3.2 Polinomi

Molti problemi di algebra riguardano i polinomi in una o più variabili, quindi prendiamoci un attimo di pausa per richiamare alcune definizioni e risultati su questi polinomi.
Un *polinomio in una variabile* è una funzione, chiamiamola $f(x)$, della forma

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0$$

o, per essere più formali,

$$f(x) = \sum_{i=0}^{n} a_i x^i.$$

Gli a_i sono costanti (in questo libro saranno sempre numeri reali) e assumiamo che a_n sia diverso da 0. Chiamiamo n il *grado* di f.
I polinomi in più variabili, diciamo tre variabili per esempio, non hanno la stessa bella forma dei polinomi monodimensionali, ma sono comunque abbastanza utili. In ogni caso, $f(x, y, z)$ è un polinomio in tre variabili se ha la forma

$$f(x, y, z) = \sum_{k,l,m} a_{k,l,m} x^k y^l z^m,$$

dove $a_{k,l,m}$ sono costanti (reali), la somma varia tra i k, l e m non negativi tali che $k + l + m \leq n$ e assumiamo che almeno uno degli $a_{k,l,m}$ diversi da 0 soddisfi $k + l + m = n$. Anche in questo caso chiamiamo n il *grado* di f; i polinomi di grado 2 sono detti *quadratici*, di grado 3 *cubici* e così via... Se il grado è 0, allora il polinomio è detto *banale* o *costante*. Se tutti gli $a_{k,l,m}$ non nulli soddisfano $n = k + l + m$, allora f è detto *omogeneo*. I polinomi omogenei hanno la proprietà che

$$f(tx_1, tx_2, \ldots, tx_m) = t^m f(x_1, x_2, \ldots, x_m)$$

per tutti gli $x_1, \ldots, x_m, t$. Per esempio, $x^2y + z^3 + xz$ è un polinomio di tre variabili (x, y e z) e ha grado 3. Non è omogeneo, perché il termine xz ha grado 2.
Un polinomio f con m variabili si dice *fattorizzato* nei due polinomi p e q se $f(x_1, \ldots, x_m) = p(x_1, \ldots, x_m)q(x_1, \ldots, x_m)$ per tutti gli $x_1, \ldots, x_m$; p e q sono detti *fattori* di f. Si dimostra facilmente che il grado di un polinomio è uguale alla somma dei gradi dei fattori. Un polinomio è *irriducibile* se non può essere fattorizzato in fattori non banali.
Le *radici* di un polinomio $f(x_1, \ldots, x_m)$ sono i valori di $(x_1, \ldots, x_m)$ che danno un valore zero, cosicché $f(x_1, \ldots, x_m) = 0$. I polinomi di una variabile possono avere tante radici quante il loro grado; in effetti se si tiene conto delle molteplicità e delle radici complesse, i polinomi di una variabile hanno sempre esattamente

tante radici quante il loro grado. Per esempio, le radici di un polinomio quadratico $f(x) = ax^2 + bx + c$ sono date dalla ben nota formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

I polinomi cubici e quelli di quarto grado hanno anch'essi delle formule per il calcolo delle loro radici, ma sono molto più intricate e non molto utili nella pratica. Non appena si arriva ai polinomi di quinto grado o maggiore, non ci sono per niente formule elementari! E i polinomi di due o più variabili sono ancora peggio; tipicamente ci sono un numero infinito di radici.
Le radici di un fattore sono un sottoinsieme delle radici del polinomio originale; questa può essere un'informazione utile per decidere se un polinomio divide un altro oppure no. In particolare, $x - a$ divide $f(x)$ se e solo se $f(a) = 0$, dato che a è radice di $x - a$. In particolare, per ogni polinomio $f(x)$ di una variabile e ogni numero naturale t, $x - t$ divide $f(x) - f(t)$.
Affrontiamo dunque qualche problema sui polinomi.

Problema 3.3. (Australian Mathematics Competition 1987, pag. 13). Siano a, b, c numeri reali tali che

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = \frac{1}{a+b+c} \tag{3.4}$$

con tutti i denominatori diversi da 0. Dimostrare che

$$\frac{1}{a^5} + \frac{1}{b^5} + \frac{1}{c^5} = \frac{1}{(a+b+c)^5}\,. \tag{3.5}$$

A prima vista questo problema sembra semplice. È data esattamente un'unica informazione, quindi dovrebbe essere una sequenza lineare di passi logici che ci porta al risultato voluto. Beh, un tentativo iniziale di dedurre la seconda equazione dalla prima potrebbe consistere nell'elevare entrambi i membri di (3.4) alla quinta potenza, il che ci porta a qualcosa di simile al risultato desiderato, ma con un'accozzaglia di termini confusi a sinistra. Non sembra che ci siano altre manipolazioni ovvie. Questo è quanto, riguardo all'approccio diretto.
A un secondo sguardo, la prima equazione sembra sospetta, come una di quelle equazioni che gli studenti delle scuole superiori sono avvertiti di non usare perché di solito sono ingannevoli. Questo ci dà il nostro primo vero indizio: la prima equazione dovrebbe limitare un bel po' a, b e c. Potrebbe valere la pena di reinterpretare l'equazione (3.4).

Un denominatore comune sembra essere un buon punto di partenza. Combinando i tre reciproci nel membro di sinistra otteniamo

$$\frac{ab+bc+ca}{abc} = \frac{1}{a+b+c}$$

e, moltiplicando a croce, otteniamo

$$ab^2 + a^2b + a^2c + ac^2 + b^2c + bc^2 + 3abc = abc. \tag{3.6}$$

A questo punto potremmo pensare alle varie disuguaglianze da usare qui: Cauchy-Schwarz, media aritmetica-media geometrica ecc. (Hardy 1975, pagg. 33-34). Non sarebbe male se a, b e c fossero vincolati a essere positivi, ma non c'è una simile restrizione: in effetti la condizione non può valere se a, b e c sono positivi, dato che $\frac{1}{a+b+c}$ sarebbe più piccolo di tutti e tre i reciproci del membro di sinistra di (3.4). Dato che (3.6) è equivalente a (3.4) ed è algebricamente più semplice, poiché (3.6) non contiene reciproci, potremmo cercare di dedurre (3.5) da (3.6). Ancora una volta l'approccio diretto non è possibile. Di solito l'unico altro modo per dedurre un'equazione da alcune altre è dimostrare un risultato intermedio o fare qualche sostituzione utile.

Ci sono alternative più esotiche, quali considerare (3.4) come un profilo della funzione $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} - \frac{1}{a+b+c}$ e quindi usare l'analisi per trovare la forma e le proprietà di tale profilo, ma è meglio restare prima sulle opzioni semplici.

Sostituire non sembra una strada opportuna: le equazioni (3.4) e (3.6) sono abbastanza semplici così come sono e le sostituzioni non le renderebbero molto più semplici. Quindi proveremo a indovinare e a dimostrare un risultato intermedio. Il miglior tipo di risultato intermedio è una parametrizzazione, dato che può essere direttamente sostituita nell'equazione desiderata.

Un modo di parametrizzare è risolvere in una delle variabili, per esempio a. L'equazione (3.6) non si risolve facilmente in a, a meno di non voler usare la formula per le equazioni di secondo grado. L'equazione (3.4) si risolve per a e possiamo dimostrare il problema risolvendo per a, b e c a turno e deducendo un risultato intermedio, che per inciso è equivalente al risultato che mostrerò qui sotto. Ma doveva accadere, no? In ogni caso, ora, proverò una strada differente.

Se fallisce una parametrizzazione, si potrebbe semplicemente riformulare (3.6) in una forma migliore. Le soluzioni di (3.6) sono sostanzialmente le radici del polinomio $a^2b + b^2a + b^2c + c^2b + c^2a + a^2c + 2abc$. Il migliore modo per trattare con le radici dei polinomi consiste nel fattorizzarli, e viceversa. Quali sono i fattori? Dato che sappiamo che (3.6) deve implicare in qualche modo (3.5), dovremmo essere abbastanza fiduciosi nel fatto che ci sia una qualche forma maneggevole di (3.6) che ci condurrà a (3.5) e l'unica forma maneggevole di un polinomio è una

scomposizione in fattori. Per trovare quali siano, dobbiamo fare esperimenti. Il polinomio è omogeneo, quindi dovrebbero esserlo anche i suoi fattori.
Il polinomio è simmetrico, perciò i fattori dovrebbero essere reciprocamente simmetrici. Il polinomio è cubico, pertanto dovrebbe esserci un fattore lineare. Dovremmo ora provare fattori della forma $a+b$, $a-b$, $a+b+c$, $a+b-c$ e così via. Anche espressioni come $a+2b$ potrebbero funzionare, ma non sono così "carine" e in ogni caso le si può provare in seguito. Diventa presto evidente, dal Teorema di fattorizzazione, che $a+b$ e, similmente, $b+c$ e $c+a$ sono radici della cubica. È facile verificare che (3.6) è fattorizzabile in $(a+b)(b+c)(c+a)$. Questo significa che la (3.4) è vera se e solo se $a+b=0$, $b+c=0$ o $c+a=0$. Il trucco consiste nel sostituire ognuna di queste possibilità in (3.5). □

Esercizio 3.5. Fattorizzare $a^3+b^3+c^3-3abc$.

Esercizio 3.6. Trovare tutti gli interi a,b,c,d tali che $a+b+c+d=0$ e $a^3+b^3+c^3+d^3=24$. Suggerimento: non è difficile indovinare *alcune* soluzioni di queste equazioni, ma per dimostrare che le abbiamo trovate *tutte*, sostituiamo la prima equazione nella seconda e fattorizziamo.

La fattorizzazione di polinomi, o la sua impossibilità, è una parte affascinante della matematica. Il seguente problema è istruttivo, perché utilizza quasi tutti i trucchi presenti nel libro per trovare una soluzione.

Problema 3.4. (**) Dimostrare che nessun polinomio della forma $f(x)=(x-a_1)^2(x-a_2)^2\cdots(x-a_n)^2+1$, dove $a_1,\ldots,a_n$ sono tutti interi, può essere fattorizzato in due polinomi non banali, ognuno a coefficienti interi.

Questo è un enunciato piuttosto generale: ci dice per esempio che il polinomio

$$(x-1)^2(x+2)^2+1=x^4+2x^3-3x^2-4x+5$$

non può essere fattorizzato in altri polinomi a coefficienti interi. Come possiamo dimostrarlo?
Beh, supponiamo che $f(x)$ sia fattorizzabile in due polinomi non banali a coefficienti interi, $p(x)$ e $q(x)$. Allora $f(x)=p(x)q(x)$ per tutti gli x. Bell'affare. Ma ricordiamo che f ha questa proprietà speciale: è una specie di quadrato più uno. Come possiamo usare questo fatto? Beh, potremmo dire che $f(x)$ è sempre positivo (o addirittura che $f(x)\geq 1$), ma questo non ci dice molto su $p(x)$ e $q(x)$ tranne che hanno lo stesso segno. Tuttavia abbiamo un altro dato: f non è solo un buon vecchio quadrato qualsiasi aumentato di uno, ma è il quadrato di una combinazione di fattori lineari. Possiamo usare questi $(x-a_i)$ a nostro vantaggio?

Beh, il migliore fattore che potremmo avere è 0, perché questo renderebbe l'intera espressione 0. In realtà, ci sono anche casi in cui avere 0 è l'ultima cosa che uno vorrebbe, perché si potrebbe desiderare di semplificare quel fattore. $(x - a_i)$ è 0 quando x è a_i, quindi abbiamo un'idea: mettiamo un a_i al posto di x. Otteniamo

$$f(a_i) = \cdots (a_i - a_i)^2 \cdots + 1 = 1.$$

Tornando a $p(x)$ e $q(x)$, questo risultato significa che

$$p(a_i)q(a_i) = 1.$$

E questo cosa ci dice? Assai poco, a meno che non si ricordi che p e q hanno coefficienti interi e che gli a_i sono anch'essi interi. Il risultato è che $p(a_i)$ e $q(a_i)$ sono ambedue interi. Abbiamo dunque due interi il cui prodotto è 1. Questo può accadere solo quando gli interi sono entrambi 1 o entrambi -1. In breve,

$$p(a_i) = q(a_i) = \pm 1 \text{ per ogni } i = 1, \ldots, n.$$

Si dovrebbe stare un po' attenti con la notazione $\pm$ qui; sappiamo che $p(a_1)$ e $q(a_1)$, per esempio, sono uguali uno all'altro, ma $p(a_1)$ e $p(a_2)$ potrebbero avere lo stesso segno o segno opposto, per quello che sappiamo finora.
Abbiamo trovato, più o meno, i valori di $p(a_1), \ldots, p(a_n)$ e $q(a_1), \ldots, q(a_n)$, quindi ogni polinomio è "appeso" a n punti. Ma i polinomi di coefficiente direttore 1 hanno tanti gradi di libertà quanto il loro grado. Ora, dato che $pq = f$, il grado di p più il grado di q è uguale al grado di f, che è $2n$. Questo significa che uno dei polinomi, diciamo p, ha grado al più n. In sintesi, abbiamo un polinomio di grado n ma costretto a passare per n punti dati. Possiamo sperare che questo conduca a una contraddizione, che è quello che cerchiamo.
Cosa sappiamo di un polinomio che ha grado al più n? Beh, che ha al massimo n radici. Sappiamo qualcosa delle radici di p? Beh, p è un fattore di f, quindi le radici di p sono anche radici di f. Quali sono le radici di f? Non ce ne sono! (Beh, nessuna sulla retta reale, almeno.) f è sempre positivo (in effetti, è sempre almeno 1) e quindi non può avere radici. Questo significa che a sua volta p non può avere radici. Cosa significa quando un polinomio non ha radici? Significa che non attraversa mai lo 0, cioè che non cambia mai segno. In altre parole, p è sempre positivo o sempre negativo. Questo ci dà due casi, ma possiamo risparmiare un po' di lavoro osservando che un caso implica l'altro. Infatti se abbiamo una fattorizzazione $f(x) = p(x)q(x)$, abbiamo automaticamente anche un'altra fattorizzazione $f(x) = (-p(x))(-q(x))$. Quindi se p è sempre negativo, possiamo sempre capovolgere la fattorizzazione e ottenere una nuova fattorizzazione in cui p è sempre positivo.

Quindi, senza perdita di generalità, supporremo che p sia sempre positivo. Sappiamo già che $p(a_i)$ è $+1$ o -1, e ora sappiamo che è positivo, quindi $p(a_i)$ deve essere $+1$ per ogni i. E $q(a_i)$ deve per forza essere uguale a $p(a_i)$, quindi anche $q(a_i)$ è $+1$ per ogni i. E ora?

Beh, $p(x)$ e $q(x)$ devono prendere il valore $+1$ almeno n volte. Questo può essere riformulato in termini di radici come segue: $p(x) - 1$ e $q(x) - 1$ hanno almeno n radici. Ma $p(x) - 1$ ha grado al più n, perché p stesso ha grado al più n. Questo significa che l'unico modo in cui $p(x) - 1$ può avere n radici è che abbia esattamente grado n. Questo significa a sua volta che $p(x)$ ha grado n e quindi $q(x)$ ha anch'esso grado n.

Diamoci l'obiettivo di assemblare quello che abbiamo capito fino a ora: abbiamo assunto $f(x) = p(x)q(x)$; p e q sono entrambi polinomi positivi a coefficienti interi di grado n e $p(a_i) = q(a_i) = 1$ o, in alternativa, $p(a_i) - 1 = q(a_i) - 1 = 0$, per ogni i. Ora conosciamo le radici di $p(x) - 1$: sono gli a_i. E sono le sole radici di $p(x) - 1$, dato che sappiamo che $p(x) - 1$ può avere al massimo n radici. Questo significa che $p(x) - 1$ è nella forma

$$p(x) - 1 = r(x - a_1)(x - a_2) \cdots (x - a_n)$$

e lo stesso vale per $q(x) - 1$

$$q(x) - 1 = s(x - a_1)(x - a_2) \cdots (x - a_n)\,,$$

dove r e s sono costanti. Per sapere di più su r e s, ricordiamo che p e q sono polinomi interi. Il coefficiente direttore di $p(x) - 1$ è r e il coefficiente direttore di $q(x) - 1$ è s. Questo significa che r e s devono essere interi.

Ora applichiamo queste formule per $p(x)$ e $q(x)$ alla nostra formula originaria $f(x) = p(x)q(x)$ e otteniamo

$$(x - a_1)^2(x - a_2)^2 \cdots (x - a_n)^2 + 1 =$$

$$(r(x - a_1)(x - a_2) \cdots (x - a_n) + 1)(s(x - a_1)(x - a_2) \cdots (x - a_n) + 1).$$

Questa equazione mette in relazione due polinomi definiti esplicitamente. La cosa migliore da fare ora è confrontare i coefficienti.

Confrontando i coefficienti di x^n otteniamo $1 = rs$ e, dato che r e s sono interi, questo significa che $r = s = +1$ o $r = s = -1$. Supponiamo per prima cosa che sia $r = s = 1$. La nostra equazione polinomiale allora diventa

$$(x - a_1)^2(x - a_2)^2 \cdots (x - a_n)^2 + 1 =$$

$$((x - a_1)(x - a_2) \cdots (x - a_n) + 1)((x - a_1)(x - a_2) \cdots (x - a_n) + 1).$$

Espandendo e semplificando, otteniamo

$$2(x-a_1)(x-a_2)\cdots(x-a_n)=0$$

che è ridicolo, perché dovrebbe valere per tutti gli x. Il caso $r=s=-1$ è analogo e quindi abbiamo concluso. □

Esercizio 3.7. Dimostrare che il polinomio $f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)-1$ non può essere fattorizzato in due polinomi più piccoli a coefficienti interi, dove gli a_i sono interi distinti. Suggerimento: se $f(x)$ si fattorizza nel prodotto di due polinomi $p(x)$ e $q(x)$, concentriamoci su $p(x)+q(x)$. Notiamo che questa strategia particolare potrebbe essere applicata anche al Problema 3.4, ma risulta essere in qualche modo non efficace in quel caso.

Esercizio 3.8. Sia $f(x)$ un polinomio con coefficienti interi e siano a e b interi. Si dimostri che $f(a)-f(b)$ può essere uguale a 1 solo se a e b sono consecutivi. Suggerimento: fattorizzare $f(a)-f(b)$.