DUE CELEBRI TEOREMI
sui numeri trascendenti

Per “matematici ciclisti”

Teorema di Hermite: la costante *e* è un numero trascendente

Teorema di Lindemann: π è un numero trascendente.

Giacomo Cavallo

INTRODUZIONE

Si può studiare la matematica con diverse motivazioni: per forza, per professione, per interesse.

Io l'ho studiata per forza, non ne ho fatto una professione, ma mi è rimasto un notevole interesse.

Il problema è che sono rimaste nella mia mente diverse domande a cui vorrei che fossero date risposte per me soddisfacenti, anche se non rigorose. Per tutta la mia vita ho cercato spiegazioni semplici di riposti concetti, ma non ho mai trovato una spiegazione che mi soddisfacesse su tre soggetti particolari:

1) *In che consiste la congettura di Riemann.* Nessuno dei testi che ho letto sull'argomento ha mai risposto in modo per me sufficiente chiaro alle mie domande, che poi sono una sola : "Che cosa c'entrano gli zeri della funzione Zeta di Riemann con i numeri primi?". A questo quesito, dopo lunghe considerazioni, ho finalmente dato una risposta per me soddisfacente, ed ho scritto un breve saggio che risponde a tutti i miei interrogativi. Non dimostra la CdR (io non sono da tanto), ma almeno dà pace al mio spirito. Guardando indietro, vedo che da lungo tempo stavo a un passo dal capire questo concetto. Occorreva solo che qualcuno mi dicesse due o tre frasi. Però, quelle due o tre frasi, evidentemente ovvie a chi divulgava la congettura, non mi erano mai state dette, oppure per mia distrazione non le avevo né ascoltate né lette.

2) *Come si dimostra che le equazioni algebriche di grado superiore al quarto non hanno soluzioni generali esprimibili per mezzo di radicali.* Qui abbiamo a che fare col teorema di Ruffini – Abel – Galois (RAG). Di spiegazioni semplici non ne ho mai trovate. Sembrerà strano, ma ho trovato molta maggior difficoltà a capire questa dimostrazione in una qualsiasi delle sue forme, che a capire in che consista la Congettura di Riemann, che pure richiede cognizioni matematiche assai più avanzate. La differenza, però, è che del teorema di RAG io voglio capire la dimostrazione, mentre per la CdR mi bastava capire in che cosa consista l'enunciato del teorema.

3) Come si dimostra la trascendenza dei due numeri π ed e? La trascendenza di π è legata all'impossibilità di quadrare il cerchio, e fu dimostrata da Lindemann nel 1882, dieci anni dopo che Hermite aveva dimostrato la trascendenza del numero e (1872),

senza rendersi conto del fatto che a questo punto aveva già compiuto la maggior parte del cammino per dimostrare anche la trascendenza di π.

In questo saggio, dedicato ai curiosi di matematica dotati di una preparazione di V Liceo Scientifico o I anno di Analisi (non proprio "pedoni", quindi, ma "ciclisti"), mi dedicherò a quest'ultimo punto, riservando le mie forze restanti per il quesito (2).

Per la comprensione di concetti e dimostrazioni matematiche devo confessare che sono convinto che aggiungere elementi storici e biografici sia inutile. La storia della matematica è interessante per se stessa quanto si vuole, ma da una dimostrazione non si dovrebbe capire se il matematico che scrive sia uomo o donna, giovane o vecchio, sano o malato, ricco o povero, sposato o scapolo (a meno che non lo lasci capire lui stesso espressamente). Con opportuna edizione, non si dovrebbe neanche capire in che epoca il matematico sia vissuto. Al massimo si possono citare riferimenti ai predecessori, cioè alla genealogia matematica della dimostrazione, non del matematico, e solo se questa genealogia è utilizzata nella dimostrazione.

NUMERI TRASCENDENTI.

Un numero è trascendente se non è algebrico, cioè se non è soluzione di un'equazione algebrica di grado qualsiasi. Si intende che i coefficienti dell'equazione siano numeri razionali (rapporti di due numeri interi) e quindi, in ultima analisi, numeri interi, come si può subito vedere moltiplicando l'intera equazione per il minimo comun multiplo dei denominatori.

Ad esempio, l'equazione a coefficienti razionali

$$\frac{2}{5}x^2 + \frac{3}{7}\,x + \frac{4}{9} = 0$$

moltiplicata per il minimo comun mutiplo dei denominatori dei coefficienti (315), diventa l'equazione a coefficienti interi:

$$112x^2 + \,135\,x + 140 = 0$$

Si noti che noi non abbiamo formule generali per la soluzione delle equazioni algebriche di grado alto a piacere. Fino al quarto grado si arriva con i radicali, con crescente difficoltà. Ma per risolvere equazioni di grado più elevato occorre chiamare in campo funzioni speciali, la cui conoscenza è riservata a pochi eletti. Ora, una dimostrazione di trascendenza prescinde dall'esistenza di una formula risolutiva. Possiamo immaginare un'equazione di grado mille, la cui soluzione verrà magari espressa per mezzo di funzioni speciali che saranno studiate fra cento anni. Ebbene, la dimostrazione di trascendenza di Hermite garantisce che mai e poi mai una tra le mille soluzioni di cosiffatta equazione sarà il numero *e* o π.

Dopo lunga ricerca, ho finalmente trovato su Internet una laconica dimostrazione dei teoremi di Hermite e di Lindemann. Essa è presentata da Steve Mayer (Novembre 2006) in un breve articolo di 6 pagine dal titolo "The transcendence of π" nel sito http://sixthform.info/maths/files/pitrans.pdf. In questa presentazione, la dimostrazione di Lindemann è preceduta da una dimostrazione in una pagina del teorema di Hermite, che riporto alla fine del mio saggio. Mayer dice di non aver escogitato lui il metodo che ci presenta, ma ha il merito di esporlo con brevità. Per quanto riguarda la chiarezza, l'articolo è certamente chiaro per chi è del mestiere. Io ci ho dovuto studiare non poco, e qui presento il risultato dei chiarimenti che mi sono dato, nella speranza (a) che siano corretti e (b) che possano essere d'aiuto a un eventuale curioso lettore. Suggerirei di leggere per prima cosa la dimostrazione di Mayer. Se vi si trovano dei punti oscuri, allora può essere utile leggere il mio saggio.

Dimostrare la trascendenza di *e* è un puro esercizio di algebra o di analisi matematica, che non dice niente allo spirito, se non è chiaro di che cosa stiamo parlando e perché questa dimostrazione sia importante.

Un'antologia introduttiva dovrebbe consistere di un certo numero di elementi, che possono essere reperiti su vari testi e su Internet stessa. In gran parte si trovano, ad esempio, su uno dei più noti testi di divulgazione matematica, "Che cosa è la matematica?", di Richard Courant e Herbert Robbins (qui citato come C&R), che, in inglese, è con qualche fatica reperibile su Internet (ma sarebbe bene che facesse parte del corredo di chi è interessato alla matematica).

1) Come si passa dai numeri naturali ai numeri razionali. Ad esempio, tracciato un segmento di retta lungo 10, si possono individuare tutti i punti tramite la loro ascissa x. Si può allora vedere dove cadono i numeri naturali (che non includono lo zero) e poi gli interi (che includono lo zero). Si tratta, ovviamente, di punti isolati. Poiché ogni punto ha "lunghezza" zero, la somma delle lunghezze totali dei dieci punti che rappresentano i numeri interi nel nostro intervallo è zero (Vedi C&R, II edizione, Capo II)

2) I numeri razionali (rapporto – *ratio* – di due interi) , pur essendo in numero infinito, costituiscono un insieme "numerabile", cioè possono essere contati, cioè possono essere messi in corrispondenza biunivoca con i numeri naturali. La chiave del metodo è tracciare una tabella $a_{hk} = \frac{h}{k}$, in cui le righe sono i numeratori e le colonne i denominatori (o viceversa) e ragionare su questa (C&R, Capo II, §4).

3) Il fatto che i numeri razionali possano essere contati indica che essi non individuano tutti i punti sul nostro segmento di retta. In effetti, se ricopriamo ogni numero di un insieme numerabile con un trattino o "coperchietto" , la somma dei coperchietti può "essere fatta tendere a zero". Per questo, l'idea è di sfruttare la serie geometrica (o altra serie affine, purché converga): sul primo numero razionale (che è un punto, di lunghezza zero) si mette un trattino di lunghezza *c da decidersi in seguito,* sul secondo un trattino di lunghezza *c/2,* sul terzo un trattino di lunghezza *c/4,* sul successivo *c/8,* poi *c/16* eccetera, ponendo cioè al denominatore le potenze di 2. La somma di tutti questi trattini, cioè la lunghezza della "copertura" di tutti i numeri razionali, quella che si chiama "misura" dell'insieme dei numeri razionali, è:

$$c + \frac{c}{2} + \frac{c}{4} + \frac{c}{8} + \frac{c}{16} \ldots . = 2c$$

Ma *c* è a nostra disposizione. Lo possiamo scegliere più piccolo di qualsiasi numero, il che significa che, a tutti gli effetti , la nostra misura vale zero. Per esempio, su un segmento lungo 10, se vogliamo che la misura valga meno di 10^{-6}, basta scegliere c < (1/2) 10^{-6}. Questo vuol dire che alla fine del procedimento, in un segmento di retta la maggior parte dei numeri non cade sotto uno di questi coperchietti. In altre parole, la maggior parte dei numeri che rappresentano i punti di un segmento di retta non è razionale. L'insieme dei numeri razionali ha "misura nulla" (C&R, Capo II, §4.2: qui viene dimostrato con lo stesso sistema, che il "continuo" non è numerabile)

4) I numeri non razionali sono detti "irrazionali". Un "procedimento diagonale" inventato da Cantor indica come si possano costruire numeri irrazionali (C&R, Capo II, § 4.2)

5) Tra i numeri irrazionali possiamo senz'altro includere i numeri algebrici, cioè i numeri non razionali che sono soluzioni di equazioni algebriche (a coefficienti interi, o razionali, le quali ultime si riducono a coefficienti interi). Ora la sorpresa è che anche i numeri algebrici possono essere contati. La chiave del metodo per contare i numeri algebrici consiste nel contare il numero di equazioni possibili di grado n, ciascuna delle quali ha esattamente n soluzioni, cioè produce n numeri algebrici (reali o complessi). Poiché l'insieme dei numeri algebrici è numerabile, anch'esso, come l'insieme dei numeri razionali, ha misura nulla. Ciò significa che i numeri che individuano i punti su un segmento di retta in massima parte non sono algebrici. Deve esistere una moltitudine di numeri irrazionali non algebrici (C&R, Capo II, §6.1).

6) Questi numeri irrazionali non algebrici sono detti "trascendenti". Un procedimento inventato da Liouville (1844) permette di costruire numeri che sono certamente trascendenti. Il procedimento, però, è complicato, e dà l'impressione che i numeri trascendenti siano delle rarità. Invece, come abbiamo appena visto, in un segmento di retta quasi tutti i punti sono individuati da un'ascissa che è un numero trascendente (C&R, Capo II, §6.1).

Tuttavia, assai più importante è dimostrare che costanti matematiche note sono – o non sono – trascendenti. Questo è stato il compito di generazioni di matematici.

TRASCENDENZA DI *e* (Hermite, 1872).

Per prima cosa spiegherò il metodo generale nelle sue grandi linee; poi lo applicherò alla dimostrazione dell'irrazionalità di *e* (dimostrerò cioè che *e* non può essere espresso come rapporto di due numeri interi). Infine darò la dimostrazione generale.

In appendice A I ho messo copia della dimostrazione data da S. Mayer. Il curioso può anzitutto leggere questa dimostrazione e vedere se gli è chiara. Se così non è, può leggere il resto di questo saggio. Se si sente abbastanza vigoroso può saltare la dimostrazione della irrazionalità di e, che ho introdotto per motivi puramente pedagogici.

I. Il metodo nelle grandi linee.

Dire che *e* è trascendente significa dunque affermare che esso non è soluzione di un'equazione di grado alto a piacere. In altre parole non esiste un'equazione

(1) $$a_0 + a_1\, x + \ a_2\, x^2 + \cdots a_m\, x^m = 0$$

di cui *e* sia una soluzione. Vale a dire, l'equazione

(2) $$a_0 + a_1\, e + \ a_2\, e^2 + \cdots a_m\, e^m = 0$$

è impossibile.

La dimostrazione è per assurdo, cioè mostra che se si assume come vera la (2) si giunge ad un assurdo. La procedura è la seguente:

1) Per mezzo dell'introduzione di un'opportuna funzione ausiliaria in cui entra come parametro un numero primo *p*, che potrà poi essere variato a piacere (e questa è l'astuta chiave del metodo), si costruisce un'equazione che può essere scritta come somma di tre termini,

$$A(p) + \ B(p) + \ C(p) = 0$$

in cui il termine C(p) contiene come fattore il membro di sinistra dell'equazione (2). Facendo l'ipotesi assurda che *e* non sia trascendente, troveremo che C(p) = 0, e quindi

(3) $$A(p) = \ -B\ (p)$$

2) Sarà relativamente facile dimostrare che il primo membro tende a zero per p tendente ad infinito.

3) Sarà un po' più macchinoso dimostrare che B(p) è dato dalla somma di vari interi, positivi e negativi. A questo punto il problema sarà dimostrare che B(p) non può mai essere zero, per quanto sia costituito dalla somma di numeri (interi) positivi e negativi.

Questo lo si otterrà dimostrando che B(p) può essere a sua volta scomposto in due parti: una prima parte, costituita dalla somma di interi positivi e negativi tutti divisibili per p (che quindi possiamo immaginare scritta come *a p+b p-c p+d p-f p* etc), ed una seconda parte, *costituita da un unico intero non divisibile per p*. Si vede allora subito che B(p) non può essere mai zero. Infatti per avere B(p) = 0 dovremmo avere (raccogliendo p a fattor comune dove possibile):

$$(a + b - c + d - f..)p + n = Mp + n = 0$$

in cui *n* non è nullo e non è divisibile per *p*.

Dovrebbe quindi essere

$$M = -n/p$$

il che è impossibile, perché *M*, somma algebrica di interi, deve essere un intero, ma *n/p* non può essere un intero perché, appunto, noi avremo dimostrato che *n* non è divisibile per p. Se poi M fosse nullo, avremmo ancora un assurdo.

4) Avremo dunque a sinistra nella (3) un *A(p)* che tende a zero al crescere di p, ed a destra un intero che non potrà mai essere zero. Di qui l'assurdo, che deriva dall'aver posto eguale a 0 il termine C(p). La (2) sarà dunque impossibile, ed avremo così dimostrato la trascendenza di *e*.

II. Irrazionalità di e.

Darò questa dimostrazione come applicazione del metodo indicato in I. Questo intero paragrafo, introdotto per motivi pedagogici, può essere saltato.

Noi vogliamo dimostrare che *e* è irrazionale, cioè non può essere data come rapporto tra due interi. In altre parole, *e* non può essere la soluzione di un'equazione del tipo

(4) $$a_0 + a_1 e = 0$$

in cui i due coefficienti a_0 ed a_1 siano numeri interi. Qui m =1 nell'equazione (1).

Procedendo per assurdo, cioè assumendo che la (4) sia vera, per prima cosa costruiremo la funzione ausiliaria

(5) $$f(x) = \frac{x^{p-1}(x-1)^p}{(p-1)!}$$

dove p è un numero primo che non divide m, che in questo caso vale 1. (Confesso che di primo acchito non sono riuscito a capire da dove fosse venuta in mente a Hermite questa funzione e la sua sorella maggiore (17), e che nessuno dei testi divulgativi trovati su Internet mi ha aiutato. Si veda tuttavia – a suo tempo - l'Appendice I.)

Per x compreso fra 0 e m, cioè tra 0 e 1, abbiamo che

(6) $$|f(x)| \leq \frac{1}{(p-1)!}$$

La f(x) è un **polinomio** di grado *2p - 1*. La potremo quindi derivare un massimo di *2p – 1* volte.

Viene ora introdotta una proprietà molto generale di una funzione F(x) definita come:

(7) $$F(x) = f(x) + f'(x) + f''(x) + \cdots f^{(2p-1)}(x)$$

Questa funzione ha *2p* termini; la sua derivata avrà un termine in meno, perché la derivata 2p di un polinomio di grado 2p-1 è zero.

Avremo:

(8) $$F'(x) = F(x) - f(x)$$

Vediamo ora come si comporta la derivata di $e^{-x}F(x)$:

(8a) $$\frac{d}{dx}(e^{-x}F(x)) = -e^{-x}F(x) + e^{-x}F'(x) = e^{-x}\big(F'(x) - F(x)\big) = -e^{-x}f(x).$$

Di conseguenza, possiamo scrivere che

(9) $$a_1 \int_0^1 e^{-x}f(x)dx = a_1[-e^{-x}F(x)]_0^1 = a_1F(0) - a_1e^{-1}F(1)$$

Si noti l'importante comparsa del fattore e^{-1} che proviene dall'integrazione. Senza questo fattore - come vedremo più avanti - rischieremmo di aver trovato un metodo che dimostra che *nessun numero* è razionale.

Accanto a quest'equazione avremmo anche la simile:

(10) $$a_0 \int_0^0 e^{-x}f(x)dx = a_0\,F(0) - a_0F(0)$$

la quale vale 0, ma noi per il momento non ce ne curiamo. Si moltiplichi la (9) per e. Sommando membro a membro la (9) moltiplicata per *e* e la (10), possiamo costruire l'equazione :

(11) $a_0 \int_0^0 e^{-x} f(x)dx \; + a_1 e \; \int_0^1 e^{-x} f(x)dx \; = a_0 \, F(0) - \; a_0 F(0) + a_1 \, e \, F(0) - a_1 F(1)$

O anche, riarrangiando i termini a II membro:

(11a)

$$a_0 \int_0^0 e^{-x} f(x)dx \; + a_1 e \; \int_0^1 e^{-x} f(x)dx \; = (\boldsymbol{a_0} \; + \boldsymbol{a_1} \, \boldsymbol{e} \,) F(0) - (\, a_0 F(0) + a_1 F(1))$$

$$= \; -(\, a_0 F(0) + a_1 F(1))$$

avendo accettato l'ipotesi che e sia un numero razionale, e quindi $\boldsymbol{a_0} \; + \boldsymbol{a_1} \, \boldsymbol{e} = \boldsymbol{0}$. E' qui appunto che entra la nostra ipotesi assurda.

E qui vediamo pure che *e*, considerato come il numero di cui vogliamo dimostrare la trascendenza, ha cancellato *e*, base esponenziale, come fattore di F(1). Il nostro metodo non funzionerebbe se volessimo dimostrare la trascendenza (in questo caso l'irrazionalità) di un altro numero.

A sinistra la situazione è gestibile: il primo termine vale 0. Nel secondo noi sostituiamo il massimo del modulo dato nella (6) e troviamo che:

(12) $|a_1 e \; \int_0^1 e^{-x} f(x)dx| \; \leq \; (\, a_1 \, e \int_0^1 e^{-x} dx) \, /(p-1)! \; \leq \; a_1 (e-1)/(p-1)!$

Notiamo che ora è comparso al denominatore un fattore (p-1)! Se noi facciamo crescere p a piacere, il primo membro può quindi essere reso piccolo a piacere.

Che succede al secondo membro? Possiamo scrivere:

(13)

$$II \; membro = \; -a_0 \, [f(0) + \; f'(0) + .. f^{2p-1}(0)] - a_1 \, [f(1) + \; f'(1) + .. f^{2p-1}(1)]$$

La conclusione desiderata può essere ottenuta mediante un ragionamento qualitativo. Non appena si incomincia a derivare il prodotto $x^{p-1}(x-1)^p$ le derivate di x^{p-1} producono fattori di *(p - 1)!* a partire dal più alto, che è p-1. Le derivate di $(x - 1)^p$, invece, incominciano a produrre fattori di *p!*, a partire da *p*. La derivata *n* di questo prodotto è costituita dalla somma di *n + 1* addendi che contengono in tutto *n* derivate distribuite fra il monomio e il binomio.

1) Caso F(0). Facendo le nostre *2p -1* derivate, fino a che $n < p - 1$ ci resterà sempre un fattore x^q , con $q = p - 1 - n$, che annullerà tutti gli *n + 1* termini della derivata per *x = 0*. Quando saremo arrivati alla derivata *n = p -1*, avremo finalmente un addendo che non andrà a zero per

x =0. Si tratterà dell'addendo che risulterà dall'aver derivato *(p - 1)* volte *unicamente* x^{p-1} . Questo termine conterrà quindi il fattore *(p -1)!*

In tutti gli altri termini resteranno a fattore delle potenze di x. Inoltre, sarà comparso qualche termine di p!, partendo da p, a seconda di quante derivate avremo eseguito di $(x-1)^p$.

Avremo dunque il termine

(14) $$(p-1)!\frac{(x-1)^p}{(p-1)!}$$

Il valore della (14) per x = 0 porge $(-1)^p$, *l'unico termine non divisibile per p* di tutta la catena di derivate che compaiono nella F(0).

Continuando a derivare questo addendo nelle derivate successive, di ordine maggiore di *(p-1),* avremo termini non nulli in cui incominceranno ad apparire fattori come *p, p(p-1)* eccetera. D'altronde, eseguendo le successive derivate degli altri addendi avremo termini non nulli solo quando avremo derivato (p-1) volte il fattore x^{p-1}, e cioè avremo introdotto un fattore (p-1)!. Dunque avremo altri termini non nulli, ma tutti questi conterranno almeno un fattore di *p!*, incominciando da *p*. Inoltre il fattore p sarà moltiplicato per un intero, in quanto, una volta eliminato il denominatore (p-1)!, avremo solo a che fare con numeri interi.

2) Caso F(1). Il ragionamento è parallelo al precedente, con la differenza che il binomio è elevato alla potenza p invece che p-1. Incominceremo quindi ad avere termini non nulli solo quando il binomio sia stato derivato p volte, e quindi comparirà un fattore p! che, diviso per (p-1)! darà un fattore p. Una volta derivato il binomio p volte, incominceremo a derivare il monomio, il quale ci darà altri termini di (p-1)!. Anche qui, una volta eliminato il denominatore (p-1)!, avremo solo a che fare con interi.

Lo stesso risultato può essere ottenuto formalmente utilizzando la regola di Leibniz applicata al prodotto $x^{p-1}(x-1)^p$. In generale, la regola afferma che le successive derivate del prodotto di due funzioni sono date da:

(15) $$(u\,v)^{(n)} = \sum_{k=0}^{n}\binom{n}{k} u^{(k)}\, v^{(n-k)}$$

La (15), applicata al polinomio $x^{p-1}(x-1)^p$ diventa:

(16) $$\frac{d^n[(x-1)^p\, x^{p-1}]}{dx^n} = \sum_{k=0}^{n}\binom{n}{k}\frac{p!}{(p-k)!}\frac{(p-1)!}{(p-1-n-k)!}(x-1)^{p-k}x^{p-1-n+k}$$

Calcolando F(0) noi cerchiamo i termini che rispondano a due requisiti:

1) non devono contenere un termine p, il che avviene se $(x-1)^p$ non è mai derivato, cioè k = 0:

2) non devono essere nulli, il che vuol dire che p-1-n+k = 0. Da cui, ponendo k =0, si ottiene che n = p-1.

Se si operano le opportune sostituzioni in (16), si ottiene:

- $\binom{n}{k} = \binom{p-1}{0} = 1$
- $\frac{p!}{(p-k)!} = \frac{p!}{p!} = 1$
- $\frac{(p-1)!}{(p-1-n-k)!} = \frac{(p-1)!}{(0)!}$ = (p-1)!
- $(x-1)^{p-k} = (x-1)^p$ che per x= 0 diviene $(-1)^p$
- $x^{p-1-n+k} = x^0 = 1$

Da cui si vede che il solo superstite è il termine (p-1)!$(-1)^p$, avendo anche posto x=0.

Nel calcolo di F(1) i termini non nulli non possono contenere fattori (x-1), e quindi p=k, il che produce un coefficiente

$$\frac{p!}{0!} = p!$$

che si semplifica col denominatore *(p – 1)!* della funzione f(x), data in (13), lasciando un fattore *p*. Se invece vogliamo i termini che non contengono p, allora deve essere k=0. Le due condizioni sono incompatibili, il che vuol dire che tutti i termini di F(1) o sono nulli o contengono il fattore *p*.

Il termine $- (-1)^p a_0$, unico sopravvissuto, potrebbe essere divisibile per p se a_0 fosse divisibile per p. Ma dobbiamo pensare alla nostra equazione (4) come data con coefficienti fissi, non variabili, per cui *p* può essere scelto come numero primo che non divida a_0, il che del resto è facile, facendo crescere p indefinitamente come richiede il completamento della dimostrazione.

Abbiamo dunque scomposto il nostro secondo membro in due parti, come indicato da principio nella sezione I: un intero $- (-1)^p a_0$ non divisibile per p, ed un addendo, dato dalla somma algebrica di tutte le altre derivate valutate in 0 ed in 1, che sono tutte o nulle o interi che contengono tutti un fattore p, moltiplicate per l'intero a_1.

Facendo tendere p ad infinito, il primo membro tende a zero, mentre il secondo è un intero che non è mai zero. Di qui si conclude che il secondo passaggio nella (11a) non è ammissibile, e quindi la (4) è impossibile, ed *e* non è un numero razionale.

III. Trascendenza di e.

Nella dimostrazione che precede, forse la maggior difficoltà, oltre a spiegarsi da dove Hermite ricavò la sua funzione ausiliaria f(x), è quella di orientarsi tra le derivate della funzione

$$f(x) = \frac{x^{p-1}(x-1)^p}{(p-1)!}$$

per scoprire che solo una di esse, valutata nel punto x=0, non è né nulla né divisibile per p. Questo lo abbiamo fatto utilizzando la formula di Leibniz.

Per dimostrare la trascendenza di e, occorre però dimostrare che *e* non è soluzione di un'equazione di grado *m* qualsiasi. Vale a dire, che l'equazione

(2) $$a_0 + a_1\, e + \ a_2\, e^2 + \cdots a_m\, e^m = 0$$

con coefficienti interi è impossibile. Incidentalmente, a_0 è evidentemente diverso da 0, altrimenti avremmo che *e = 0*; mentre a_m è diverso da zero, perché questo è solo il nome del coefficiente della potenza più alta di e. Del resto abbiamo già dimostrato che *m* non può valere *1*, altrimenti *e* sarebbe un numero razionale.

Fortunatamente, gran parte dei concetti introdotti nella precedente sezione resta valida.

Introduciamo ora la funzione:

(17) $$f(x) = \frac{x^{p-1}(x-1)^p\,(x-2)^p \ldots (x-m)^p}{(p-1)!}$$

La f(x) è ora un polinomio di grado *mp + p -1*, e può quindi essere derivata altrettante volte.

Per x compreso fra 0 e m, abbiamo il seguente limite al valore assoluto:

(18) $$|f(x)| \leq \frac{m^{p-1} m^{mp}}{(p-1)!} = \frac{m^{mp+p+1}}{(p-1)!}$$

Utilizzeremo ancora la funzione:

(19) $$F(x) = f(x) + f'(x) + f''(x) + \cdots + f^{(mp+p-1)}(x)$$

Per cui varranno ancora la

(8) $$F'(x) = F(x) - f(x)$$

e la

(8a) $$\frac{d}{dx}(e^{-x}F(x)) = -e^{-x}F(x) + e^{-x}F'(x) = e^{-x}\big(F'(x) - F(x)\big) = -e^{-x}f(x).$$

mentre la (9) verrà generalizzata per un generico coefficiente a_j nella

(20) $$a_j \int_0^j e^{-x} f(x)dx = a_j[-e^{-x}F(x)]_0^j = a_j F(0) - a_j e^{-j} F(j)$$

Come facemmo per i soli due termini a_0 e a_1, moltiplichiamo per e^j e sommiamo da j =1 a j=m. Avremo:

(21)

$$a_0 \int_0^0 e^{-x} f(x)dx + a_1 e \int_0^1 e^{-x} f(x)dx + \cdots + a_m e^m \int_0^m e^{-x} f(x)dx$$
$$= (\boldsymbol{a_0} + \boldsymbol{a_1}\, \boldsymbol{e} \ldots + \boldsymbol{a_m e^m}) F(0) - (a_0 F(0) + a_1 F(1) \ldots + a_m F(m))$$
$$= -(a_0 F(0) + a_1 F(1) \ldots + a_m F(m))$$

In cui si è ancora una volta utilizzata l'ipotesi che $a_0 + a_1\, e \ldots + a_m e^m = 0$, che si vuole provare assurda.

Il secondo membro può ora essere scritto come:

(22) $$II\ Membro = -[a_0 \left(f(0) + f'(0) \ldots + f^{(mp+p-1)}(0)\right) +$$
$$a_1 \left(f(1) + f'(1) \ldots + f^{(mp+p-1)}(1)\right) +$$
$$a_m \left(f(m) + f'(m) \ldots + f^{(mp+p-1)}(m)\right)]$$

La situazione non è molto cambiata da quella che seguiva la (13).

Considerando la funzione F(0), questa è data dalla somma di diversi termini. Perché un termine non sia nullo per *x = 0* occorre che x^{p-1} sia derivato *(p - 1)* volte, il che produrrà un fattore *(p-1)!*. Perché il termine non sia multiplo di p, occorre che nessuno dei fattori $(x - j)^p$, per $0 < j \leq m$, sia derivato alcuna volta.

Esiste quindi un solo termine non nullo e non multiplo di p, avendo cancellato il fattore *(p -1)!* al numeratore, dato dalla derivazione *p-1* volte di x^{p-1} , con il *(p-1)!* dato dalla definizione di f(x). Il termine varrà:

(23) $$f^{(p-1)}(0) = (-1)^p (-2)^p \ldots (-m)^p$$

il quale non sarà multiplo di *p*, per esempio scegliendo *p > m*.

Per ogni *F(j)*, con $0 < j \leq m$, tutti i termini saranno o nulli o interi multipli di *p*, in quanto un termine non nullo per *x = j* $(0 < j \leq m)$ non dovrà più contenere fattori del

tipo *(x – j)*, il che potrà essere ottenuto solo derivando p volte il termine $(x - j)^p$. Ne risulterà un termine *p!* che si semplificherà con il denominatore *(p-1)!*, producendo *p*.

Inoltre, data la nostra scelta della f(x), p sarà moltiplicato per un numero costituito da vari addendi, tutti contenenti altri termini di p! (se ve ne sono), nonché per un prodotto di differenze di numeri numeri interi. Si tratterà quindi di un numero intero.

Come nel caso precedente, dunque, avremo a II membro una somma contenente solo addendi divisibili per p, più un unico termine dato da

(24) $$-a_0\, f^{(p-1)}(0) \;=\; -a_0(-1)^p(-2)^p \ldots(-m)^p$$

che, per opportuna scelta di p (ad esempio maggiore di m e di a_0) , potrà essere scelto come non divisibile per p, e renderà impossibile l'annullarsi del II membro.

Volendo, avremmo potuto applicare anche qui la regola di Leibniz, che, per il caso generale, porge:

(16') $$\frac{d^n[(x-1)^p\,(x-2)^p \ldots x^{p-1}]}{dx^n} = \sum_{h,k,l}^{n} \frac{n!}{h!k!\ldots l!}\frac{p!}{(p-h)!}\frac{p!}{(p-k)!}\ldots\frac{(p-1)!}{(p-i)!}(x-1)^{p-h}(x-2)^{p-k}\ldots x^{p-1-i}$$

In cui $h+k+\ldots i\ldots = n$.

Caso F(0). Qui non vogliamo alcun fattore p (e quindi h=0, k=0…) e non vogliamo alcun fattore x (e quindi i = p-1).

Ma i = n-h-k-… e quindi abbiamo che n = p-1. Sostituendo i vari termini troviamo che tutti i termini numerici valgono 1, eccetto $\frac{(p-1)!}{(p-i)!}$ = (p-1)!. Da cui otterremo la (24), ponendo x=0.

Caso F(2), valido per qualsiasi F(j) , vista la simmetria della (16').

Vogliamo qui k=p, per avere termini non nulli. Quindi il termine $\frac{p!}{(p-k)!} = p!$, e $n \geq p$,

Questo termine cancellerà il (p-1)! che si trova a denominatore della funzione, lasciando un termine p. Tutti gli altri fattori sono numeri interi.

In quanto al primo membro, in valore assoluto esso sarà inferiore a:

(25) $\frac{m^{mp+p+1}}{(p-1)!}\left[\, a_0 e^0 \int_0^0 e^{-x}dx \;+\; a_1 e^1 \int_0^1 e^{-x}dx + \;\; a_2 e^2 \int_0^2 e^{-x}dx + \cdots + a_m e^m \int_0^m e^{-x}dx \,\right] = \frac{m^{mp+p+1}}{(p-1)!}\left[a_0(e^0-1) + a_1(e-1) + \; a_2(e^2 - \; 1) \ldots + a_m \;\; (e^m - \; 1)\right]$.

Ciò che importa in questa formula è il fattore $\frac{m^{mp+p-1}}{(p-1)!}$, il rimanente della formula essendo costituito da costanti.

In particolare, utilizzando la

$$a_0 + a_1\, e + \;\; a_2\, e^2 + \cdots a_m\, e^m = 0$$

si ricava che il termine fra parentesi quadre è dato da

(26) $$-\left[a_0 + \; a_1 + \; a_2 \ldots + \; a_m\right]$$

cioè dalla somma dei coefficienti dell'equazione, cambiata di segno.

Aumentando p a piacere, il I membro tende a zero, per quanto al numeratore la potenza m^{mp+p-1} cresca indefinitamente. Tuttavia, ad esempio dalla formula di Stirling, sappiamo

(27) $$(p-1)! = p!/(p \;> \sqrt{(2\pi/p)}p^p e^{-p}$$

da cui si ha che *(p-1)!* cresce più rapidamente di qualsiasi potenza *p* di un numero costante, quale è *m*.

Abbiamo quindi ricreato la condizione prevista nella sezione I, con a sinistra un termine che tende a zero per p che tende ad infinito e a destra una somma che tende ad un numero intero non nullo. Per p abbastanza grande i due membri non possono quindi essere eguali, ciò che dipende criticamente dall'aver posto $\boldsymbol{a_0 \;\; + a_1\, e} \ldots + \boldsymbol{a_m e^m = 0}$ nella (21). Resta così dimostrata la trascendenza di e.

APPENDICE A: da Steve Mayer, November 2005.

Theorem 3 (Hermite). *e is transcendental over $\mathbb{Q}$*

Proof. Suppose $a_m e^m + \ldots + a_1 e + a_0 = 0 \qquad (a_i \in \mathbb{Z})$. WLOG $a_0 \neq 0$

Define $$f(x) = \frac{x^{p-1}(x-1)^p(x-2)^p \ldots (x-m)^p}{(p-1)!}$$

where for the moment p is arbitrary and prime.

Define $$F(x) = f(x) + f'(x) + \ldots + f^{(mp+p-1)}(x).$$

Now if $0 < x < m$,

$$\begin{aligned} |f(x)| &\leq \frac{m^{p-1} m^{mp}}{(p-1)!} \\ &= \frac{m^{mp+p-1}}{(p-1)!} \end{aligned}$$

Also $$\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{-x}F(x)\right) = e^{-x}\left[F'(x) - F(x)\right] = -e^{-x}f(x)$$

so that

$$\begin{aligned} a_j \int_0^j e^{-x} f(x)\,\mathrm{d}x &= a_j \left[-e^{-x}F(x)\right]_0^j \\ &= a_j F(0) - a_j e^{-j} F(j). \end{aligned}$$

Multiplying by e^j and summing over $j = 0, 1, \ldots m$ we get

$$\begin{aligned} \sum_{j=0}^{m} a_j e^j \int_0^j e^{-x} f(x)\,\mathrm{d}x &= F(0).0 - \sum_{j=0}^{m} a_j F(j) \\ &= -\sum_{j=0}^{m} \sum_{i=0}^{mp+p-1} a_j f^{(i)}(j). \end{aligned}$$

We claim that each $f^{(i)}(j)$ is an integer, divisible by p except when $j = 0$ and $i = p-1$. For only non-zero terms arise from terms where the factor $(x-j)^p$ has been differentiated p times, and then $p!$ cancels $(p-1)!$ and leaves p, except in the exceptional case.

We show that in the exceptional case the term is NOT divisible by p. Clearly $f^{(p-1)}(0) = (-1)^p \ldots (-m)^p$. We CHOOSE p larger than m, when this product cannot have a prime factor p. Hence the right-hand side of the above equation is an integer $\neq 0$. But as $p \to \infty$ the left-hand side tends to 0, using the above estimate for $|f(x)|$. This is a contradiction. ■

TRASCENDENZA DI π (Lindemann, 1882).

La dimostrazione della trascendenza di π presenta un interesse speciale in quanto risponde (negativamente) al problema della quadratura del cerchio per mezzo di riga e compasso, coronando così duemila anni di ricerche.

Una breve antologia introduttiva dovrebbe contenere due spiegazioni:

1) In che cosa consisteva il problema classico della quadratura del cerchio.

2) Come un punto "costruibile" con riga e compasso nel piano Euclideo sia la soluzione di un'equazione algebrica. Dunque il problema non poteva essere risolto soltanto per mezzo di riga e squadra, a meno che π fosse la soluzione di un'equazione algebrica.

In generale molta chiara informazione, con altrettanto chiare dimostrazioni, è contenuta nel Capo III di C&R. Più in particolare, il punto (2) è nel §2.1 del Capo III. Il problema della quadratura del cerchio è brevemente trattato nel §2.5.

La trascendenza di π, dimostrata da Lindemann del 1882, significa che non esistono equazioni algebriche a coefficienti interi una delle cui soluzioni sia π. Ne segue che "non si può quadrare il cerchio".

I. Il metodo nelle grandi linee.

Dire che π è trascendente significa affermare che esso non è soluzione di un'equazione algebrica di grado alto a piacere. In altre parole non esiste un'equazione algebrica

(1) $$a_0 + a_1\, x + \ a_2\, x^2 + \cdots a_m\, x^m = 0$$

di cui π sia una soluzione. Vale a dire, l'equazione

(2) $$a_0 + a_1\, \pi + \ a_2\, \pi^2 + \cdots a_m\, \pi^m = 0$$

è impossibile.

La dimostrazione è per assurdo, cioè mostra che se si assume come vera la (2) si giunge ad un assurdo. I passi sono i seguenti:

1) Se π soddisfa un'equazione algebrica, anche $i\pi$ soddisfa un'equazione algebrica. Questo ci è utile, perché sappiamo che

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

2) Introducendo le altre n-1 radici α_i dell'equazione a cui soddisfa $i\pi$, possiamo costruire l'equazione

$$(e^{\alpha_1} + 1)(e^{\alpha_2} + 1)(e^{\alpha_3} + 1) \dots (e^{\alpha_n} + 1) = 0$$

Il prodotto sarà nullo, perché a primo fattore avremo posto $e^{i\pi} + 1 = 0$.

3) Svolgendo il prodotto, troviamo che gli esponenti sono dati anzitutto dalle radici dell'equazione a cui soddisfa $i\pi$, poi dalle somme delle radici a coppie, poi dalle somme delle radici a terzetti, e via dicendo, per finire con la somma di tutte le radici. Alcune radici possono essere nulle, e danno ciascuna un addendo di valore e^0=1.

4) Se conosciamo le radici non nulle possiamo costruire l'equazione totale a cui esse soddisfanno.

Essa è della forma:

$$P_1(x)\, P_2(x) \dots P_n(x)=0$$

In cui le radici di $P_1(x)$ sono le radici dell'equazione originaria; le radici di $P_2(x)$ sono tutte le somme delle radici prese a coppie; le radici di $P_3(x)$ sono tutte le somme delle radici prese a terzetti eccetera.

5) La P(x) ci permetterà di costruire in modo naturale una funzione ausiliaria, assai simile alla funzione che Hermite utilizzò per dimostrare la trascendenza della costante *e*. Anche in essa entrerà come parametro un numero primo *p*, che potrà poi essere variato a piacere. Ci ridurremo così a tre termini,

$$A(p) + B(p) + C(p) = 0$$

in cui il termine C(p) contiene come fattore il membro di sinistra di un'equazione direttamente derivata dalla (2). Facendo l'ipotesi assurda che π non sia trascendente, troveremo che C(p) = 0, e quindi

(3) $$A(p) = -B(p)$$

2) Sarà relativamente facile dimostrare che il primo membro tende a zero per p tendente ad infinito.

3) Sarà un po' più macchinoso dimostrare che B(p) è dato dalla somma di vari interi, positivi e negativi. A questo punto il problema sarà dimostrare che B(p) non può mai essere zero, per quanto sia costituito dalla somma di numeri (interi) positivi e negativi.

Questo lo si otterrà dimostrando che B(p) può essere a sua volta scomposto in due parti: una prima parte, costituita dalla somma di interi positivi e negativi tutti divisibili per p (che quindi possiamo immaginare scritta come *a p+b p-c p+d p-f p* etc), ed una seconda parte, *costituita da un unico intero non divisibile per p*. Si vede allora subito che B(p) non può essere mai zero. Infatti per avere B(p) = 0 dovremmo avere (raccogliendo p a fattor comune dove possibile):

$$(a + b - c + d - f..)p + n = Mp + n = 0$$

in cui *n* non è nullo e non è divisibile per *p*.

Dovrebbe quindi essere

$$M = -n/p$$

il che è impossibile, perché *M*, somma algebrica di interi, deve essere un intero, ma *n/p* non può essere un intero perché, appunto, noi avremo dimostrato che *n* non è divisibile per p. Se poi M fosse nullo, avremmo ancora un assurdo.

4) Avremo dunque a sinistra nella (3) un *A(p)* che tende a zero al crescere di p, ed a destra un intero che non potrà mai essere zero. Di qui l'assurdo, che deriva dall'aver posto eguale a 0 il termine C(p). La (2) sarà dunque impossibile, ed avremo così dimostrato la trascendenza di π .

II. Trascendenza di π

Se π fosse un numero algebrico, esso dovrebbe soddisfare ad un'equazione di grado n, della forma

(2) $$a_0 + a_1\,\pi + \ a_2\,\pi^2 + \cdots a_m\,\pi^m = 0$$

Anzitutto notiamo che se π è soluzione di un'equazione algebrica a coefficienti interi, anche iπ soddisfa un'equazione algebrica a coefficienti interi, cioè è un numero algebrico.

Questa affermazione viene di solito data di schianto, ma a me è stato necessario qualche tempo per dimostrarla. Tuttavia, guardando indietro, non è una dimostrazione difficile, e chi la conosce già la può saltare a piè pari.

Per fissare le idee supponiamo che y sia radice di un'equazione di quarto grado:

$$a_4z^4 + a_3z^3 + a_2z^2 + a_1z + a_0 = 0$$

Avremmo allora:

(i) $$a_4y^4 + a_3y^3 + a_2y^2 + a_1y + a_0 = 0$$

Vediamo che cosa succede ponendo iy al posto di y. Ricordando che $i^2 = -1; i^3 = -i; i^4 = i$ avremmo a primo membro

(ii) $$a_4y^4 - i\,a_3y^3 - a_2y^2 + ia_1y + a_0$$

Ora scriviamo l'equazione i cui coefficienti sono a_4 al posto di a_4; ia_3 al posto di a_3; $-a_2$ al posto di a_2; -i a_1 al posto di a_1. Se in questa equazione moltiplicheremo i nuovi coefficienti per le corrispondenti potenze di iy, troveremo:

(iii) $$a_4y^4 + (i\,a_3)(-iy^3) - (-a_2)y^2 + (-ia_1)(iy) + a_0 = a_4y^4 + a_3y^3 + a_2y^2 + a_1y + a_0$$

Ma il secondo membro è la (i) e quindi vale zero. Ne segue che iy è soluzione dell'equazione

(iv) $$a_4z^4 + i\,a_3z^3 - a_2z^2 - ia_1z + a_0 = 0$$

In genere, però, questo non è soddisfacente perché si vogliono coefficienti reali.

Tuttavia, se moltiplichiamo questa equazione per la complessa coniugata, troviamo:

(v) $$((a_4z^4 - a_2z^2 + a_0) + i\,(a_3z^3 - a_1z))\,((a_4z^4 - a_2z^2 + a_0) - i\,(a_3z^3 - a_1z)) = (a_4z^4 - a_2z^2 + a_0)^2 + (a_3z^3 - a_1z)^2$$

Poiché iy è una soluzione del primo fattore a primo membro, il primo membro si annullerà per iy, che quindi sarà soluzione di

(6) $$(a_4z^4 - a_2z^2 + a_0)^2 + (a_3z^3 - a_1z)^2 = 0$$

che ora è un'equazione a coefficienti reali.

Sia $P_1(x)=0$ l'equazione a cui soddisfa iπ.

Supporremo che il grado sia n, e quindi $P_1(x)$ abbia n radici α_i, di cui α_1 sia $i\pi$.

Come è noto, dalla formula di Euler

$$e^{ix} = \cos x + i \sin x$$

si ottiene, per $x = \pi$

$$e^{i\pi} = -1$$

Cioè

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

Si formi ora il prodotto:

(8) $$(e^{\alpha_1} + 1)(e^{\alpha_2} + 1)(e^{\alpha_3} + 1) \dots (e^{\alpha_n} + 1) = 0$$

Il prodotto è nullo, perché il primo fattore non è altri che $e^{i\pi} + 1 = 0$.

Per fissare le idee supporremo di svolgere il prodotto per n=3. Avremmo:

(9) $(e^{\alpha_1} + 1)(e^{\alpha_2} + 1)(e^{\alpha_3} + 1) =$

$$e^{\alpha_1+\alpha_2+\alpha_3} + e^{\alpha_1+\alpha_2} + e^{\alpha_1+\alpha_3} + e^{\alpha_2+\alpha_3} + e^{\alpha_1} + e^{\alpha_2} + e^{\alpha_3} + 1$$

Qui notiamo che

$$\alpha_1 + \alpha_2 + \alpha_3$$

è radice dell'equazione

$$x - (\alpha_1 + \alpha_2 + \alpha_3) = 0$$

che chiameremo $P_3(x)=0$.

Le tre somme di coppie di radici sono soluzioni di

$$\big(x - (\alpha_1 + \alpha_2)\big)(x - (\alpha_1 + \alpha_3)\big(x - (\alpha_2 + \alpha_3)\big) = 0$$

che chiameremo $P_2(x)=0$.

Infine, le tre radici singole sono soluzioni dell'originale $P_1(x)=0$.

In altre parole, gli esponenti della (9) sono le soluzioni dell'equazione

$$P(x)= P_1(x)\, P_2(x)\, P_3(x)=0$$

che è un'equazione di settimo grado.

Notiamo ancora che se $\alpha_2 = -\alpha_3$, il prodotto diventa

$$e^{\alpha_1} + e^{\alpha_1-\alpha_3} + e^{\alpha_1+\alpha_3} + e^{0} + e^{\alpha_1} + e^{-\alpha_2} + e^{\alpha_3} + 1 = 0$$

Cioè:

$$e^{\alpha_1} + e^{\alpha_1-\alpha_3} + e^{\alpha_1+\alpha_3} + e^{\alpha_1} + e^{-\alpha_3} + e^{\alpha_3} + 2 = 0$$

in cui gli esponenti sono le radici di

$$[x-\alpha_1]\big[\big(x-(\alpha_1-\alpha_3)\big)(x-(\alpha_1+\alpha_3)(x)\big][(x-\alpha_1)(x+\alpha_3)(x-\alpha_3)]$$

$$= x\big(x-(\alpha_1-\alpha_3)\big)(x-(\alpha_1+\alpha_3))(x-\alpha_1)^2(x^2-{\alpha_3}^2) = 0$$

Vediamo quindi che dal prodotto possono sorgere radici nulle, che contribuiscono con un addendo 1 al risultato del prodotto.

Generalizzando ad n radici avremo che gli esponenti che risultano svolgendo il prodotto (8) sono le radici di un'equazione

(10) $$P_1(x)\, P_2(x) \ldots P_n(x) = 0$$

le cui radici sono:

- le radici singole di $P_1(x)$,
- tutte le possibili somme di tali radici prese a coppie (radici di $P_2(x)$);
- tutte le possibili somme delle radici prese a terzetti (radici di P3(x));
  e via dicendo.

Come abbiamo notato nel precedente esempio, alcune di queste somme possono annullarsi, creando radici nulle. Le radici nulle, se ce ne saranno, contribuiranno ciascuna un addendo 1 all'ultimo termine del prodotto (8), ed il prodotto (10) si ridurrà ad un polinomio di grado *m*, che chiameremo P(x).

Svolgendo il prodotto, l'equazione potrà essere scritta come

(11) $$P(x) = c_m x^m + c_{m-1} x^{m-1} + \ldots + c_1 x + c_0 = 0$$

La quale, avendo trattato a parte le radici del tipo $x^q = 0$, avrà un coefficiente c_0 diverso da zero. Inoltre l'equazione avrà tutti i coefficienti razionali. *Se i coefficienti non fossero razionali, non sarebbe da noi considerata un'equazione algebrica*.

Non solo, ma moltiplicando tutti i coefficienti per il loro minimo comune multiplo, essi diventano *interi*, e tali li considereremo.

Il fatto che noi *possiamo* scrivere l'equazione algebrica P(x) = 0 con coefficienti *interi* è legato all'ipotesi – che si dimostrerà assurda - che π stesso sia soluzione di P_1 (x), cioè di un'equazione algebrica.

Infatti, per le equazioni algebriche vale il teorema di Viète, uno dei primi teoremi della matematica moderna, poi esteso da Newton, che i coefficienti di un'equazione algebrica, che sono numeri razionali per ipotesi, sono *funzioni simmetriche elementari* (termine definito più oltre) delle radici di un'equazione algebrica. Una funzione simmetrica delle radici, naturalmente, è tale se il suo valore non cambia scambiando fra loro le radici.

Ad esempio, un'equazione di secondo grado come $x^2 - 2 = 0$ ha due soluzioni irrazionali (cioè $r_1 = \sqrt{2}$ e $r_2 = -\sqrt{2}$). Però le due funzioni simmetriche "elementari" delle due radici, cioè in questo semplice caso la loro somma $r_1 + r_2 = 0$ e il loro prodotto $r_1 r_2 = -2$ sono due numeri razionali. Qualsiasi funzione razionale unicamente di queste funzioni sarà ancora simmetrica e rappresentata da un numero razionale.

Il teorema di Viète viene dimostrato senza difficoltà ove si ricordi che un'equazione di grado *n* può essere scritta tanto in termini dei suoi coefficienti (che sono numeri razionali o addirittura interi) quanto in termini delle sue radici, che chiameremo r_1, r_2, ...r_n :

(12) $$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = a_n (x - r_1)(x - r_2) \ldots (x - r_n)$$

Da cui, svolgendo il prodotto a secondo membro e identificando i coefficienti delle potenze di ugual grado, dette s_i *le funzioni simmetriche elementari*:

(13)
$$s_1 = r_1 + r_2 + r_3 + \cdots + r_n$$

$$s_2 = r_1 r_2 + r_1 r_3 + \cdots + r_{n-1} r_n$$

$$s_3 = r_1 r_2 r_3 + r_1 r_2 r_4 + \cdots + r_{n-2} r_{n-1} r_n$$

$$s_n = r_1 r_2 r_3 \ldots r_{n-1} r_n$$

otteniamo le n relazioni:

(14) $$s_1 = -\frac{a_{n-1}}{a_n} \; ; \; s_2 = \frac{a_{n-2}}{a_n} \quad ; \quad s_i = (-1)^i \frac{a_{n-i}}{a_n} \; ; \; s_n = (-1)^n \frac{a_0}{a_n}$$

che sono numeri razionali, se i coefficienti dell'equazione sono numeri razionali o interi. Le s_i, in più, sarebbero interi se fosse $a_n = 1$.

Le funzioni razionali delle funzioni simmetriche elementari sono a loro volta funzioni simmetriche. Ora, l'equazione $P_1(x)=0$ ha come radici le radici α_i dell'equazione algebrica originale, che esiste per ipotesi. Le funzioni simmetriche delle radici α_i possono essere espresse in termini delle loro funzioni simmetriche elementari (i coefficienti dell'equazione divisi per a_n, con segni opportuni) e quindi sono numeri razionali.

L'equazione $P_2(x)= 0$, di grado $h = \binom{n}{2}$, ha come radici tutte le possibili somme $\alpha_i + \alpha_j$, in cui, per contarle una volta sola, supporremo $j > i$. L'equazione può dunque essere svolta come

$$P_2(x) = \prod_{i,j} \left(x - \left(\alpha_i + \alpha_j\right)\right)$$

I coefficienti del polinomio $P_2(x)$, una volta eseguito il prodotto, sono le funzioni simmetriche elementari delle radici di tale equazione. Queste funzioni, oltre ad essere funzioni simmetriche elementari delle somme di coppie di radici α_i, sono anche funzioni simmetriche delle α_i stesse, e quindi sono anch'esse numeri razionali. Che si tratti di funzioni simmetriche delle α_i lo si può vedere perché nel prodotto per $P_2(x)$, permutando in qualsiasi modo le α_i, non otteniamo altro effetto che quello di permutare fra loro le somme di coppie di radici, in quanto esse sono tutte presenti, e sono tutte trattate allo stesso modo. Con ciò, vengono semplicemente permutati fra loro i fattori di $P_2(x)$, mentre i coefficienti del polinomio resteranno invariati. Dunque anche i coefficienti di $P_2(x)$ sono numeri razionali. Lo stesso vale per tutte le $P_i(x)$ e quindi per il loro prodotto.

Oggi che esistono programmi commerciali e anche gratuiti on-line che svolgono operazioni anche complicate, usare un PC per rendersi conto della portata di questa strana formula è quasi obbligatorio. Mi domando quanti studenti possano recitare e derivare le formule indicate sopra senza essersi mai resi conto del loro significato.

Per fissare le idee e per non essere troppo banali, si può per esempio costruire un'equazione di terzo grado usando come elementi un paio di equazioni di cui conosciamo, o possiamo facilmente calcolare, le soluzioni. Noi utilizzeremo:

(1') $x - 2 = 0$ con soluzione 2

(2') $x^2 - x - 1/2 = 0$ con soluzioni $\frac{1}{2}\left(1 - \sqrt{3}\right)$ e $\frac{1}{2}(1 + \sqrt{3})$

Il prodotto delle due equazioni, un'equazione di terzo grado, avrà le tre soluzioni date sopra. Il prodotto, eliminato il denominatore 2, sarà:

(3') $$2x^3 - 6x^2 + 3x + 2 = 0$$

La stessa equazione può essere ricostruita facendo il prodotto

$$(x - 2\,)(\mathrm{x} - \frac{1}{2}\left(1 - \sqrt{3}\right))(x - \frac{1}{2}\left(1 + \sqrt{3}\right))$$

e moltiplicando per il denominatore comune 2. Magicamente troviamo come coefficienti solo numeri interi o razionali. Questa sarà la nostra $P_1(x)$.

Ci sono tre somme di coppie di radici, mediante le quali possiamo formare il polinomio $P_2(x)$:

(4') $$P_2(x) = \left(x - \left(2 + \tfrac{1}{2}(1 - \sqrt{3})\right)\right)\left(x - \left(2 - \tfrac{1}{2}(1 + \sqrt{3})\right)\right)(x - (\tfrac{1}{2}(1 - \sqrt{3}) + \tfrac{1}{2}(1 + \sqrt{3})))$$

$$= x^3 - 6x^2 + \frac{21}{2}x - \frac{11}{2}$$

Infine c'è un'unica somma delle tre radici, da cui si ricava:

(5') $$P_3(x) = \left(x - \left(2 + \tfrac{1}{2}(1 - \sqrt{3}) + \tfrac{1}{2}(1 + \sqrt{3})\right)\right) = x - 3$$

Possiamo ora scrivere il P(x):

(6') $$\mathrm{Px}) = \mathrm{P_1(x)\ P_2(x)\ P_3(x)} = (2x^3 - 6x^2 + 3x + 2)\left(x^3 - 6x^2 + \frac{21}{2}x - \frac{11}{2}\right)(x - 3) =$$
$$= 4x^7 - 48x^6 + 228x^5 - 540x^4 + 645x^3 - 306\,x^2 - 49x + 66 = 0$$

Il problema è che questo polinomio P(x) ha primo coefficiente diverso da 1, per cui le funzioni elementari simmetriche (14) non saranno numeri interi, avendo tutte il primo coefficiente al denominatore.

C'è tuttavia un modo di trasformare il polinomio P(x) in un polinomio che abbia 1 come coefficiente della potenza più alta, restando interi gli altri coefficienti.

Per questo, si consideri che l'equazione

$$(x - a) = 0$$

ha le stesse radici di

$$Ax - Aa = 0$$

Noi possiamo quindi riscrivere la (12) come :

(12') $P(x) = a_n(x - r_1)(x - r_2) \dots (x - r_n) = (a_n x - a_n r_1)(x - r_2) \dots (x - r_n)$

Questo ci indica la strada: moltiplicando per a_n anche gli altri fattori, avremo moltiplicato l'intera equazione per ${a_n}^{n-1}$ ed avremo ottenuto il risultato che tutte le radici contengono un fattore a_n, il che ci assicura che tutte le s_i sono numeri interi, avendo ora introdotto la variabile $X = a_n x$.

(12'') $${a_n}^{n-1} P(x) = (a_n x - a_n r_1)(a_n x - a_n r_2) \dots (a_n x - a_n r_n)$$

La (6') va dunque moltiplicata per ${a_7}^{7-1} = 4^6$, mentre $x = \frac{X}{4}$.

Da cui risulta:

(13') $X^7 - 48\,X^6 + 912\,X^5 - 8640\,X^4 + 41280\,X^3 - 78336X^2 - 50176\,X + 270336 = 0$

il quale è un polinomio di settimo grado con primo coefficiente 1 e a coefficienti interi, anche se quasi proibitivi. Nel nostro caso le sette radici sono ora:

$X_1 = 4$, $X_2 = 8$, $X_3 = 12$, $X_4 = 2(1 - \sqrt{3})$, $X_5 = 2(1 + \sqrt{3})$, $X_6 = 2(5 - \sqrt{3})$, $X_7 = 2(5 + \sqrt{3})$,

che valgono ciascuna il quadruplo delle radici già trovate, come previsto.

La (12'') ci dà dunque un importante risultato. Moltiplicando P(x) per ${a_n}^{n-1}$ noi avremo trasformato P(x) nel prodotto

(12''') $$(a_n x - a_n r_1)(a_n x - a_n r_2) \dots (a_n x - a_n r_n)$$

Le cui radici sono "quasi" numeri interi, in quanto hanno tutte denominatore 1. Sono infatti chiamati "interi algebrici", un soggetto che a buon diritto costituisce un'area della teoria dei campi.

Ma ciò che importa è che le funzioni simmetriche delle radici $a_n r_i$ sono ora numeri interi, come si vede subito, notando che le funzioni simmetriche elementari sono date da uno dei coefficienti diviso a_n.

Torniamo ora al nostro teorema.

Le m radici dell'equazione (11) siano dette $\beta_1, \beta_2 \dots \beta_m$, per cui l'equazione (8) potrà essere scritta:

(13) $$e^{\beta_1} + \cdots + e^{\beta_m} + e^0 + \cdots + e^0 + 1 = 0$$

Dove i termini e^0 (= 1) derivano dalle radici nulle, o, se vogliamo da eventuali termini x^q nell'equazione P(x) =0, i quali, come annunciato, sono così trattati a parte.

Quindi l'equazione può essere scritta come

(14) $$\sum e^{\beta_i} + k = 0$$

In cui k è un intero ≥ 1, dato da *k = 1 + (numero delle radici nulle).*

Introduciamo ora, seguendo il metodo di Hermite, una appropriata funzione ausiliaria:

(15) $$f(x) = \frac{c_m^{mp-1} x^{p-1}([P(x)]^p)}{(p-1)!}$$

Come per il caso del teorema di Hermite, *p* è un parametro che ci sarà utile in seguito. Questa funzione non è molto diversa da quella usata da Hermite, in quanto ora abbiamo:

(15b) $$\frac{c_m^{mp-1} x^{p-1}([P(x)]^p)}{(p-1)!} = \frac{c_m^{mp-1} x^{p-1}(c_m^p[(x-\beta_1)(x-\beta_2)\ldots(x-\beta_m)]^p)}{(p-1)!}$$

In pratica, le β_i hanno sostituito i numeri naturali, e in più è comparso un coefficiente c_m^{mp-1}.

La funzione è un polinomio di grado *mp+p-1*. Il primo coefficiente c_m^{mp-1} è necessario, come abbiamo visto più sopra, per trasformare la funzione $[P(x)]^p = 0$ da un polinomio a coefficienti interi con primo coefficiente arbitrario ad un polinomio in cui il primo coefficiente è 1 e le funzioni simmetriche elementari delle radici sono numeri interi. Abbiamo cioè ricostruito la situazione della (17) con nuovi interi, gli "interi algebrici", che ci garantiscono che le funzioni simmetriche delle radici sono numeri interi.

In effetti, ponendo : $c_m x = X$ e $c_m \beta_i = B_i$

(15c) $$\frac{c_m^{mp-1} x^{p-1}(c_m^p[(x-\beta_1)(x-\beta_2)\ldots(x-\beta_m)]^p)}{(p-1)!} =$$

$$\frac{1}{(p-1)!} \frac{X^{p-1}}{c_m^{p-1}} [(X - B_1)(X - B_2) \ldots (X - B_m)]$$

Le derivate rispetto alla variabile x che ci occorreranno, notando che c_m è una costante, seguono la regola:

(15d) $$\frac{d^n f(X)}{dx^n} = c_m^n \frac{d^n f(X)}{dX^n}$$

Anche in questo caso creeremo la funzione:

(16) $$F(x) = f(x) + f'(x) + f''(x) + \cdots + f^{(mp+p-1)}(x)$$

Ed avremo ancora

(17) $$\frac{d}{dx}(e^{-x}F(x)) = -e^{-x}F(x) + e^{-x}F'(x) = e^{-x}\big(F'(x) - F(x)\big) = -e^{-x}f(x)$$

Da cui

(18) $$-\int_0^x e^{-y}f(y)dy = e^{-x}F(x) - F(0)$$

Moltiplicando per e^x , che può entrare sotto il segno di integrale, l'equazione diventa:

$$-\int_0^x e^{x-y}f(y)dy = F(x) - e^x F(0)$$

Sostituendo y = λx, otteniamo:

(19) $$-x\int_0^1 e^{(1-\lambda)x}f(\lambda x)d\lambda = F(x) - e^x F(0)$$

Sostituiamo ad x le radici β_i e sommiamo. Avremo:

(20) $$-\sum_{j=1}^m \beta_j \int_0^1 e^{(1-\lambda)\beta_j}f\big(\lambda\beta_j\big)d\lambda = \sum_{j=1}^m F\big(\beta_j\big) - (\sum e^{\beta_j})F(0)$$

Che, ricordando la (14), diventa:

(21) $$-\sum_{j=1}^m \beta_j \int_0^1 e^{(1-\lambda)\beta_j}f\big(\lambda\beta_j\big)d\lambda = \sum_{j=1}^m F\big(\beta_j\big) + kF(0)$$

Ricordiamo che alla base di quest'ultimo passaggio stanno tre considerazioni:

- Che si è fatta l'ipotesi che π sia soluzione di un'equazione algebrica (2)
- Che valga la (7): $e^{i\pi} + 1 = 0$, proprietà di cui π gode in modo esclusivo (il che ci assicura che potremo usare questa dimostrazione solo per π).
- Che valga la (14): $\sum e^{\beta_i} + k = 0$

I membro dell'equazione (21)

Grazie alla $f(\lambda\beta_j)$ presente sotto segno di integrale col suo denominatore *(p - 1)!*, il membro di sinistra tende a zero per p tendente ad infinito. Ma, come per il caso del teorema di Hermite, per *p* sufficientemente grande il membro di destra tenderà ad un intero diverso da zero. Dall'impossibilità di eguagliare i due termini sarà dimostrata l'assurdità dell'ipotesi (2) e quindi la trascendenza di π.

Si può vedere un po' meglio come $\beta_j \int_0^1 e^{(1-\lambda)x} f(\lambda x) d\lambda$ tenda a zero per p tendente ad infinito. Abbiamo:

$$|f(\lambda\beta_j)| \leq |f(\beta_j)|$$

In tutte le varie differenze di coppie di radici nonché al posto di x sostituiamo $b = \beta_{j\,max}$. Da cui:

$$|f(\beta_j)| \leq C \frac{b^{mp+p-1}}{(p-1)!}$$

in cui C è la potenza di c_m inclusa nella definizione di f(x).

Il massimo ottenuto per $|f(\lambda\beta_j)|$ può uscire dal segno integrale, lasciandovi $e^{(1-x)b}$, il cui integrale è (e^b-1)/b. Il b al denominatore cancella il b che precederebbe l'integrale, col risultato che la somma è inferiore a

$$m\ C\ e^b \frac{b^{mp+p-1}}{(p-1)!}$$

Dove m è il numero di addendi. Ci siamo così riportati (all'incirca) alla discussione che segue l'equazione (2). Va detto che se si eseguono numericamente gli integrali, per esempio con un opportuno programma, bisogna talvolta salire a valori molto elevati di p, per incominciare a vedere un predominio del denominatore, nel caso in cui b sia grande. Ma prima o poi ci si arriva.

II membro dell'equazione (21)

I termine del II membro

Per dimostrare questa seconda parte incominciamo col considerare che il primo termine è dato da:

$$\sum_{j=1}^{m} F(\beta_j) = \sum_{j=1}^{m} [f(\beta_j) + f'(\beta_j) + f''(\beta_j) + \cdots + f^{(mp+p-1)}(\beta_j)]$$

Che può anche essere scritto:

$$= [f(\beta_1) + f(\beta_2) + \ \ldots + f(\beta_m)] +$$
$$+ [f'(\beta_1) + f'(\beta_2) + \cdots + f'(\beta_m)] +$$
$$+ [\, f''(\beta_1) + f''(\beta_2) + \cdots + \ f''(\beta_m)] + \cdots$$
$$+ [f^{(mp+p-1)}(\beta_1) + f^{(mp+p-1)}(\beta_2) + \cdots + f^{(mp+p-1)}(\beta_m)]$$

Le prime p parentesi quadre sono nulle, perché le β_j sono radici di P(x) e quindi annullano f(x) e tutte le derivate fino a quella di ordine p, in quanto, per $0 \leq t < p$, esiste sempre un termine $(x - \beta_j)$ con esponente non nullo. Restano ora i termini del tipo di:

$$\sum_{j=1}^{m} f^{(t)}(\beta_j)$$

per $p \leq t \leq mp + p - 1$.

Perché una $f^{(t)}(\beta_j)$ non sia nulla, occorre che il termine $(x - \beta_j)^p$ sia stato derivato p volte. Questo ci porta un fattore $p!\, c_m^p$ che si semplifica con il denominatore della f(x), dandoci un fattore p. Le altre derivate opereranno su altri binomi, ma produrranno sempre per primi dei fattori p. Poiché le radici sono presenti nella funzione in forma simmetrica, la somma dei vari termini $f^{(t)}(\beta_j)$, cioè ogni parentesi quadra, costituirà un'unica funzione simmetrica delle radici, che per la scelta da noi fatta del polinomio non sono più β_j, ma $B_j = c_m \beta_j$. La somma sarà quindi un numero intero, moltiplicato per p - che è a fattor comune dei vari termini.

Si veda in appendice un'illustrazione di quanto detto, applicata al caso più elementare che abbiamo presentato sopra.

Sarà dunque:

$$\sum_{j=1}^{m} f^{(t)}(\beta_j) = Mp$$

II termine del II membro

(22) $F(0) = f(0) + f'(0) + f''(0) + \cdots + f^{(mp+p-1)}(0)$

1) Le varie derivate $f^{(t)}(0)$ sono nulle per x = 0 fino a che rimanga un termine x^q, avendo derivato (p-1-q) volte il monomio x^{p-1} di f(x). Quindi sono nulle per $0 \leq t \leq p - 2$.

2) Per t= p-1, le p-1 derivate avranno prodotto a numeratore il fattore $(p-1)!\,c_m^{p-1}$ che si semplificherà con il denominatore. Del polinomio P(x) resterà solo il termine che non moltiplica una potenza di x, cioè c_0, il quale non è nullo, perché si è già tenuto conto delle radici nulle per costituire il termine k nella (14). Poiché il polinomio è elevato alla potenza p, e mettendo insieme quel che resta, avremo:

(23) $$f^{(p-1)}(0) = c_m^{mp-1}c_0^p$$

3) Per $p \leq t \leq mp + p - 1$, incominceremo a derivare il termine P(x). Questo porterà dei fattori *mp*, *mp(mp-1),* e via dicendo, tutti divisibili per p, mentre il denominatore sarà già stato eliminato dalle prime *(p-1)* derivate.

Potremo quindi dire che il secondo termine del secondo membro è:

(24) $$F(0) = Mp + \mathrm{k}c_m^{mp-1}c_0^p$$

costituito da un intero multiplo di p *più un unico addendo* non nullo e non multiplo di p, se avremo l'accortezza di scegliere p primo e maggiore di k, c_0, c_m.

A questo punto avremo, come per il caso della costante *e*, a sinistra un termine che tende a zero per p abbastanza grande, ed a destra un termine che non tenderà a zero per p primo e maggiore di k, c_0, c_m . Di qui l'assurdo, che – tenendo presenti le considerazioni fatte sulla (21) - dimostra il teorema di Lindemann.

APPENDICE B

Illustrazione della relazione:

$$\sum_{j=1}^{m} f^{(t)}\left(\beta_j\right) = 0, \;\; 0 < t < p$$

$$\sum_{j=1}^{m} f^{(t)}\left(\beta_j\right) = Mp, \;\; p \leq t \leq mp + p - 1$$

Anzitutto, con un programma commerciale, eseguiamo le derivate della funzione non banalissima, più semplice possibile, con m=2. Le due radici saranno indicate a e b invece che β_j, per indicare che si tratta di "interi algebrici". Abbiamo:

$$f(x,p) = x^{p-1}\frac{((x-\ a)^p(x-\ b)^p)}{(p-1)!}$$

Primo esempio:

Calcolo di $f^{(t)}(0)$

	p=2
t = p – 1 =1	a^2b^2
t = p = 2	*-4ab(a+b)*
t = 2p =4	$-48(a+b)$
t = 2p+p-1 = 5	5040

Da questa tabella si vede come le varie derivate siano divisibili per p, cioè 2, eccetto la prima, se p è maggiore di a e b.

Prendiamo ora p =10.

	p=10
t = p – 1= 9	a^2b^2
t = p =10	*-100* a^9b^9*(a+b)*
t = 2p =20	Polinomio simmetrico in *a, b* di nono grado, 10 termini, incomincia con

	$-67044257280000a^9 - 3016991577600000a^8b$
t = 2p+p-1 =29	243655257763990904832000000

Dalla tabella risultano la simmetria e la divisibilità per 10, eccetto che in p -1=9.

Calcolo della somma delle due derivate $f^{(t)}(a) + f^{(t)}(b)$:

	p=2
t = p – 1= 1	0
t = p =2	$2(a^3 - a^2b - ab^2 + b^3)$
t = 2p =4	*24(a+b)*
t = 2p+p-1 =5	240

Dalla tabella risultano la simmetria e la divisibilità per 2.

	p=10
t = p – 1= 9	0
t = p =10	Polinomio simmetrico in *a*, *b* di 19° grado, con venti termini, che incomincia con $10a^{19} - 100a^{18}b + 450a^{17}b^2 - 1200a^{16}b^3$
t = 2p =20	Polinomio simmetrico in a,b, dieci termini, nono grado, incomincia con $619274395643904000a^9 - 2930403919322880000a^8b$
t = 2p+p-1 =29	4873105155279818096640000 0

Dalla tabella risultano la simmetria e la divisibilità per 10.

APPENDICE C

Teorema di Lindemann (Steve Mayer, 2006)

Theorem 4 (Lindemann). *π is transcendental over $\mathbb{Q}$*

Proof. If π satisfies an algebraic equation with coefficents in $\mathbb{Q}$, so does $i\pi$ ($i = \sqrt{-1}$). Let this equation be $\theta_1(x) = 0$, with roots $i\pi = \alpha_1, \ldots, \alpha_n$. Now $e^{i\pi} + 1 = 0$ so

$$(e^{\alpha_1} + 1) \ldots (e^{\alpha_n} + 1) = 0$$

We now construct an algebraic equation with integer coefficients whose roots are the exponents of e in the expansion of the above product. For example, the exponents in pairs are $\alpha_1 + \alpha_2$, $\alpha_1 + \alpha_3$, $\ldots$, $\alpha_{n-1} + \alpha_n$. The αs satisfy a polynomial equation over $\mathbb{Q}$ so their elementary symmetric functions are rational. Hence the elementary symmetric functions of the sums of pairs are symmetric functions of the αs and are also rational. Thus the pairs are roots of the equation $\theta_2(x) = 0$ with rational coefficients. Similarly sums of 3 αs are roots of $\theta_3(x) = 0$, etc. Then the equation

$$\theta_1(x)\theta_2(x) \ldots \theta_n(x) = 0$$

is a polynomial equation over $\mathbb{Q}$ whose roots are all sums of αs. Deleting zero roots from this, if any, we get

$$\theta(x) = 0$$
$$\theta(x) = cx^r + c_1x^{r-1} + \ldots c_r$$

and $c_r \neq 0$ since we have deleted zero roots. The roots of this equation are the non-zero exponents of e in the product when expanded. Call these $\beta_1, \ldots \beta_r$. The original equation becomes

$$e^{\beta_1} + \ldots e^{\beta_r} + e^0 + \ldots e^0 = 0$$

ie

$$\sum e^{\beta_i} + k = 0$$

where k is an integer > 0 $\quad (\neq 0$ since the term $1 \ldots 1$ exists)
Now define

$$f(x) = c^s x^{p-1} \frac{[\theta(x)]^p}{(p-1)!}$$

where $s = rp - 1$ and p will be determined later.
Define

$$F(x) = f(x) + f'(x) + \ldots + f^{(s+p)}(x).$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[e^{-x}F(x)\right] = -e^{-x}f(x) \text{ as before.}$$

Hence we have

$$e^{-x}F(x) - F(0) = -\int_0^x e^{-y}f(y)\,\mathrm{d}y$$

Putting $y = \lambda x$ we get

$$F(x) - e^xF(0) = -x\int_0^1 e^{(1-\lambda)x}f(\lambda x)\,\mathrm{d}\lambda.$$

Let x range over the β_i and sum. Since $\sum e^{\beta_i} + k = 0$ we get

$$\sum_{j=1}^{r} F(\beta_j) + kF(0) = -\sum_{j=1}^{r} \beta_j \int_0^1 e^{(1-\lambda)\beta_j} f(\lambda\beta_j)\,\mathrm{d}\lambda.$$

CLAIM. *For large enough p the LHS is a non-zero integer.*

$\sum_{j=1}^{r} f^{(t)}(\beta_j) = 0 \quad (0 < t < p)$ by definition of f. Each derivative of order p or more has a factor p and a factor c^s, since we must differentiate $[\theta(x)]^p$ enough times to get $\neq 0$. And $f^{(t)}(\beta_j)$ is a polynomial in β_j of degree at most s. The sum is

symmetric, and so is an integer provided each coefficient is divisible by c^s, which it is. (symmetric functions are polynomials in coefficients = polynomials in $\frac{c_i}{c}$ of degree $\leq s$). Thus we have

$$\sum_{j=1}^{r} f^{(t)}(\beta_j) = pk_t \quad t = p, \ldots, p+s.$$

Thus $LHS = (\text{integer}) + kF(0)$. What is $F(0)$?

$$\begin{aligned} f^{(t)}(0) &= 0 \qquad t = 0, \ldots, p-2. \\ f^{(p-1)}(0) &= c^s c_r^p \qquad (c_r \neq 0) \\ f^{(t)}(0) &= p(\text{some integer}) \qquad t = p, p+1, \ldots. \end{aligned}$$

So the LHS is an integer multiple of $p + c^s c_r^p k$. This is not divisible by p if $p > k, c, c_r$. So it is a non-zero integer. But the RHS $\to 0$ as $p \to \infty$ and we get the usual contradiction. ■