

$$\lim_{\zeta \to \infty} |f(\zeta)||\zeta| = 0.$$

Di conseguenza dal Lemma del grande arco applicato al Teorema dei residui si ottiene che la somma dei residui di $f(\zeta)$ nelle singolarità isolate è nulla:

$$Res(f,z) + \sum_{j=0}^{R-1} Res(f,\sigma_j) = 0. \tag{2.15}$$

Per definizione $W(z) = \frac{-1}{s!} Res(f,\sigma_r)$. Essendo z un polo semplice il relativo residuo si calcola come

$$Res(f,z) = \lim_{\zeta\to\infty} (\zeta - z) f(\zeta) = \lim_{\zeta\to\infty} \frac{(\zeta-\sigma_r)^s U(z)}{U(\zeta)} = (z-\sigma_r)^s.$$

Sostituendo queste espressioni in (2.15) otteniamo

$$\begin{aligned} W(z) &= \frac{-1}{s!}\left[-(z-\sigma_r)^s - \sum_{j\neq r, j=0}^{R-1} Res(f,\sigma_j)\right] \\ &= \frac{(z-\sigma_r)^s}{s!} + \frac{U(z)}{s!}\frac{1}{2\pi i}\sum_{j\neq r, j=0}^{R-1}\int_{C_j} \frac{(\zeta-\sigma_r)^s}{(\zeta-z)U(\zeta)}\,d\zeta \end{aligned}$$

dove C_j è la circonferenza di centro σ_j e raggio sufficientemente piccolo come in C_r. In $z = \sigma_r$ la somma su j è regolare e siccome $U(z)$ ha uno zero di ordine S in σ_r, tutte le sue derivate $U_j(\sigma_r)$ si annullano per $0 \le j < S$. Segue che l'unica componente non nulla di $W_j(\sigma_r)$ deriva dal termine $\frac{(z-\sigma_r)^s}{s!}$. Si ricava

$$\left[\frac{\partial^j}{\partial z^j}\frac{(z-\sigma_r)^s}{s!}\right]_{z=\sigma_r} = \left[\frac{s(s-1)\dots(s-j+1)(z-\sigma_r)^{s-j}}{s!}\right]_{z=\sigma_r} = \begin{cases} 0 & j \neq s \\ \frac{s!}{s!} = 1 & j = s \end{cases}$$

Si evince che $W_j(\sigma_r) = 1$ se $j = s$ e 0 altrimenti.
Poniamo $t = S - s - 1$. La formula integrale di Cauchy per le derivate di ordine superiore è

$$g^{(t)}(a) = \frac{t!}{2\pi i}\int_\gamma \frac{g(\zeta)}{(\zeta-a)^{t+1}}\,d\zeta$$

per cui $W(z)$ risulta

$$\begin{aligned} W(z) &= \left(\frac{-1}{s!}\right)\frac{1}{2\pi i}\int_{C_r}\left[\frac{U(z)(\zeta-\sigma_r)^S}{(\zeta-z)U(\zeta)}\right]\frac{1}{(\zeta-\sigma_r)^{t+1}}\,d\zeta \\ &= \frac{-1}{s!t!}\frac{d^t}{d\zeta^t}\left[\frac{U(z)(\zeta-\sigma_r)^S}{(\zeta-z)U(\zeta)}\right]_{\zeta=\sigma_r} = \frac{-U(z)}{s!t!}\frac{d^t}{d\zeta^t}\left[\frac{(\zeta-\sigma_r)^S}{(\zeta-z)U(\zeta)}\right]_{\zeta=\sigma_r} \end{aligned}$$

Calcoliamo la derivata t-esima del prodotto sfruttando la regola di Leibniz:

$$\begin{aligned}
&\frac{d^t}{d\zeta^t}\left[\frac{(\zeta-\sigma_r)^S}{(\zeta-z)U(\zeta)}\right]_{\zeta=\sigma_r} = \frac{d^t}{d\zeta^t}\left[(\zeta-z)^{-1}\prod_{i=0,i\neq r}^{R-1}(\zeta-\sigma_j)^{-S}\right]_{\zeta=\sigma_r}\\
&= \sum_{j_0+\cdots+j_{R-1}=t}\binom{t}{j_0,\ldots,j_{R-1}}(-1)^{j_r}j_r!(\sigma_r-z)^{-1-j_r}\\
&\quad\prod_{i=0,i\neq r}^{R-1}(-1)^{j_i}j_i!\binom{S+j_i-1}{j_i}(\sigma_r-\sigma_i)^{-S-j_i}\\
&= t!(-1)^t\sum_{j_0+\cdots+j_{R-1}=t}v(j_0,\ldots,j_{R-1})(\sigma_r-z)^{-1-j_r}
\end{aligned}$$

dove $v(j_0,\ldots,j_{R-1}) = \prod_{i=0,i\neq r}^{R-1}\binom{S+j_i-1}{j_i}(\sigma_r-\sigma_i)^{-S-j_i}$. Allora

$$W(z) = (-1)^{t-1}(s!)^{-1}U(z)\sum_{j_0+\cdots+j_{R-1}=t}v(j_0,\ldots,j_{R-1})(\sigma_r-z)^{-1-j_r}.$$

Dato che $j_r+1\geq 1$, $W(z)\in\mathbb{C}(z)$ ha grado al più $RS-1$.
In $z=\sigma_i$ $(i\neq r)$ $W(z)$, come $U(z)$, ha uno zero di ordine S, questo garantisce che $W_j(\sigma_i)=0$ per ogni $j<S$.
Dalla definizione di ρ e la proprietà $\binom{n}{k}\leq 2^n$ si ha

$$\begin{aligned}
|v(j_0,\ldots,j_{R-1})| &\leq \prod_{i=0,i\neq r}^{R-1}2^{S+j_i-1}\rho^{-S-j_i} \leq \prod_{i=0,i\neq r}^{R-1}\left(\frac{2}{\rho}\right)^{S+j_i}\\
&\leq \left(\frac{2}{\rho}\right)^{(R-1)S+j_0+\cdots+j_{R-1}} \leq \left(\frac{2}{\rho}\right)^{RS}.
\end{aligned}$$

Si consideri il termine $Q(z)=(\sigma_r-z)^{-j_r-1}U(z)$. Osserviamo che

$$Q(z) = \prod_{i=0,i\neq r}^{R-1}(z-\sigma_i)^S(z-\sigma_r)^{S-j_r-1}(-1)^{j_r+1}.$$

Essendo $j_r+1\leq t+1 = S-s\leq S$, $Q(z)$ è un polinomio monico, a meno del segno, di grado $N=RS-j_r-1<RS$. Per le relazioni di Viète i coefficienti di un polinomio monico sono funzioni simmetriche delle radici esprimibili come somme di $\binom{N}{j}$ prodotti di j radici. Ogni radice σ_i di $Q(z)$ ha modulo minore di σ quindi i coefficienti hanno valore assoluto limitato superiormente da

$$\sum_{j=0}^{N}\binom{N}{j}\sigma^j = (1+\sigma)^N.$$

Il numero di termini nella sommatoria di $W(z)$ non supera S^R poiché ogni j_i ha al più S scelte. Dunque $W(z)=\sum_{j=0}^{RS-1}w_jz^j$ e i suoi coefficienti soddisfano

$$|w_j| \leq S^R(\sigma+1)^{RS}\left(\frac{2}{\rho}\right)^{RS} \leq \left(\frac{2(\sigma+1)S^{1/S}}{\rho}\right)^{RS} \leq \left(\frac{8\sigma}{\rho}\right)^{RS}$$

in quanto $S^{1/S} \leq 2$ e $\sigma + 1 \leq 2\sigma$. Si conclude che il polinomio $W(z)$ costruito verifica le condizioni richieste. □

Capitolo 3

Teorema di Baker e conseguenze

Si dispone ora di tutti gli elementi necessari per la dimostrazione del Teorema di Baker 1.8 che sarà l'obiettivo di questo capitolo. Proponiamo di seguito una sintesi del percorso logico seguito, prima di passare alla trattazione formale.
La struttura della dimostrazione si basa su un procedimento per assurdo ipotizzando l'esistenza di una relazione lineare tra $1, \log \alpha_1, \ldots, \log \alpha_n$ sul campo dei numeri algebrici.
In primo luogo viene riscritta la funzione ausiliaria $\phi(z)$ introducendo delle nuove grandezze S e R, strettamente dipendenti da un intero h e sfruttando l'espressione di α_n derivata dall'ipotesi iniziale. Associando univocamente ogni indice i a una combinazione di parametri mediante la scomposizione in base S si ottiene l'espressione

$$\phi(z) = \sum_{i=0}^{RS-1} p_i z^{\nu_i} e^{\psi_i z}.$$

Successivamente in virtù dell'indipendenza lineare dei logaritmi sul campo razionale e del Lemma B6 si prova l'esistenza di R valori distinti per le combinazioni ψ_i. Questo passaggio insieme al Lemma B7 permette di definire rigorosamente le proprietà di un polinomio $W(z)$ e dei suoi parametri w_j.
La ricerca della contraddizione risiede in un coefficiente intero non nullo della funzione ausiliaria, denominato p_t. Applicando il Teorema di Leibniz per le derivate successive del prodotto di funzioni, Baker riesce a esprimere questo termine come una combinazione lineare delle derivate nell'origine della funzione ausiliaria, $\phi_j(0)$ e dei parametri w_j:

$$p_t = \sum_{j=0}^{RS-1} w_j \phi_j(0).$$

Infine si ricorre a specifiche stime analitiche per valutare il modulo di p_t. Mentre per costruzione p_t è un intero non nullo (e dunque $|p_t| \geq 1$), le maggiorazioni teoriche sulle derivate del Lemma B5 e sui coefficienti del Lemma B7 mostrano che, per un valore di h sufficientemente grande, si otterrebbe $\log |p_t| < 0$. Questa incoerenza logica invalida l'ipotesi iniziale confermando la tesi del teorema.

Il capitolo si chiude con l'analisi dei principali corollari discendenti dalla dimostrazione e delle prospettive di ricerca future nella teoria dei numeri trascendenti.

3.1 Dimostrazione del teorema principale

Ricordiamo l'enunciato del Teorema di Baker:

Siano $\alpha_1, \ldots, \alpha_n$ numeri algebrici non nulli tali che $\log \alpha_1, \ldots, \log \alpha_n$ siano linearmente indipendenti sui razionali. Allora $1, \log \alpha_1, \ldots, \log \alpha_n$ sono linearmente indipendenti sul campo di tutti i numeri algebrici.

Dimostrazione. Supponiamo per assurdo che esistano $\beta_1, \ldots, \beta_n \in \overline{\mathbb{Q}}$ tali che $\beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n = 0$. Senza perdita di generalità, assumendo $\beta_n \neq 0$ possiamo isolare l'ultimo termine e riformulare l'espressione come segue

$$\alpha_n = e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}}.$$

Consideriamo la funzione ausiliaria ϕ costruita nel capitolo precedente. Dal Lemma B5 si evince

$$|\phi_j(0)| < e^{-h^{8n}} \quad \forall\, 0 \leq j \leq h^{8n}. \tag{3.1}$$

Definiamo i parametri $S = L + 1$ e $R = S^n$. Ogni intero i nell'intervallo $0 \leq i < RS = S^{n+1}$ ammette una scomposizione univoca in base S:

$$i = \lambda_0 + \lambda_1 S + \cdots + \lambda_n S^n, \quad \text{con } \lambda_i \in \{0, \ldots, S - 1 = L\}.$$

Per ogni indice i poniamo $\nu_i = \lambda_0$, $p_i = p(\lambda_0, \ldots, \lambda_n)$ e $\psi_i = \lambda_1 \log \alpha_1 + \cdots + \lambda_n \log \alpha_n$. Di conseguenza

$$\begin{aligned} \phi(z) = \Phi(z, \ldots, z) &= \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\lambda_0, \ldots, \lambda_n) z^{\lambda_0} e^{\lambda_n \beta_0 z} \alpha_1^{\gamma_1 z} \cdots \alpha_{n-1}^{\gamma_{n-1} z} \\ &= \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p_i z^{\nu_i} e^{\lambda_n \beta_0 z + \sum_{r=1}^{n-1} (\lambda_r + \lambda_n \beta_r) z \log \alpha_r}. \end{aligned}$$

Dall'espressione di α_n:

$$\begin{aligned} \sum_{r=1}^{n-1} \lambda_r \log \alpha_r + \lambda_n (\beta_0 + \sum_{r=1}^{n-1} \beta_r \log \alpha_r) = \sum_{r=1}^{n-1} \lambda_r \log \alpha_r + \lambda_n \log \alpha_n = \psi_i \\ \implies \phi(z) = \sum_{i=0}^{RS-1} p_i z^{\nu_i} e^{\psi_i z}. \end{aligned} \tag{3.2}$$

Siano $(\lambda_1, \ldots, \lambda_n)$ e $(\lambda'_1, \ldots, \lambda'_n)$ due vettori distinti associati rispettivamente a ψ_i e ψ_j. La loro differenza è espressa da

$$\psi_i - \psi_j = (\lambda_1 - \lambda'_1) \log \alpha_1 + \cdots + (\lambda_n - \lambda'_n) \log \alpha_n.$$

Sostituendo $t_k = \lambda_k - \lambda'_k$ si ottiene $|t_k| \leq L$ per ogni k. Siccome i vettori sono distinti, i coefficienti t_k non sono tutti nulli e il Lemma B6 garantisce che $|\psi_i - \psi_j| > c_{11}^{-L}$. Il numero complessivo di combinazioni $(\lambda_1, \ldots, \lambda_n)$ è pari a $S^n = R$ poiché ciascuna coordinata λ_k può assumere $L + 1 = S$ valori. A n-uple distinte corrispondono valori distinti di ψ_i in virtù dell'indipendenza lineare dei logaritmi $\log \alpha_1, \ldots, \log \alpha_n$ su $\mathbb{Q}$. Dunque esistono esattamente R valori distinti per ψ_i e li indichiamo come $\sigma_0, \ldots, \sigma_{R-1}$.
Introducendo σ e ρ come nel Lemma B7 si ha che

$$\sigma = \max\{1, |\sigma_0|, \ldots, |\sigma_{R-1}|\} \quad \text{dove } \sigma_i = \lambda_1 \log \alpha_1 + \cdots + \lambda_n \log \alpha_n \quad \text{e } 0 \leq \lambda_i \leq L$$
$$\implies |\sigma_i| \leq \sum_{k=1}^{n} \lambda_k |\log \alpha_k| \leq L c_{14} \quad \text{con } c_{14} = \sum_{k=1}^{n} |\log \alpha_k|;$$
$$\rho = \min_{0 \leq i < j < R} \{1, |\sigma_i - \sigma_j|\} \quad \text{e} \quad |\sigma_i - \sigma_j| > c_{15}^{-L} \quad \text{con } c_{15} = c_{11}$$

allora $\sigma \leq L c_{14}$ e $\rho \geq c_{15}^{-L}$.
Fissato un indice t per cui $p_t \neq 0$, poniamo $s = \nu_t$ e sia $r \in \{0, \ldots, R-1\}$ l'unico indice tale che $\sigma_r = \psi_t$. Le proprietà del polinomio $W(z)$, definito nel Lemma B7, implicano che

$$p_t = \sum_{i=o}^{RS-1} p_i W_{\nu_i}(\psi_i).$$

Il Teorema di Leibniz afferma che $\frac{d^j}{dz^j}(f(z)g(z)) = \sum_{k=0}^{j} \binom{j}{k} f^{(k)}(z) g^{(j-k)}(z)$. Questa formula permette di calcolare la derivata j-esima nell'origine della funzione $\phi(z)$ (3.2):

$$\begin{aligned}
\phi_j(0) \quad &= \left[\sum_{i=0}^{RS-1} p_i \sum_{k=0}^{j} \binom{j}{k} \nu_i(\nu_i - 1) \ldots (\nu_i - k + 1) z^{\nu_i - k} \psi_i^{j-k} e^{\psi_i z} \right]_{z=0} \\
(k \neq \nu_i) &= 0 \\
(k = \nu_i) &= \sum_{i=0}^{RS-1} p_i \binom{j}{\nu_i} \nu_i! \psi_i^{j-\nu_i} = \sum_{i=0}^{RS-1} p_i \frac{j!}{(j-\nu_i)!} psi_i^{j-\nu_i}.
\end{aligned}$$

Essendo $W(z) = \sum_{j=0}^{RS-1} w_j z^j$ risulta

$$\begin{aligned}
W_{\nu_i}(\psi_i) &= \frac{d^{\nu_i}}{d\psi_i^{\nu_i}} \left(\sum_{j=0}^{RS-1} w_j \psi_i^j \right) = \sum_{j=0}^{RS-1} w_j \psi_i^{j-\nu_i} j(j-1) \ldots (j - \nu_i + 1) \\
&= \sum_{j=0}^{RS-1} \frac{j!}{(j-\nu_i)!} w_j \psi_i^{j-\nu_i}.
\end{aligned}$$

Quindi

$$p_t = \sum_{i=o}^{RS-1} p_i W_{\nu_i}(\psi_i) = \sum_{i=0}^{RS-1} p_i \sum_{j=0}^{RS-1} \frac{j!}{(j-\nu_i)!} w_j \psi_i^{j-\nu_i} = \sum_{j=0}^{RS-1} w_j \phi_j(0).$$

Osserviamo che per h sufficientemente grande $RS = (L+1)^{n+1} \leq (h^{2-1/(4n)} + 1)^{n+1} \leq h^{2n+2}$. In particolare la maggiorazione (3.1) vale per ogni $0 \leq j \leq RS$. Dal Lemma B7 i coefficienti di $W(z)$ soddisfano

$$|w_j| \leq \left(\frac{8\sigma}{\rho}\right)^{RS} \leq (8Lc_{14}c_{15}^{L})^{RS} \leq (8h^{2-1/(4n)}c_{14}c_{15}^{h^{2-1/(4n)}})^{h^{2n+2}}.$$

Per $h \to \infty$ il termine dominante è $c_{15}^{h^2}$ e vale la stima $8h^{2-1/(4n)}c_{14} = O(c_{15}^{h^2})$. Segue che per valori di h opportunamente grandi

$$|w_j| \leq (c_{15}^{2h^2})^{h^{2n+4}} = c_{16}^{h^{2n+4}}$$

con $c_{16} = c_{15}^2$. Per costruzione p_t è un intero non nullo, necessariamente $|p_t| \geq 1$ e pertanto

$$\begin{aligned} 0 \leq \log|p_t| = \log\left|\sum_{j=0}^{RS-1} w_j\phi_j(0)\right| &\leq \log\left(\sum_{j=0}^{RS-1} c_{16}^{h^{2n+4}}e^{-h^{8n}}\right) \\ \leq \log(RSc_{16}^{h^{2n+4}}e^{-h^{8n}}) &\leq \log(h^{2n+2}) + h^{2n+4}c_{17} - h^{8n} \end{aligned}$$

dove $c_{17} = \log c_{16}$. Per $h \to \infty$ il termine $-h^{8n}$ domina le altre componenti positive della disuguaglianza. Ciò implica

$$\log|p_t| < 0$$

per h sufficientemente grande portando dunque a una contraddizione. Si conclude che $1, \log\alpha_1, \ldots, \log\alpha_n$ sono linearmente indipendenti su $\overline{\mathbb{Q}}$. □

A conclusione di questa sezione sottolineiamo alcune proprietà delle costanti $c_1, c_2, \ldots$ introdotte nel corso dei capitoli. Come indicato in precedenza tali numeri sono positivi e dipendono esclusivamente dai parametri fissi del problema: i numeri algebrici α_i e β_j, i coefficienti del loro polinomio minimo e le determinazioni scelte per i logaritmi.
Precisiamo un aspetto fondamentale di queste costanti che si è potuto constatare nello sviluppo delle dimostrazioni ovvero che esse sono effettive e permettono di stabilire limiti precisi. Infine, le costanti la cui definizione dipende dai coniugati dei numeri algebrici coinvolti possono essere ulteriormente semplificate e rese esplicite utilizzando la stima della Proposizione 1.16 sui numeri algebrici. Questo le rende direttamente collegate alle proprietà intrinseche dei numeri algebrici considerati.

3.2 Sviluppi successivi

Il Teorema di Baker permette un'ampia applicazione. Di seguito vengono riportati alcuni corollari immediati.

Corollario 3.1. *Qualsiasi combinazione lineare non nulla di logaritmi di numeri algebrici con coefficienti algebrici è trascendente.*

In altri termini per ogni insieme di numeri algebrici non nulli $\alpha_1, \ldots, \alpha_n$ e per ogni insieme di numeri algebrici $\beta_0, \ldots, \beta_n$ non tutti nulli si ha

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n \neq 0$$

se $\beta_0 \neq 0$ oppure se $\beta_0 = 0$ e $\beta_1, \ldots, \beta_n$ sono linearmente indipendenti sui razionali. Ci riferiamo alla condizione $\beta_0 = 0$ come al caso omogeneo e a $\beta_0 \neq 0$ come quello disomogeneo.

Corollario 3.2. *Se $\beta_0, \beta_1, \ldots, \beta_n, \alpha_1, \ldots, \alpha_n$ sono numeri algebrici non nulli, allora $e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ è trascendente.*

Corollario 3.3. *Se $1, \beta_1, \ldots, \beta_n$ sono linearmente indipendenti su $\mathbb{Q}$ con $\beta_1, \ldots, \beta_n$ numeri algebrici e $\alpha_1, \ldots, \alpha_n$ numeri algebrici diversi da $0, 1$, allora $\alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$ è trascendente.*

Corollario 3.4. *$e^{\pi\alpha+\beta}$ è trascendente per ogni α, β algebrici, non entrambi nulli. $\pi + \log \alpha$ è trascendente per ogni α algebrico non nullo.*

I teoremi classici di Hermite e Lindemann sulla trascendenza di e e π e il Teorema di Gel'fond–Schneider sono inclusi nei corollari precedenti come casi speciali.

Una caratteristica significativa di questi risultati risiede nell'effettività del metodo dimostrativo che permette di calcolare un limite inferiore sufficientemente forte per $|\Lambda|$. Baker fornì inizialmente, per $\Lambda \neq 0$, una stima della forma $|\Lambda| \gg e^{-(\log B)^\kappa}$, dove $\kappa > n + 1$ e B rappresenta l'altezza associata ai coefficienti dei β_i. Dalla pubblicazione di questo primo lavoro, la ricerca si è concentrata sul raffinamento delle costanti quantitative e sulla dipendenza esplicita dai parametri fondamentali: il numero di logaritmi, il grado del campo numerico, l'altezza massima dei coefficienti e le altezze dei numeri algebrici coinvolti. Tali ottimizzazioni sono state dettate principalmente dalla necessità di determinare esplicitamente le soluzioni di svariate classi di equazioni diofantee. Nel 1971, Naum Il'ich Fel'dman migliorò il risultato enunciando il seguente teorema.

Teorema 3.5. *Per ogni insieme di numeri algebrici $\alpha_1, \ldots, \alpha_n$ si ha $|\Lambda| \geq B^{-C}$ per tutti i numeri algebrici $\beta_0, \ldots, \beta_n$ con altezze al massimo $B > 1$, dove C è effettivamente computabile in termini di α e del grado dei β.*

Un importante avanzamento metodologico è consistito nell'introduzione dei polinomi di Fel'dman. Questi strumenti permettono di sostituire l'altezza massima B con una quantità più precisa, capace di tenere conto del rapporto tra i coefficienti e i logaritmi. Successivi progressi nelle stime sono stati apportati dallo stesso Baker, da Michel Waldschmidt e, in particolare, da Tarlok Nath Shorey, il quale introdusse il parametro E per gestire i casi in cui i numeri algebrici α_j sono molto vicini all'unità. Tali evoluzioni hanno trasformato il teorema di Baker da un puro risultato teorico di trascendenza in un vero e proprio algoritmo operativo. Il metodo si è dimostrato così capace di risolvere effettivamente intere classi di equazioni diofantee e problemi profondi legati al numero di classi di campi quadratici immaginari.

Nel presente lavoro si è scelto di seguire l'approccio dimostrativo sviluppato originariamente da Baker, il quale nasce come estensione del metodo di Gel'fond. Altre strategie si basano invece sulla generalizzazione del metodo di Schneider o sull'argomento di Daniel Bertrand e David Masser. Questi ultimi hanno dedotto il teorema a partire dal criterio di Schneider-Lang relativo ai valori algebrici di funzioni meromorfe su prodotti cartesiani. Tra i più recenti sviluppi della materia vi è l'introduzione, dovuta a Michel Laurent, dei determinanti di interpolazione, i quali sostituiscono le costruzioni delle funzioni ausiliarie. In questo contesto, anziché risolvere un sistema di equazioni, si considera unicamente il determinante della matrice associata al sistema lineare, eliminando così la necessità di ricorrere al principio dei cassetti di Dirichlet.

Pur godendo di una lunga storia, e nonostante la raffinatezza dei metodi moderni, la teoria dei numeri trascendenti rimane tuttora un campo aperto. Numerosi sono i problemi che attendono ancora una soluzione come le congetture riguardanti l'indipendenza algebrica di e e π o la trascendenza della costante di Eulero γ. L'indipendenza algebrica di e e π sarebbe immediatamente dimostrata se si riuscisse a provare la Congettura di Schanuel, che racchiude in sé quasi tutti i problemi noti sulla trascendenza.

Congettura di Schanuel. *Se $\alpha_1, \ldots, \alpha_n$ sono numeri complessi linearmente indipendenti su $\mathbb{Q}$ allora il grado di trascendenza di $\mathbb{Q}(\alpha_1, \ldots, \alpha_n, e^{\alpha_1}, \ldots, e^{\alpha_n})$ è almeno n.*

Bibliografia

[1] A. Baker. *Transcendental number theory*. Cambridge University Press, 1975.

[2] A. Baker e G. Wüstholz. *Logarithmic forms and diophantine geometry*. Cambridge University Press, 2007.

[3] A. J. Harper. *A version of Baker's theorem on linear forms in logarithms*. Appunti del corso dell'anno 2010.

[4] S. Natarajan e R. Thangadurai. *Pillars of transcendental number theory*. Springer Singapore, 2020.

[5] K. Soundararajan. *Transcendental number theory*. Appunti del corso dell'anno 2011.

[6] M. Waldschmidt. *Diophantine approximation on linear algebraic groups: transcendence properties of the exponential function in several variables*. Springer-Verlag, 2000.