Appunti per il corso di Metodi Algebrici

Enrico Gregorio

Anno accademico 2002–2003

gregorio@sci.univr.it

Dipartimento di Informatica—Settore di Matematica
Università di Verona

CAPITOLO 1

Serie di potenze

1.1. Polinomi

Se A è un anello commutativo, come sempre in questi appunti, abbiamo definito l'*anello dei polinomi* $A[x]$ come l'insieme delle successioni

$$(a_n) = (a_0, a_1, \dots, a_n, \dots)$$

in A che sono *definitivamente nulle*, cioè per le quali esiste un $\bar{n}$ tale che

$$a_n = 0, \text{ per ogni } n > \bar{n}.$$

Su questo insieme abbiamo definito le operazioni di addizione e moltiplicazione con

$$(a_n) + (b_n) = (s_n)$$
$$(a_n)(b_n) = (p_n)$$

dove $s_n = a_n + b_n$ e

$$p_n = \sum_{i=0}^{n} a_i b_{n-i} = \sum_{i+j=n} a_i b_j.$$

Non è difficile dimostrare che $A[x]$ è un anello, in cui l'elemento neutro per l'addizione è la successione costante 0 e quello per la moltiplicazione è $(1, 0, \dots, 0, \dots)$.

Si è poi notato che l'applicazione $j\colon A \to A[x]$ definita da

$$j(a) = (a, 0, \dots, 0, \dots)$$

è un omomorfismo iniettivo di anelli e che è possibile *identificare*, tramite j, A con un sottoanello di $A[x]$. Se poi si pone

$$x = (0, 1, 0, \dots, 0, \dots),$$

si ha che

$$x^n = (0, \dots, 0, \overset{n}{\underset{\downarrow}{1}}, 0, \dots, 0, \dots)$$

e ci si accorge che ogni polinomio $f \in A[x]$ si può scrivere in modo unico come

$$f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

(intendendo che non è rilevante per l'unicità se aggiungiamo o togliamo termini con coefficienti nulli). Se il polinomio non è *nullo*, possiamo supporre comunque che $a_n \neq 0$; in tal caso n è il *grado* di f, che si indica anche con $\partial(f)$. Il grado del polinomio nullo è, per convenzione, $-\infty$.

Scrivere i polinomi in questo modo permette di eseguire i calcoli in modo usuale, senza ricorrere alla formula complicata per i prodotti: basta osservare che $A[x]$ è un anello commutativo e quindi *ridurre i termini simili*.

Quando l'anello A è un campo, sappiamo che $A[x]$ è un *dominio euclideo* e quindi ogni suo ideale è principale. Se I è un ideale di $A[x]$ e $I \neq \{0\}$, basta prendere in I un elemento di grado minimo (e possiamo scegliere l'*unico* polinomio monico con tale grado) f e si ha

$$I = fA[x] = \{\, fg \mid g \in A[x] \,\}.$$

Dal momento che $\partial(fg) = \partial(f) + \partial(g)$, quando A è un dominio, sappiamo che gli elementi invertibili di $A[x]$ sono esattamente gli elementi invertibili di A.

Tuttavia, riflettendo sulle dimostrazioni che permettono di concludere che $A[x]$ è un anello, non è difficile convincersi che anche l'insieme di *tutte* le successioni in A è, con le operazioni definite allo stesso modo, un anello, che denoteremo con $A[\![x]\!]$.

Proviamo allora a calcolare il prodotto delle successioni (a_n) e (b_n), dove

$$a_n = \begin{cases} 1 & \text{se } n = 0 \\ -1 & \text{se } n = 1 \\ 0 & \text{se } n > 1 \end{cases} \qquad \text{e} \qquad b_n = 1, \text{ per ogni } n.$$

Questo prodotto è (p_n), dove

$$p_n = \sum_{i+j=n} a_i b_j.$$

In particolare

$$\begin{aligned} p_0 &= a_0 b_0 = 1 = 1 \cdot 1, \\ p_1 &= a_1 b_0 + a_0 b_1 = (-1) \cdot 1 + 1 \cdot 1 = 0, \\ p_2 &= a_2 b_0 + a_1 b_1 + a_0 b_2 = 0 \cdot 1 + (-1) \cdot 1 + 1 \cdot 1 = 0. \end{aligned}$$

Più in generale, sapendo che $a_n = 0$, per $n > 1$, si ha, sempre per $n > 1$,

$$p_n = \sum_{i+j=n} a_i b_j = a_1 b_{n-1} + a_0 b_n = (-1) \cdot 1 + 1 \cdot 1 = 0.$$

In altre parole la successione (p_n) è proprio l'elemento neutro del nuovo anello.

1.2. Serie

Il risultato precedente non dovrebbe essere troppo sorprendente, se ci si ricorda della teoria delle serie di potenze. In effetti, è noto che

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$$

per ogni numero reale (o anche complesso) z tale che $|z| < 1$. Quello che abbiamo scritto prima è una "dimostrazione" di questa identità. Occorre una piccola osservazione: una serie di potenze è determinata univocamente dalla successione dei coefficienti. In questo caso si tratta proprio della costante 1.

Che poi stiamo lavorando in modo analogo alle serie di potenze risulta dal fatto che il prodotto da noi definito sulle successioni corrisponde proprio al *prodotto di Cauchy* delle serie.

Proviamo di nuovo a dimostrare l'identità:

$$\begin{aligned} (1-z) \sum_{n=0}^{\infty} z^n &= \sum_{n=0}^{\infty} z^n - \sum_{n=0}^{\infty} z^{n+1} \\ &= 1 + \sum_{n=1}^{\infty} z^n - \sum_{n=1}^{\infty} z^n = 1. \end{aligned}$$

Nell'ultimo passaggio abbiamo spezzato la prima somma, mettendo in evidenza il termine per $n = 0$, e abbiamo cambiato l'indice nella seconda, con $n + 1 \to n$. Dov'è l'inghippo in questa "dimostrazione"?

Non si tratta di una dimostrazione scorretta, tuttavia abbiamo usato proprietà delle serie che vanno giustificate; altrimenti si possono ottenere assurdità come

$$\ldots \frac{1}{z^n} + \cdots + \frac{1}{z^2} + \frac{1}{z} + 1 + z + z^2 + \cdots + z^n + \cdots = 0.$$

Infatti, la somma a sinistra fino all'1 è quella di una progressione geometrica di ragione $1/z$, quindi è

$$\frac{1}{1-\dfrac{1}{z}} = \frac{z}{z-1};$$

quella dall'1 in poi è $1/(1-z)$ e, avendo contato due volte il termine 1, la somma è

$$\frac{z}{z-1} + \frac{1}{1-z} - 1 = 0.$$

È evidente che il ragionamento è assurdo; lo è per molti motivi, ma il principale è che non abbiamo tenuto in minimo conto la convergenza. Infatti la serie a sinistra converge solo quando $|z| > 1$ e quella di destra solo quando $|z| < 1$: non esistono valori nei quali entrambe le serie convergano.

Nel caso della serie geometrica non ci sono difficoltà, per via dei noti teoremi sulle serie assolutamente convergenti, che permettono manipolazioni algebriche su queste serie.

Eppure il ragionamento della sezione precedente funzionava senza parlare di convergenza, poiché si basava solo sulle proprietà algebriche delle operazioni. Abbiamo applicato la definizione e trovato esattamente quello che ci serviva.

Proviamo a risolvere un'altra questione: data una successione (a_n) in $A[\![x]\!]$, cerchiamo la sua inversa (b_n). Dobbiamo quindi risolvere le infinite equazioni

$$\begin{aligned}
&a_0 b_0 = 1,\\
&a_1 b_0 + a_0 b_1 = 0\\
&a_2 b_0 + a_1 b_1 + a_0 b_2 = 0\\
&\dots\\
&\sum_{i=0}^{n} a_{n-i} b_i = 0\\
&\dots
\end{aligned}$$

Dalla prima ricaviamo che $b_0 = a_0^{-1}$, cioè che a_0 deve essere invertibile in A. In tal caso possiamo anche risolvere tutte le altre: per $n > 0$ si deve avere

$$b_n = -a_0^{-1} \sum_{i=0}^{n-1} a_{n-i} b_i$$

e, per induzione, abbiamo definito la successione (b_n) che cercavamo. È ovvio che, con questa scelta, $(a_n)(b_n) = 1$, dove 1 denota l'elemento neutro di $A[\![x]\!]$, cioè la successione $(1, 0, \dots, 0, \dots)$.

Ad esempio, per $a_0 = 1$, $a_1 = -1$ e $a_n = 0$ per $n > 1$, le equazioni sono: $b_0 = 1$, $b_1 = (-1)\cdot((-1)\cdot 1) = 1$, e, per induzione,

$$b_n = -a_0^{-1}(a_1 b_{n-1}) = (-1)\cdot((-1)\cdot 1) = 1.$$

Però possiamo anche trovare l'inversa della successione $(n!)$ che ha senso per esempio quando $A = \mathbf{R}$:

$$\begin{aligned}
b_0 &= (0!)^{-1} = 1\\
b_1 &= -(0!)^{-1}(1!\, b_0) = -1\\
b_2 &= -(0!)^{-1}(2!\, b_0 + 1!\, b_1) = -1\\
b_3 &= -(0!)^{-1}(3!\, b_0 + 2!\, b_1 + 1!\, b_2) = -3\\
b_4 &= -(0!)^{-1}(4!\, b_0 + 3!\, b_1 + 2!\, b_2 + 1!\, b_3) = -13\\
&\dots
\end{aligned}$$

Questo non ha alcun corrispettivo nelle serie di potenze: infatti la serie di potenze $\sum_n (n!)z^n$ converge solo per $z = 0$. Tuttavia la manipolazione algebrica dà un risultato ben definito; vedremo presto che cosa significa.

Un altro esempio: cerchiamo l'inversa della successione $(a_n) = (1, -1, -1, 0, \dots, 0, \dots)$. Avremo

$$\begin{aligned} b_0 &= a_0^{-1} = 1 \\ b_1 &= -a_0^{-1}(a_1 b_0) = 1 \\ b_2 &= -a_0^{-1}(a_2 b_0 + a_1 b_1) = 2 \\ b_3 &= -a_0^{-1}(a_3 b_0 + a_2 b_1 + a_1 b_2) = 3 \\ &\dots \\ b_{n+2} &= -a_0^{-1}(a_2 b_n + a_1 b_{n+1}) = b_n + b_{n+1} \end{aligned}$$

quindi la successione cercata è esattamente la *successione di Fibonacci*. Esercizio: trovare il raggio di convergenza della serie che ha come coefficienti quelli della successione di Fibonacci.

1.3. Serie formali

L'anello $A[\![x]\!]$ si chiama *anello delle serie formali a coefficienti in A*. Cominciamo a vedere in modo più rigoroso alcune delle sue proprietà. Per rendere le cose più semplici, tratteremo il caso in cui A è un campo; questo è comunque il caso più importante.

Useremo alcune notazioni fisse; gli elementi di $A[\![x]\!]$ saranno denotati con lettere come f e g. Se $\mathsf{f} \in A[\![x]\!]$, essa è una successione, i cui termini si indicano con f_n. Perciò

$$\mathsf{f} = (\mathsf{f}_0, \mathsf{f}_1, \dots, \mathsf{f}_n, \dots).$$

Il minimo n tale che $\mathsf{f}_n \neq 0$ si chiama *ordine* di f e si indica con $o(\mathsf{f})$. Per lo zero di $A[\![x]\!]$, che denoteremo con $\mathsf{0}$, l'ordine è ∞.

Un particolare elemento di ordine n è

$$x^n = (0, \dots, 0, \underset{\substack{\uparrow \\ n}}{1}, 0, \dots, 0, \dots).$$

In effetti ogni polinomio è un elemento di $A[\![x]\!]$, quindi $A[x]$ è un sottoanello di $A[\![x]\!]$.

Cominciamo con un fatto che abbiamo già visto.

Proposizione 1.3.1. *Un elemento $\mathsf{f} \in A[\![x]\!]$ è invertibile se e solo se $\mathsf{f}_0 \neq 0$, cioè se e solo se $o(\mathsf{f}) = 0$.*

Proposizione 1.3.2. *Se $\mathsf{f}, \mathsf{g} \in A[\![x]\!]$, allora:*

(1) $o(\mathsf{f} + \mathsf{g}) \geq \min\{o(\mathsf{f}), o(\mathsf{g})\}$;
(2) $o(\mathsf{fg}) = o(\mathsf{f}) + o(\mathsf{g})$.

Dimostrazione. Se $n < \min\{o(\mathsf{f}), o(\mathsf{g})\}$, abbiamo $\mathsf{f}_n = \mathsf{g}_n = 0$ e quindi $(\mathsf{f} + \mathsf{g})_n = 0$.

Se $\mathsf{f} = \mathsf{0}$, allora $\mathsf{fg} = \mathsf{0g} = \mathsf{0}$ e l'uguaglianza è provata. Possiamo allora supporre che entrambe le successioni siano non nulle.

Poniamo $a = o(\mathsf{f})$ e $b = o(\mathsf{g})$; allora

$$(\mathsf{fg})_{a+b} = \sum_{i+j=a+b} \mathsf{f}_i \mathsf{g}_j = \mathsf{f}_a \mathsf{g}_b \neq 0;$$

inoltre, se $n < a + b$,

$$(\mathsf{fg})_n = \sum_{i+j=n} \mathsf{f}_i \mathsf{g}_j = 0,$$

perché $\mathsf{f}_i = 0$, per $i < a$, e $\mathsf{g}_j = 0$, per $j < b$. □

Ricordiamo che due elementi a, b di un anello commutativo si dicono *associati* se esiste un elemento invertibile u tale che $b = au$.

LEMMA 1.3.3. *Se $\mathsf{f} \in A[\![x]\!]$ è non nullo, con $o(\mathsf{f}) = d$, allora f è associato a x^d.*

DIMOSTRAZIONE. Abbiamo $\mathsf{f}_i = 0$ per $i < d$ e $\mathsf{f}_d \neq 0$. Possiamo allora considerare la successione $\mathsf{g} \in A[\![x]\!]$ tale che

$$\mathsf{g}_n = \mathsf{f}_{d+n}.$$

Si verifica immediatamente che $\mathsf{f} = x^d\mathsf{g}$; poiché g è invertibile, si ha la tesi. □

È chiaro che x^d non è associato a x^k, se $d \neq k$.

PROPOSIZIONE 1.3.4. *L'anello $A[\![x]\!]$ è un dominio a ideali principali e ammette un unico ideale massimale, formato dagli elementi non invertibili.*

DIMOSTRAZIONE. Che $A[\![x]\!]$ sia un dominio segue dalla Proposizione 1.3.2, perché, se $\mathsf{f} \neq \mathsf{0}$ e $\mathsf{g} \neq \mathsf{0}$, l'ordine di fg è finito.

Consideriamo l'insieme degli elementi non invertibili, che sono quelli di ordine > 0. La somma di due tali elementi è allora non invertibile. Se moltiplichiamo un elemento non invertibile per uno qualunque, otteniamo un elemento non invertibile (questo vale in ogni anello, esercizio). Dunque gli elementi non invertibili di $A[\![x]\!]$ formano un ideale, che è per forza massimale (esercizio; suggerimento: se un ideale contiene propriamente questo, deve contenere un elemento invertibile).

Vediamo ora che ogni ideale I di $A[\![x]\!]$ è principale. La dimostrazione è simile a quella che abbiamo dato per $A[x]$. Se $I = \{\mathsf{0}\}$, allora $I = \mathsf{0}A[\![x]\!]$. Supponiamo allora che $I \neq \{\mathsf{0}\}$; in tal caso in I troviamo elementi di ordine finito e quindi uno di ordine minimo, diciamo d. Pertanto, per il lemma 1.3.3, $x^d \in I$ e dunque $x^d A[\![x]\!] \subseteq I$.

Verifichiamo allora l'inclusione inversa. Sia $\mathsf{f} \in I$; allora $o(\mathsf{f}) \geq d$. Possiamo allora scrivere $\mathsf{f} = x^k\mathsf{g}$, dove g è invertibile. Ma allora

$$\mathsf{f} = x^k\mathsf{g} = x^d x^{k-d}\mathsf{g} \in x^d A[\![x]\!],$$

come si voleva. □

COROLLARIO 1.3.5. *L'insieme degli ideali di $A[\![x]\!]$ è totalmente ordinato per inclusione.*

L'idea di considerare l'anello $A[\![x]\!]$ è di trattare in modo algebrico le serie di potenze, dimenticandosi della convergenza; i risultati che si ottengono possono nondimeno avere applicazioni anche alle funzioni sviluppabili in serie di potenze, purché si sia cauti.

Identificheremo la successione $\mathsf{f} \in A[\![x]\!]$ con la serie (formale) di potenze

$$\sum_n \mathsf{f}_n x^n;$$

questa notazione ha solo lo scopo di facilitare i calcoli e le identità, come vedremo fra poco.

Le serie formali infatti hanno notevole importanza nello studio delle successioni definite per ricorrenza. Torniamo all'esempio della successione di Fibonacci, per $A = \mathbf{R}$:

$$\mathsf{F}_0 = 1, \quad \mathsf{F}_1 = 1, \quad \mathsf{F}_n = \mathsf{F}_{n-1} + \mathsf{F}_{n-2} \ (n > 1).$$

Abbiamo già visto che $\mathsf{F} = (1 - x - x^2)^{-1}$; cerchiamo di scrivere i termini della successione in un altro modo.

Cerchiamo di scrivere

$$\frac{1}{(1 - x - x^2)} = \frac{A}{1 - \alpha x} + \frac{B}{1 - \beta x}.$$

Questa è una procedura ben nota, ad esempio nel calcolo degli integrali delle funzioni razionali. È sufficiente che

$$\begin{aligned} A + B &= 1 \\ A\beta + B\alpha &= 0 \\ \alpha + \beta &= 1 \\ \alpha\beta &= -1 \end{aligned}$$

come polinomi, in particolare che α e β siano radici del polinomio $x^2 - x - 1$. Se poniamo

$$\alpha = \varphi = \frac{1+\sqrt{5}}{2}$$

che si chiama spesso *numero di Fidia*, avremo

$$\beta = \frac{1-\sqrt{5}}{2} = 1 - \varphi.$$

Allora $B = 1 - A$ e, da $A(1-\varphi) + (1-A)\varphi = 0$ troviamo $A(2\varphi - 1) = \varphi$, cioè

$$A = \frac{\varphi}{\sqrt{5}}, \qquad B = -\frac{1-\varphi}{\sqrt{5}}.$$

Perché scrivere così? La risposta è che conosciamo già quanto vale $(1-kx)^{-1}$ come serie formale di potenze:

$$(1-kx)^{-1} = \sum_n k^n x^n$$

perché si tratta della solita serie geometrica. Quindi:

$$\begin{aligned} \frac{1}{1-x-x^2} &= \frac{1}{\sqrt{5}}\left(\frac{\varphi}{1-\varphi x} - \frac{1-\varphi}{1-(1-\varphi)x}\right) \\ &= \frac{1}{\sqrt{5}}\left(\varphi \sum_n \varphi^n x^n + (1-\varphi)\sum_n (1-\varphi)^n x^n\right) \\ &= \frac{1}{\sqrt{5}} \sum_n \bigl(\varphi^{n+1} - (1-\varphi)^{n+1}\bigr) x^n \end{aligned}$$

e dunque

$$\mathsf{F}_n = \frac{1}{\sqrt{5}}\bigl(\varphi^{n+1} - (1-\varphi)^{n+1}\bigr)$$

Per esercizio, controllare per i casi piccoli. Siccome

$$\left|\frac{(1-\varphi)^{n+1}}{\sqrt{5}}\right| < 1/2,$$

un metodo per calcolare F_n è di arrotondare $\varphi^{n+1}/\sqrt{5}$ all'intero più vicino.

Le serie di potenze formali danno un metodo potente per lo studio delle successioni definite per ricorrenza, come si vedrà.

1.4. Derivate e primitive

Non possiamo considerare una serie di potenze formale come una funzione; tuttavia possiamo lo stesso usare alcuni metodi dell'analisi.

Se $\mathsf{f} \in A[\![x]\!]$, definiamo la *derivata* di f come

$$\mathsf{f}' = \sum_n n \mathsf{f}_n x^{n-1}.$$

Per essere pignoli occorrerebbe scrivere $\mathsf{f}' = \mathsf{g}$, dove

$$\mathsf{g}_n = (n+1)\mathsf{f}_{n+1}.$$

Definiamo anche la *serie esponenziale*

$$\mathsf{E} = \sum_n \frac{1}{n!} x^n,$$

cioè $\mathsf{E}_n = 1/n!$. È allora evidente che $\mathsf{E}' = \mathsf{E}$. Naturalmente, per fare questo, è necessario che la caratteristica del campo A sia 0. Anche questa è un'ipotesi che faremo sempre, da ora in poi.

PROPOSIZIONE 1.4.1. *Se* $\mathsf{f}, \mathsf{g} \in A[\![x]\!]$ *e* $\mathsf{f}' = \mathsf{g}'$*, allora esiste* $a \in A$ *tale che* $\mathsf{f} - \mathsf{g} = a$*.*

DIMOSTRAZIONE. È banale, basta scrivere la condizione. □

Una possibilità in più che abbiamo è di *sostituire un serie in un'altra*. In generale questo non ha senso, ad esempio non possiamo sostituire una costante, perché ciò equivarrebbe a calcolare una somma infinita. Invece, se prendiamo f e g dove $o(\mathsf{g}) > 0$, la cosa si può fare:

$$\mathsf{f}(\mathsf{g}) = \sum_n \mathsf{f}_n \mathsf{g}^n.$$

Infatti $o(\mathsf{g}^n) \geq n$, quindi ogni x^n ha solo un numero finito di coefficienti non nulli. Ad esempio è possibile scrivere

$$\mathsf{f}(-x) = \sum_n (-1)^n \mathsf{f}_n x^n,$$

perciò

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

Esercizio: dimostrare che $\mathsf{E}(-x) = \mathsf{E}^{-1}$.

PROPOSIZIONE 1.4.2. *Se* $f, g \in A[\![x]\!]$ *hanno ordine maggiore di zero, allora*

$$\mathsf{E}(f+g) = \mathsf{E}(f) \cdot \mathsf{E}(g).$$

DIMOSTRAZIONE. Calcoliamo senza preoccuparci troppo delle verifiche, che comunque sono possibili; ci ricordiamo solo del teorema del binomio che vale in ogni anello commutativo:

$$\mathsf{E}(f+g) = \sum_n \frac{(f+g)^n}{n!} = \sum_n \left(\sum_{i=0}^n \frac{1}{n!} \binom{n}{i} f^i g^{n-i} \right).$$

D'altra parte

$$\mathsf{E}(f) \cdot \mathsf{E}(g) = \left(\sum_n \frac{f^n}{n!} \right) \left(\sum_n \frac{g^n}{n!} \right) = \sum_n \left(\sum_{i=0}^n \frac{f^i}{i!} \frac{g^{n-i}}{(n-i)!} \right).$$

La verifica è conclusa ricordando che

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

perciò le due serie sono uguali. □

PROPOSIZIONE 1.4.3. *Se* $\mathsf{f}, \mathsf{g} \in A[\![x]\!]$*, allora*

$$(\mathsf{fg})' = \mathsf{f}'\mathsf{g} + \mathsf{fg}'.$$

PROPOSIZIONE 1.4.4. *Se* $\mathsf{f}, \mathsf{g} \in A[\![x]\!]$ *e* $o(\mathsf{g}) > 0$*, allora*

$$(\mathsf{f}(\mathsf{g}))' = \mathsf{f}'(\mathsf{g}) \cdot \mathsf{g}'.$$

La dimostrazione non è troppo difficile, ma la omettiamo.

Una serie con ordine 1 è quella *logaritmica*:

$$\mathsf{L} = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} x^n,$$

cioè $\mathsf{L}_0 = 0$ e $\mathsf{L}_n = (-1)^{n-1}/n$ per $n \geq 1$. Questa serie dovrebbe fare le veci di $\log(1+x)$. In effetti è possibile dimostrare che $\mathsf{E}(\mathsf{L}) = 1 + x$.

Calcoliamo la derivata di L:

$$\mathsf{L}' = \sum_n (-1)^n x^n = (1+x)^{-1}.$$

Per $a \in A$, definiamo

$$(1+x)^a = \mathsf{E}(a\mathsf{L}).$$

Le proprietà di questa serie sono:

(1) $(1+x)^0 = 1$;
(2) $(1+x)^a \cdot (1+x)^b = (1+x)^{a+b}$.

Vogliamo trovare una espressione dei coefficienti della serie $(1+x)^a$. Poniamo

$$\binom{a}{0} = 1, \qquad \binom{a}{n} = \frac{a(a-1)\dots(a-n+1)}{n!} \ (n > 0).$$

Dalla proposizione sulle "funzioni composte" abbiamo il seguentre risultato.

Corollario 1.4.5. *La derivata di* $(1+x)^a$ *è* $a(1+x)^{a-1}$.

Proposizione 1.4.6. *Vale la formula del binomio di Newton*

$$(1+x)^a = \sum_n \binom{a}{n} x^n.$$

Dimostrazione. Poniamo $\mathsf{f}_a = \sum_n \binom{a}{n} x^n$. Allora

$$\mathsf{f}_a' = \sum_{n\geq 1} n\binom{a}{n} x^{n-1}.$$

Ma abbiamo anche

$$n\binom{a}{n} = a\binom{a-1}{n-1}$$

quindi

$$\mathsf{f}_a' = \sum_{n\geq 1} a\binom{a-1}{n-1} x^{n-1} = a\sum_m \binom{a-1}{m} x^m = a\mathsf{f}_{a-1}.$$

La stessa dimostrazione che vale nei numeri interi dice che

$$\binom{a-1}{n} + \binom{a-1}{n-1} = \binom{a}{n} \quad (n \geq 1).$$

Perciò:

$$\sum_{n\geq 1} \binom{a-1}{n} x^n + \sum_{n\geq 1} \binom{a-1}{n-1} x^n = \sum_{n\geq 1} \binom{a}{n} x^n,$$

cioè

$$(\mathsf{f}_{a-1} - 1) + x\mathsf{f}_{a-1} = \mathsf{f}_a$$

e quindi

$$\mathsf{f}_{a-1} = (1+x)^{-1}\mathsf{f}_a$$

da cui

$$\mathsf{f}_a' = a(1+x)^{-1}\mathsf{f}_a.$$

Ne segue che la derivata di $\mathsf{f}_a \cdot (1+x)^{-a}$ è

$$\begin{aligned}(\mathsf{f}_a \cdot (1+x)^{-a})' &= \mathsf{f}_a' \cdot (1+x)^{-a} + \mathsf{f}_a \cdot ((1+x)^{-a})' \\ &= a(1+x)^{-1}\mathsf{f}_a \cdot (1+x)^{-a} + \mathsf{f}_a \cdot (-a)(1+x)^{-a-1} \\ &= 0.\end{aligned}$$

Di conseguenza $\mathsf{f}_a \cdot (1+x)^{-a} \in A$ (è "costante") e un esame dei coefficienti dice che

$$\mathsf{f}_a \cdot (1+x)^{-a} = 1.$$

Basta allora moltiplicare per $(1+x)^a$ per ottenere il risultato. □

Esiste anche l'operatore di *integrazione*:

$$\int \mathsf{f} = \sum_n \frac{\mathsf{f}_n}{n+1} x^{n+1}.$$

In altre parole $\int \mathsf{f}$ è la successione g dove $\mathsf{g}_0 = 0$ e $\mathsf{g}_n = \mathsf{f}_{n-1}$, per $n \geq 1$. Chiaramente la derivata dell'integrale di f è f.

1.5. Relazioni di ricorrenza

Una *relazione di ricorrenza lineare* è un'espressione del tipo

$$t_n = c_1 t_{n-1} + c_2 t_{n-2} + \cdots + c_k t_{n-k}, \tag{*}$$

con $c_1, c_2, \ldots, c_k \in A$. Una *soluzione* della relazione di ricorrenza data è $f \in A[\![x]\!]$ tale che

$$\mathsf{f}_n = c_1 \mathsf{f}_{n-1} + c_2 \mathsf{f}_{n-2} + \cdots + c_k \mathsf{f}_{n-k}$$

per $n \geq k$. Diremo che k è l'*ordine* della relazione di ricorrenza.

Proposizione 1.5.1. *Se* f *e* g *sono soluzioni di una data relazione di ricorrenza e* $a, b \in A$, *allora anche* $a\mathsf{f} + b\mathsf{g}$ *è una soluzione.*

Siccome $A[\![x]\!]$ è uno spazio vettoriale su A, le soluzioni di una relazione di ricorrenza formano un suo sottospazio.

Proposizione 1.5.2. *La successione* $(a^n) = \sum_n a^n x^n$ *è una soluzione di* (*) *se e solo se* a *è una radice del polinomio*

$$t^k - c_1 t^{k-1} - c_2 t^{k-2} - \cdots - c_{k-1} t - c_k.$$

Se per esempio A è il campo complesso, possiamo scrivere tutte le radici del polinomio e trovare k soluzioni linearmente indipendenti. Se le radici sono distinte il problema è risolto, perché si può dimostrare che lo spazio delle soluzioni ha dimensione proprio k. Se una radice a ha molteplicità d, si può verificare che

$$\sum_n a^n x^n, \quad \sum_n n a^n x^n, \quad \sum_n n^2 a^n x^n, \quad \cdots \sum_n n^{d-1} a^n x^n$$

sono anch'esse soluzioni di (*), linearmente indipendenti. Dunque anche in questo caso le soluzioni si possono determinare.

Come esempio, rivediamo la successione di Fibonacci. Abbiamo che il polinomio *caratteristico* è

$$t^2 - t - 1$$

le cui radici sono φ e $1 - \varphi$. Quindi ogni soluzione è della forma

$$\sum_n (a\varphi^n + b(1-\varphi)^n) x^n$$

e possiamo determinare a e b imponendo le condizioni iniziali $\mathsf{F}_0 = 1$, $\mathsf{F}_1 = 1$. Infatti questo dice che

$$\begin{cases} a + b = 1 \\ a\varphi + b(1-\varphi) = 1 \end{cases}$$

che dà subito $a = \varphi/\sqrt{5}$ e $b = -(1-\varphi)/\sqrt{5}$.

Se la relazione di ricorrenza è invece

$$t_n = -4t_{n-1} - 5t_{n-2} - 2t_{n-3},$$

il polinomio caratteristico è $x^3 + 4x^2 + 5x + 2 = (x+1)^2(x+2)$ e quindi le soluzioni sono tutte e sole quelle della forma

$$\sum_n (a(-1)^n + bn(-1)^n + c(-2)^n)x^n$$

e basterà imporre le condizioni iniziali.

1.6. Numeri di Catalan e Bell

Sia $*$ un'operazione non necessariamente associativa sull'insieme X. Se $a_1, a_2, \ldots, a_n$ sono elementi di X, quanti prodotti con questi fattori posso eseguire? Ad esempio, se $n = 3$ oppure 4, abbiamo le seguenti possibilità:

$$a_1 * (a_2 * a_3), \qquad (a_1 * a_2) * a_3;$$

$$\begin{cases} a_1 * (a_2 * (a_3 * a_4)), & a_1 * ((a_2 * a_3) * a_4), \\ (a_1 * a_2) * (a_3 * a_4), & \\ (a_1 * (a_2 * a_3)) * a_4, & ((a_1 * a_2) * a_3) * a_4. \end{cases}$$

Chiamiamo C_n il numero di modi con n fattori e cerchiamo una relazione di ricorrenza anche per questa successione. Porremo $\mathsf{C}_0 = 0$, $\mathsf{C}_1 = 1$ e $\mathsf{C}_2 = 1$.

PROPOSIZIONE 1.6.1. *Per $n \geq 2$ si ha*

$$\mathsf{C}_n = \sum_{i=1}^{n-1} \mathsf{C}_i \mathsf{C}_{n-i}.$$

DIMOSTRAZIONE. Si deve considerare l'ultima operazione eseguita, cioè il simbolo $*$ che sta fuori dalle parentesi; se a sinistra ci sono i fattori, a destra ce ne sono $n - i$. I modi per scrivere un'espressione di questo tipo sono allora $\mathsf{C}_i\mathsf{C}_{n-i}$. □

Consideriamo allora $\mathsf{C} = \sum_n \mathsf{C}_n x^n$. Possiamo scrivere

$$\begin{aligned} \mathsf{C} &= \mathsf{C}_0 + \mathsf{C}_1 x + \sum_{n\geq 2} \mathsf{C}_n x^n \\ &= x + \sum_{n\geq 2} \left(\sum_{i=1}^{n-1} \mathsf{C}_i \mathsf{C}_{n-i} \right) x^n \\ &= x + \sum_{n\geq 2} \left(\sum_{i=0}^{n} \mathsf{C}_i \mathsf{C}_{n-i} \right) x^n \\ &= x + \sum_{n} \left(\sum_{i=0}^{n} \mathsf{C}_i \mathsf{C}_{n-i} \right) x^n \\ &= x + \mathsf{C}^2. \end{aligned}$$

Ne segue che

$$\mathsf{C}^2 - \mathsf{C} + x = 0$$

da cui

$$0 = \mathsf{C}^2 - \mathsf{C} + \frac{1}{4} - \frac{1}{4} + x$$

e quindi

$$\left(\mathsf{C} - \frac{1}{2}\right)^2 = \frac{1}{4} - x$$

Conosciamo una serie di potenze che elevata al quadrato dia $1 - 4x$, cioè

$$(1 - 4x)^{\frac{1}{2}}.$$

Poiché $A[\![x]\!]$ è un dominio, abbiamo che

$$\mathsf{C} - \frac{1}{2} = \frac{1}{2}(1 - 4x)^{\frac{1}{2}} \quad \text{oppure} \quad \mathsf{C} - \frac{1}{2} = -\frac{1}{2}(1 - 4x)^{\frac{1}{2}}$$

e dobbiamo scegliere la seconda, perché $\mathsf{C}_0 = 0$. Dunque

$$\mathsf{C} = \frac{1}{2} - \frac{1}{2}(1 - 4x)^{\frac{1}{2}} = \frac{1}{2} - \frac{1}{2}\sum_n \binom{1/2}{n}(-4)^n x^n.$$

Si tratta di semplificare il coefficiente (per $n \geq 1$)

$$\begin{aligned}
\mathsf{C}_n = -\frac{1}{2}\binom{1/2}{n}(-4)^n &= -\frac{1}{2}(-1)^n 2^{2n}\frac{\frac{1}{2}\left(\frac{1}{2} - 1\right)\dots\left(\frac{1}{2} - n + 1\right)}{n!} \\
&= -\frac{1}{2}(-1)^n 2^{2n}\frac{1(-1)(-3)\dots(3-2n)}{2^n n!} \\
&= -\frac{1}{2}(-1)^n 2^{2n}(-1)^{n-1}\frac{(2n-3)!!}{2^n n!} \\
&= 2^{n-1}\frac{(2n-3)!!}{n!}
\end{aligned}$$

dove $r!!$ denota il *semifattoriale* di r:

$$r!! = \begin{cases} 1 \cdot 3 \cdot \ldots \cdot r & \text{se } r \text{ è dispari,} \\ 2 \cdot 4 \cdot \ldots \cdot r & \text{se } r > 0 \text{ è pari,} \\ 1 & \text{se } r = 0. \end{cases}$$

Ora, se r è dispari, diciamo $r = (2s - 1)$, abbiamo

$$r!! = 1 \cdot 3 \cdot \ldots \cdot (2s-1)\frac{2 \cdot 4 \cdot \ldots \cdot 2s}{2^s \cdot s!} = \frac{(2s)!}{2^s \cdot s!}.$$

Nel nostro caso $r = 2n - 3 = 2(n-1) - 1$ e quindi $s = n - 1$; quindi

$$\mathsf{C}_n = 2^{n-1}\frac{(2n-2)!}{2^{n-1}(n-1)!\,n!} = \frac{1}{n}\frac{(2n-2)!}{(n-1)!\,(n-1)!} = \frac{1}{n}\binom{2n-2}{n-1}.$$

Come altra applicazione calcoliamo i *numeri di Bell*, anche se in realtà non arriveremo a una formula esplicita. Vogliamo determinare quante sono le possibili relazioni di equivalenza in un insieme con n elementi. Una relazione di equivalenza è data da una partizione e quindi ci basta contare queste. Sia dunque B_n il numero delle partizioni di un insieme con n elementi, che possiamo supporre sia $\{1, 2, \dots, n\}$ se $n > 0$. Avremo $\mathsf{B}_0 = 1$, perché c'è un'unica partizione dell'insieme vuoto; inoltre è chiaro che $\mathsf{B}_1 = 1$ e anche che $\mathsf{B}_2 = 2$, dal momento che le partizioni di $\{1, 2\}$ sono

$$\{\{1\}, \{2\}\} \qquad \text{e} \qquad \{\{1, 2\}\}.$$

Per $n = 3$ le partizioni sono:

$$\{\{1\}, \{2\}, \{3\}\}, \quad \{\{1\}, \{2, 3\}\}, \quad \{\{2\}, \{1, 3\}\}, \quad \{\{3\}, \{1, 2\}\}, \quad \{\{1, 2, 3\}\},$$

dunque $\mathsf{B}_3 = 5$. Cerchiamo una relazione di ricorrenza; se $\mathscr{F}$ è una partizione di $X = \{1, 2, \dots, n\}$, con $n > 1$, l'elemento n deve stare in esattamente un elemento A di $\mathscr{F}$. Questo A è un sottoinsieme di $\{1, 2, \dots, n\}$ ed avrà i elementi. Se togliamo A dalla partizione, otteniamo una partizione dell'insieme $X \setminus A$, che ha $n - i$ elementi, e di queste ce sono B_{n-i}. I possibili modi di scegliere un sottoinsieme A di X che abbia i elementi e al quale n appartenga sono esattamente il numero di sottoinsiemi di $\{1, 2, \dots, n-1\}$ che hanno $i - 1$ elementi, cioè $\binom{n-1}{i-1}$. Dunque, sommando,

$$\mathsf{B}_n = \sum_{i=1}^{n} \binom{n-1}{i-1} \mathsf{B}_{n-i}.$$

Ad esempio

$$\begin{aligned}
\mathsf{B}_3 &= \binom{3-1}{1-1}\mathsf{B}_{3-1} + \binom{3-1}{2-1}\mathsf{B}_{3-2} + \binom{3-1}{3-1}\mathsf{B}_{3-3} \\
&= \binom{2}{0}\mathsf{B}_2 + \binom{2}{1}\mathsf{B}_1 + \binom{2}{2}\mathsf{B}_0 \\
&= 2 + 2 + 1 = 5; \\
\mathsf{B}_4 &= \binom{4-1}{1-1}\mathsf{B}_{4-1} + \binom{4-1}{2-1}\mathsf{B}_{4-2} + \binom{4-1}{3-1}\mathsf{B}_{4-3} + \binom{4-1}{4-1}\mathsf{B}_{4-3} \\
&= \binom{3}{0}\mathsf{B}_3 + \binom{3}{1}\mathsf{B}_2 + \binom{3}{2}\mathsf{B}_1 + \binom{3}{3}\mathsf{B}_0 \\
&= 5 + 3 \cdot 2 + 3 \cdot 1 + 1 = 15.
\end{aligned}$$

Invece di calcolare direttamente B, calcoleremo l'espressione di

$$\hat{\mathsf{B}} = \sum_n \frac{\mathsf{B}_n}{n!} x^n$$

cioè porremo $\hat{\mathsf{B}} = \mathsf{B}_n/n!$. Questa si chiama *funzione generatrice esponenziale*. Dimostreremo con metodi analoghi ai precedenti che

$$\hat{\mathsf{B}} = \mathsf{E}(\mathsf{E} - 1).$$

Notiamo che la sostituzione eseguita ha senso in astratto, visto che $o(\mathsf{E} - 1) = 1$. Visto che però stiamo lavorando nei reali, questo dimostra che

$$\sum_n \frac{\mathsf{B}_n}{n!} x^n = e^{e^x - 1}.$$

Calcoliamo la derivata di $\hat{\mathsf{B}}$:

$$\hat{\mathsf{B}}' = \sum_{n \geq 1} \frac{\mathsf{B}_n}{n!} n x^{n-1} = \sum_{n \geq 1} \left(\sum_{i=1}^{n} \binom{n-1}{i-1} \mathsf{B}_{n-1} \right) \frac{1}{(n-1)!} x^{n-1}.$$

Invece il prodotto $\mathsf{E} \cdot \hat{\mathsf{B}}$ vale:

$$\mathsf{E} \cdot \hat{\mathsf{B}} = \left(\sum_n \frac{1}{n!} x^n \right) \left(\sum_n \frac{\mathsf{B}_n}{n!} x^n \right) = \sum_m \left(\sum_{j=0}^{m} \frac{1}{j!} \frac{\mathsf{B}_{m-j}}{(m-j)!} \right) x^m.$$

Cambiamo j in $i - 1$ e m in $n - 1$, ottenendo che

$$\mathsf{E} \cdot \hat{\mathsf{B}} = \sum_{n \geq 1} \left(\sum_{i=1}^{n} \frac{1}{(i-1)!} \frac{\mathsf{B}_{n-i}}{(n-i)!} \right) x^{n-1} = \sum_{n \geq 1} \left(\sum_{i=1}^{n} \binom{n-1}{i-1} \frac{\mathsf{B}_{n-i}}{(n-1)!} \right) x^{n-1}$$

e quindi che $\hat{\mathsf{B}}' = \mathsf{E} \cdot \hat{\mathsf{B}}$. Ma allora

$$(\mathsf{E}(1-\mathsf{E}) \cdot \hat{\mathsf{B}})' = \mathsf{E}(1-\mathsf{E}) \cdot (-\mathsf{E}) \cdot \hat{\mathsf{B}} + \mathsf{E}(1-\mathsf{E}) \cdot \mathsf{E} \cdot \hat{\mathsf{B}} = 0.$$

Di conseguenza $\mathsf{E}(1-\mathsf{E}) \cdot \hat{\mathsf{B}}$ è costante e basta calcolare in 0 per ottenere che la costante è 1. Finalmente

$$\hat{\mathsf{B}} = (\mathsf{E}(1-\mathsf{E}))^{-1} = \mathsf{E}(\mathsf{E}-1).$$

Terminiamo con il calcolo di un'altra successione di una certa importanza; vogliamo trovare il numero T_n delle permutazioni su n oggetti che non abbiano punti uniti; diremo che $\sigma \in S_n$ non ha punti uniti se $\sigma(i) \neq i$, per ogni $i = 1, 2, \dots, n$. Un punto unito corrisponde quindi a un ciclo di lunghezza 1 nella decomposizione in cicli disgiunti; quindi dobbiamo calcolare il numero delle permutazioni che si scrivono come

$$\sigma = (a_{11}a_{12}\dots a_{1d_1})(a_{21}a_{22}\dots a_{2d_2})\dots(a_{k1}a_{k2}\dots a_{kd_k})$$

dove $1 < d_i \leq n$ e $d_1 + d_2 + \cdots + d_k = n$.

PROPOSIZIONE 1.6.2. *Si ha* $\mathsf{T}_0 = 1$, $\mathsf{T}_1 = 0$ *e*

$$\mathsf{T}_{n+1} = n(\mathsf{T}_n + \mathsf{T}_{n-1}) \quad (n > 0).$$

DIMOSTRAZIONE. Sia $n > 0$ e scriviamo una permutazione $\sigma \in S_{n+1}$ senza punti uniti in cicli disgiunti come prima; possiamo supporre che $a_{11} = n+1$ e quindi, per ipotesi, $1 \leq \sigma(n+1) \leq n$. Possiamo allora ripartire l'insieme S'_{n+1} delle permutazioni di S_{n+1} senza punti uniti in sottoinsiemi disgiunti

$$A_i = \{\, \sigma \in S'_{n+1} \mid \sigma(n+1) = i \,\}.$$

A loro volta gli insiemi A_i sono ripartiti in due sottoinsiemi disgiunti:

$$B_i = \{\, \sigma \in A_i \mid d_1 = 2 \,\} \quad \text{e} \quad C_i = \{\, \sigma \in A_i \mid d_1 > 2 \,\}.$$

Una permutazione $\sigma \in B_i$ è allora della forma

$$\sigma = (n+1\ i)(a_{21}a_{22}\dots a_{2d_2})\dots(a_{k1}a_{k2}\dots a_{kd_k})$$

e quindi

$$\sigma' = (a_{21}a_{22}\dots a_{2d_2})\dots(a_{k1}a_{k2}\dots a_{kd_k})$$

è una permutazione senza punti uniti di un insieme con $n-1$ elementi, precisamente $\{1, \dots, n\} \setminus \{i\}$; anzi, da ogni permutazione senza punti uniti di questo insieme possiamo ottenerne una di $\{1, \dots, n, n+1\}$ aggiungendo il ciclo $(n+1\ i)$. Perciò il numero di elementi in B_i è esattamente T_{n-1}.

Analogamente una permutazione $\sigma \in C_i$ è della forma

$$\sigma = (n+1\ i\ a_{13}\dots a_{1d_1})(a_{21}a_{22}\dots a_{2d_2})\dots(a_{k1}a_{k2}\dots a_{kd_k})$$

e ad essa possiamo associare la permutazione senza punti uniti

$$\sigma'' = (i\ a_{13}\dots a_{1d_1})(a_{21}a_{22}\dots a_{2d_2})\dots(a_{k1}a_{k2}\dots a_{kd_k})$$

di $\{1, \dots, n\}$. Ancora, ad una permutazione senza punti uniti di $\{1, \dots, n\}$ possiamo associarne una di C_i. Dunque C_i ha esattamente T_n elementi.

Poiché i si può scegliere in n modi, abbiamo la tesi. □

Cercheremo ancora la funzione generatrice esponenziale

$$\hat{\mathsf{T}} = \sum_n \frac{\mathsf{T}}{n!} x^n.$$

Dalla relazione di ricorrenza ricaviamo che

$$\sum_{n\geq 1} \frac{\mathsf{T}_{n+1}}{n!} x^n = \sum_{n\geq 1} \frac{\mathsf{T}_{n-1}}{(n-1)!} x^n + \sum_{n\geq 1} \frac{\mathsf{T}_n}{(n-1)!} x^n.$$

Ricordando che $\mathsf{T}_0 = \mathsf{T}_1 = 0$, abbiamo

$$\sum_{n\geq 1} \frac{\mathsf{T}_{n+1}}{n!} x^n = \sum_{n\geq 2} \frac{\mathsf{T}_n}{(n-1)!} x^{n-1} = \sum_{n\geq 2} \frac{\mathsf{T}_n}{n!} n x^{n-1} = \hat{\mathsf{T}}'.$$

Invece

$$\sum_{n\geq 1} \frac{\mathsf{T}_{n-1}}{(n-1)!} x^n = x \sum_{n\geq 1} \frac{\mathsf{T}_{n-1}}{(n-1)!} x^{n-1} = x\hat{\mathsf{T}}$$

e

$$\sum_{n\geq 1} \frac{\mathsf{T}_n}{(n-1)!} x^n = x \sum_{n\geq 1} \frac{\mathsf{T}_n}{n!} n x^{n-1} = x\hat{\mathsf{T}}'.$$

Dunque $\hat{\mathsf{T}}' = x\hat{\mathsf{T}} + x\hat{\mathsf{T}}'$, quindi

$$(1-x)\hat{\mathsf{T}}' = x\hat{\mathsf{T}},$$

da cui

$$\hat{\mathsf{T}}' = \frac{x}{1-x}\hat{\mathsf{T}}.$$

Il trucco per "risolvere" questa equazione è di fingere che sia un'ordinaria equazione differenziale:

$$\frac{y'}{y} = \frac{x}{1-x} = -1 + \frac{1}{1-x}$$

da cui

$$\log|y| = -x - \log(1-x) + log|c|$$

e quindi

$$y = ce^{-x}\frac{1}{1-x}.$$

La condizione iniziale dice che $y(0) = 1$, quindi $c = 1$. Il metodo in realtà è simile a quello usato per la successione dei numeri di Bell. Si mostra cioè che la derivata di $\mathsf{E} \cdot (1-x) \cdot \hat{\mathsf{T}}$ è zero. Di conseguenza

$$\hat{\mathsf{T}} = \frac{\mathsf{E}(-x)}{1-x}.$$

Da questa possiamo ricavare esplicitamente T. Infatti

$$\hat{\mathsf{T}} = \left(\sum_n \frac{(-1)^n}{n!} x^n\right)\left(\sum_n x^n\right)$$

e dunque

$$\mathsf{T}_n = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}.$$

CAPITOLO 2

Campi finiti

2.1. Estensioni algebriche

Ricordiamo che un *campo* è un anello commutativo F con $1 \neq 0$ in cui ogni elemento non nullo è invertibile. Diremo che un campo K è *estensione* di un campo F se F è un sottoanello di K. Scriveremo, per abbreviare, K/F, anche se questa notazione non ha nulla a che fare con insiemi quoziente.

Se K è un'estensione di F, ogni polinomio nell'indeterminata x a coefficienti in F può essere considerato anche a coefficienti in K; per essere più precisi, $F[x]$ è un sottoanello di $K[x]$.

Un'altra proprietà importante è che K può essere considerato come spazio vettoriale su F, quindi si possono applicare tutte le nozioni che conosciamo sugli spazi vettoriali.

DEFINIZIONE 2.1.1. L'estensione K/F si dice *finita* se la dimensione di K come spazio vettoriale su F è finita. In tal caso questa dimensione si indica con $[K : F]$.

È talvolta importante considerare estensioni di estensioni. Vale un'importante proprietà.

TEOREMA 2.1.2. *Sia K un'estensione di F e sia L un'estensione di K. Allora L/F è finita se e solo se L/K e K/F sono finite e, in tal caso, $[L : F] = [L : K][K : F]$.*

DIMOSTRAZIONE. Supponiamo che L/K e K/F siano finite; allora esistono una base $\mathscr{B} = \{l_1, \ldots, l_m\}$ di L come spazio vettoriale su K e una base $\mathscr{C} = \{k_1, \ldots, k_n\}$ di K come spazio vettoriale su F. Dimostriamo che

$$\mathscr{D} = \{\, l_i k_j \mid 1 \leq i \leq m, 1 \leq j \leq n \,\}$$

è una base di L come spazio vettoriale su F.

(1) *$\mathscr{D}$ è un insieme di generatori.* Infatti, se $l \in L$, possiamo scrivere $l = \sum_{i=1}^{m} \alpha_i l_i$, con $\alpha_1, \ldots, \alpha_m \in K$, dal momento che $\mathscr{B}$ è una base di L/K. Siccome $\mathscr{C}$ è una base di K/F, ciascun k_i può essere scritto come

$$\alpha_i = \sum_{j=1}^{n} \beta_{ij} k_j$$

con $\beta_{ij} \in F$. Ne segue che

$$l = \sum_{i=1}^{m} \alpha_i l_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} \beta_{ij} k_j \Big) l_i = \sum_{i,j} \beta_{ij} l_i k_j.$$

(2) *$\mathscr{D}$ è linearmente indipendente.* Supponiamo che

$$\sum_{i,j} \beta_{ij} l_i k_j = 0$$

con $\beta_{ij} \in F$. Allora

$$0 = \sum_{i,j} \beta_{ij} l_i k_j = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} \beta_{ij} k_j \Big) l_i$$

e quindi, essendo $\mathscr{B}$ linearmente indipendente,

$$\sum_{j=1}^{n} \beta_{ij} k_j = 0 \qquad (i = 1, 2, \ldots, m).$$

Siccome anche $\mathscr{C}$ è linearmente indipendente, abbiamo che

$$\beta_{ij} = 0$$

per ogni i e j.

Viceversa, supponiamo che L/F sia finita. Allora esiste un insieme di generatori finito di L come spazio vettoriale su F e, a maggior ragione, sarà anche un insieme di generatori di L come spazio vettoriale su K. Per finire, K è un sottospazio vettoriale di L/F, quindi ha dimensione finita. □

DEFINIZIONE 2.1.3. Sia K un'estensione di F. Un elemento $b \in K$ si dice *algebrico* su F se esiste un polinomio $f \in F[x]$ non nullo tale che $f(b) = 0$. Un elemento $b \in K$ si dice *trascendente* su F se non è algebrico. Diciamo che l'estensione K/F è algebrica se ogni elemento di K è algebrico su F.

Come esempio classico, $\sqrt{2} \in \mathbf{R}$ è algebrico su $\mathbf{Q}$, poiché è radice del polinomio $x^2 - 2$. Un famoso risultato di Lindemann e Weierstrass dice che π è trascendente su $\mathbf{Q}$.

TEOREMA 2.1.4. *Se K/F è finita, allora è algebrica.*

DIMOSTRAZIONE. Sia $n = [K : F]$ e sia $b \in K$. Allora gli elementi 1, b, $\ldots$, b^n sono linearmente dipendenti (sono più della dimensione) e perciò esistono $a_0, a_1, \ldots, a_n \in F$, non tutti nulli, tali che $a_0 \cdot 1 + a_1 b + \cdots + a_n b^n = 0$. Questo significa proprio che il polinomio $f = a_0 + a_1 x + \cdots + a_n x^n$ è non nullo e che $f(b) = 0$. □

Il viceversa non è vero, come vedremo più avanti.

Si pone un problema: dato un polinomio non nullo in $F[x]$, è possibile trovare un'estensione K/F dove il polinomio ammetta una radice? Ad esempio, il polinomio $x^2 + 1 \in \mathbf{R}[x]$ non ha radici, ma ne ha nell'estensione $\mathbf{C}$.

Proprio come abbiamo ampliato i reali per ottenere i complessi, si può risolvere il problema di prima in generale.

LEMMA 2.1.5. *Sia $b \in K$ algebrico su F; allora l'insieme $I_b = \{\, f \in F[x] \mid f(b) = 0 \,\}$ è un ideale non nullo di $F[x]$.*

DIMOSTRAZIONE. Chiaramente il polinomio nullo appartiene a I_b; se $f, g \in I_b$, abbiamo che $v_b(f + g) = v_b(f) + v_b(g) = 0 + 0$, quindi $f + g \in I_b$. Se poi $f \in I_b$ e $g \in F[x]$, abbiamo che $v_b(fg) = v_b(f)v_b(g) = 0$. □

Ricordiamo che ogni ideale di $F[x]$ è principale, quindi esiste un unico polinomio monico $f_b \in I_b$ tale che $I_b = f_b F[x]$, cioè ogni polinomio in I_b è della forma $f_b g$, per un opportuno $g \in F[x]$. Questo polinomio è il polinomio monico di grado minimo in I_b e si chiama *polinomio minimo di b*.

PROPOSIZIONE 2.1.6. *Sia $b \in K$ algebrico su F; allora il polinomio minimo $f_b \in F[x]$ è irriducibile.*

DIMOSTRAZIONE. Se $f_b = gh$, con $g, h \in F[x]$, allora $0 = f_b(b) = g(b)h(b)$, quindi $g \in I_b$ oppure $h \in I_b$. Nel primo caso, siccome f_b ha grado minimo, dovremo avere che il grado di g non è minore del grado di f_b, quindi g e f_b hanno lo stesso grado e h ha grado 0, cioè è invertibile. L'altro caso è analogo. □

Esercizio: dimostrare che, se $g \in I_b$ è irriducibile e monico, allora è il polinomio minimo di b.

Attenzione: occorre sempre dire qual è l'estensione che si sta considerando. Ad esempio, il polinomio minimo di $i\sqrt[4]{2} \in \mathbf{C}$ su $\mathbf{Q}$ è $x^4 + 2$, mentre su $\mathbf{R}$ è $x^2 + \sqrt{2}$.

Se consideriamo una famiglia di sottocampi del campo K, la loro intersezione è un sottocampo. Analogamente per sottoanelli. Di conseguenza i concetti di cui parleremo ora sono ben definiti.

DEFINIZIONE 2.1.7. Se $b \in K$, estensione di F, indichiamo con $F[b]$ il minimo sottoanello di K che contenga sia F che b. Con $F(b)$ indichiamo il minimo sottocampo di K che contiene sia F che b.

È ovvio che $F[b] \subseteq F(b)$. Quando saranno uguali? Ci occorre una descrizione di $F[b]$. Consideriamo l'omomorfismo di valutazione

$$\begin{aligned} v_b \colon F[x] &\to K \\ f &\mapsto f(b) \end{aligned}$$

e sia B la sua immagine. Allora B è un sottoanello di K che contiene sia F che b; dunque $F[b] \subseteq B$. Ma un elemento di B si scrive come $f(b)$, per un opportuno $f \in F[x]$, cioè $a_0 + a_1 b + \cdots + a_n b^n$ e questo elemento deve appartenere ad ogni sottoanello di K che contenga sia F che b. Perciò $B = F[b]$.

PROPOSIZIONE 2.1.8. *Se b è algebrico su F, allora $F[b]$ è un campo. Se b è trascendente su F, allora $F[b]$ è isomorfo a $F[x]$.*

DIMOSTRAZIONE. Sia b algebrico. Il teorema di omomorfismo dice che l'immagine di v_b è un anello isomorfo a $F[x]/I_b$, perché I_b è il nucleo di v_b. Siccome I_b è l'ideale principale generato da un elemento irriducibile, esso è massimale, quindi l'immagine di v_b, che è proprio $F[b]$ è un campo.

Se invece b è trascendente, il nucleo di v_b è l'ideale nullo, quindi v_b è iniettivo. Ne segue che $F[b]$ è isomorfo a $F[x]$. □

COROLLARIO 2.1.9. *Se b è algebrico su F, allora $F[b] = F(b)$.*

Notiamo come la dimostrazione "astratta" permetta di evitare la verifica che ogni elemento di $F[b]$ è invertibile. La stessa dimostrazione però può essere resa "algoritmica": vogliamo trovare un metodo per calcolare gli inversi. Prendiamo dunque $b \in K$ algebrico su F e il suo polinomio minimo f_b. Un elemento di $F[a]$ è della forma $c = g(b)$, per un opportuno $g \in F[x]$. L'algoritmo della divisione in $F[x]$ dice che esistono e sono unici $q, r \in F[x]$ tali che

$$g = f_b q + r, \qquad \partial(r) < \partial(f_b).$$

Ne segue che

$$c = g(b) = v_b(g) = v_b(f_b q + r) = v_b(f_b) v_b(q) + v_b(r) = v_b(r).$$

Perciò $c \neq 0$ se e solo se $r(b) \neq 0$. Nel caso in cui $c \neq 0$ possiamo allora dire che $\mathrm{mcd}(f_b, r) = 1$, in quanto $r \neq 0$ ha grado minore del grado di f_b che è irriducibile. Per il teorema di Bézout, che vale anche il $F[x]$, esistono due polinomi $s, t \in F[x]$ tali che

$$sr + t f_b = 1.$$

Valutiamo questa identità in b:

$$1 = v_b(1) = v_b(sr + t f_b) = v_b(s) v_b(r) + v_b(t) v_b(f_b) = s(b) r(b).$$

Non solo abbiamo trovato l'inverso di c, cioè $s(b)$; abbiamo anche visto che ogni elemento di $F[b]$ si scrive come $r(b)$, dove $r \in F[x]$ e $\partial(r) < \partial(f_b)$. Questa scrittura è unica: infatti, se $r_1(b) = r_2(b)$ con $\partial(r_1), \partial(r_2) < \partial(f_b)$, abbiamo che $r_1 - r_2 \in I_b$ e ha grado minore di $\partial(f_b)$, quindi $r_1 - r_2 = 0$.

TEOREMA 2.1.10. *Sia $b \in K$ algebrico su F. Allora $F[b]$ è il minimo sottocampo di K che contiene sia F che b. Inoltre*

$$[F[b] : F] = \partial(f_b)$$

cioè la dimensione di $F[b]$ come spazio vettoriale su F è uguale al grado n del polinomio minimo di b su F e una base è data da $\{1, b, \dots, b^{n-1}\}$.

DIMOSTRAZIONE. Ci basta dimostrare l'asserzione sulla base. Abbiamo visto che ogni elemento di $F[b]$ si scrive in modo unico come $r(b)$, dove $r \in F[x]$ e $\partial(r) < \partial(f)$. Un polinomio r di questo tipo è della forma $a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$ e quindi gli elementi di $F[b]$ si scrivono in modo unico come

$$a_0 \cdot 1 + a_1 b + \cdots + a_{n-1}b^{n-1}$$

e questo dice proprio che ogni elemento di $F[b]$ si scrive in modo unico come combinazione lineare a coefficienti in F di 1, b, $\dots$, b^{n-1}. □

Ritorniamo allora al problema di trovare un'estensione opportuna di F in cui un certo polinomio $f \in F[x]$ abbia radici. La costruzione sarà iterativa; intanto, siccome è possibile scomporre f in fattori irriducibili, non è restrittivo supporre che f sia irriducibile. In questo caso l'ideale $I = fF[x]$ è massimale in $F[x]$ e quindi l'anello quoziente $F[x]/I$ è un campo. Inoltre l'applicazione

$$\begin{aligned} \alpha\colon F &\to F[x]/I \\ a &\mapsto [a] \end{aligned}$$

è un omomorfismo iniettivo di anelli. Perciò, invece di scrivere $[a]$ possiamo scrivere a, identificando questi elementi. In tal modo $K = F[x]/I$ diventa un'estensione di F. Poniamo $b = [x] \in K$. Abbiamo allora, scrivendo $f = a_0 + a_1 x + \cdots + a_n x^n$,

$$\begin{aligned} v_b(f) &= a_0 + a_1 b + \cdots + a_n b^n = [a_0] + [a_1]b + \cdots + [a_n]b^n \\ &= [a_0] + [a_1][x] + \cdots + [a_n][x]^n = [a_0 + a_1 x + \cdots + a_n x^n] = [0] = 0. \end{aligned}$$

Il problema è risolto! Infatti K è proprio un'estensione di F in cui f ammette una radice. Possiamo allora scrivere $f = (x - b)f_1$, dove $f_1 \in K[x]$ ha grado più piccolo di f. Questo dà il via alla procedura iterativa.

TEOREMA 2.1.11. *Se $f \in F[x]$ è monico e ha grado $n > 0$, esiste un'estensione K di F tale che:*

(1) *$f = (x - b_1)(x - b_2)\dots(x - b_n)$ in $K[x]$;*
(2) *$K = F[b_1, b_2, \dots, b_n]$*

Se L/F è un'altra estensione con le stesse proprietà, allora esiste un isomorfismo $\alpha\colon K \to L$ tale che $\alpha(a) = a$, per ogni $a \in F$.

Non daremo la dimostrazione di questo teorema, che richiederebbe troppo tempo. L'estensione così determinata, che è allora unica a meno di isomorfismi, si chiama un *campo di riducibilità completa di f*. Con $F[b_1, b_2, \dots, b_n]$, dove gli elementi $b_i \in K$, intendiamo il minimo sottoanello di K che contiene $F \cup \{b_1, \dots, b_n\}$; è facile dimostrare per induzione che $F[b_1, b_2, \dots, b_n]$ è un sottocampo se gli elementi b_i sono algebrici su F. Analogamente $F(b_1, \dots, b_n)$ è il minimo sottocampo di K che contiene $F \cup \{b_1, \dots, b_n\}$.

C'è un'altra proprietà importante delle estensioni algebriche.

TEOREMA 2.1.12. *Sia K/F un'estensione di campi. Se $b, b_1, b_2 \in K$ sono algebrici su F, allora $b_1 + b_2$, $b_1 - b_2$, $b_1 b_2$ e $-b$ sono algebrici su F; se $b \neq 0$, anche b^{-1} è algebrico su F.*

DIMOSTRAZIONE. Consideriamo $F(b_1, b_2)$; è ovvio dalla definizione che $F(b_1, b_2) = F(b_1)(b_2)$. Siccome b_2 è algebrico su F, esso è a maggior ragione algebrico su $F(b_1)$; quindi

$$[F(b_1, b_2) : F] = [F(b_1)(b_2), F(b_1)][F(b_1) : F]$$

è finita e quindi ogni elemento di $F(b_1, b_2)$ è algebrico su F. Quindi lo sono $b_1 + b_2$, $b_1 - b_2$, $b_1 b_2$. Analogamente per $-b$ e b^{-1}. □

COROLLARIO 2.1.13. *Se K è un'estensione di F, allora l'insieme degli elementi di K che sono algebrici su F è un sottocampo di K.*

Diamo senza dimostrazione un famoso teorema dovuto a Steinitz. Prima però una definizione: un campo K si dice *algebricamente chiuso* se ogni polinomio a coefficienti in K, di grado > 0, ammette una radice. Il cosiddetto teorema fondamentale dell'algebra dice che $\mathbf{C}$ è algebricamente chiuso. Un'estensione K di F si dice una *chiusura algebrica* di F se K è algebricamente chiuso e ogni elemento di K è algebrico su F.

TEOREMA 2.1.14. *Ogni campo F ha una chiusura algebrica. Se K e L sono chiusure algebriche di F, allora esiste un isomorfismo $\alpha\colon K \to L$ tale che $\alpha(a) = a$, per ogni $a \in F$.*

Possiamo allora dire, usando gli ultimi due enunciati, che l'insieme $\bar{\mathbf{Q}}$ dei numeri complessi algebrici su $\mathbf{Q}$ è una chiusura algebrica di $\mathbf{Q}$. Dal momento che esistono numeri complessi trascendenti su $\mathbf{Q}$, abbiamo $\bar{\mathbf{Q}} \neq \mathbf{C}$. Chiaramente, invece, $\mathbf{C}$ è una chiusura algebrica di $\mathbf{R}$, poiché ogni numero complesso è radice di un polinomio a coefficienti reali: infatti $a + bi$ è radice di $x^2 - 2ax + (a^2 + b^2)$.

2.2. Campi finiti

Ricordiamo la definizione di *caratteristica* di un anello A: esiste un unico omomorfismo di anelli $\chi_A\colon \mathbf{Z} \to A$, definito da $\chi_A(n) = n1$, multiplo secondo n di $1 \in A$. Il nucleo di questo omomorfismo è un ideale di $\mathbf{Z}$, quindi della forma $k\mathbf{Z}$, dove $k \geq 0$. La caratteristica di A è allora questo numero k. Se k è la caratteristica di A, $ka = 0$, per ogni $a \in A$.

Sappiamo che, se A è un dominio, la sua caratteristica è zero oppure un numero primo. In particolare questo vale per un campo. Se in più sappiamo che il campo F è finito, la sua caratteristica deve essere un numero primo p: infatti l'omomorfismo χ_A non può essere iniettivo, quindi $\ker \chi_A \neq \{0\} = 0\mathbf{Z}$.

PROPOSIZIONE 2.2.1. *Se la caratteristica dell'anello commutativo A è un numero primo p, allora l'applicazione*

$$\begin{aligned} \Phi\colon A &\to A \\ a &\mapsto a^p \end{aligned}$$

è un omomorfismo di anelli.

DIMOSTRAZIONE. Che $\Phi(1) = 1$ e $\Phi(ab) = (ab)^p = a^p b^p = \Phi(a)\Phi(b)$ è chiaro. Inoltre per il teorema del binomio

$$\Phi(a + b) = (a + b)^p = \sum_{i=1}^{p} \binom{p}{i} a^i b^{p-i}.$$

Se $0 < i < p$, abbiamo che

$$\binom{p}{i} = \frac{p!}{i!\,(p - i)!}$$

è divisibile per p, quindi $\binom{p}{i} a^i b^{p-i} = 0$. Di conseguenza $(a + b)^p = a^p + b^p$. □

Naturalmente allora anche $\Phi^2 = \Phi \circ \Phi$, $\ldots$, $\Phi^{n+1} = \Phi^n \circ \Phi$, eccetera, sono omomorfismi.

Un'altro fatto importante è che un anello commutativo di caratteristica k ha un sottoanello isomorfo a $\mathbf{Z}/k\mathbf{Z}$, precisamente l'immagine di χ_A, che si chiama *sottoanello primo* di A; nel caso in cui k è un numero primo, questo è un sottocampo.

TEOREMA 2.2.2. *Sia F un campo finito e sia p la sua caratteristica. Allora*

$$|F| = p^n$$

per un opportuno intero $n \geq 1$.

DIMOSTRAZIONE. Sia F_0 il sottocampo primo di F; allora $F_0 \cong \mathbf{Z}/p\mathbf{Z}$ ha p elementi. La dimensione di F come spazio vettoriale su F_0 è finita, perché F stesso è un insieme di generatori; se questa dimensione è $n = [F : F_0]$, ogni elemento di F si scrive in modo unico come combinazione lineare di n elementi a coefficienti in F_0; il numero di tali combinazioni lineari è proprio p^n. □

Il *piccolo teorema di Fermat* ammette una generalizzazione.

TEOREMA 2.2.3. *Sia F un campo con q elementi. Allora, per ogni $a \in F$, si ha $a^q = a$.*

DIMOSTRAZIONE. L'insieme $F \setminus \{0\}$ è, rispetto alla moltiplicazione, un gruppo con $q - 1$ elementi. Perciò $a^{q-1} = 1$, per ogni $a \in F \setminus \{0\}$. Di conseguenza $a^q = a$ e l'identità vale anche per $a = 0$. □

Il teorema sull'unicità di un campo di riducibilità completa ha come conseguenza che due campi con lo stesso numero (finito) di elementi sono isomorfi. Infatti questi due campi hanno lo stesso sottocampo primo (per meglio dire sottocampi primi isomorfi, ma si possono identificare). Il polinomio $x^q - x$ ha coefficienti nel sottocampo primo; basta allora verificare che un campo con q elementi è il campo di riducibilità completa di $x^q - x$ sul sottocampo primo.

Sia dunque $F = \{a_1, a_2, \ldots, a_q\}$ un campo con q elementi. Per quanto visto prima, ogni elemento di F è radice del polinomio $f = x^q - x$; dunque f si scrive come prodotto di fattori di primo grado e certamente $F = F_0(a_1, \ldots, a_q)$, dove F_0 è il sottocampo primo.

LEMMA 2.2.4. *Sia F un campo di caratteristica p e sia q una potenza di p; allora il polinomio $x^q - x$ non ha radici di molteplicità ≥ 2.*

DIMOSTRAZIONE. Supponiamo che $a \in F$ sia una radice ed eseguiamo la divisione di $f = x^q - x$ per $x - a$:

$$\begin{array}{c|cccccc|c} & 1 & 0 & 0 & \ldots & 0 & -1 & 0 \\ a & & a & a^2 & \ldots & a^{q-2} & a^{q-1} & a^q - a \\ \hline & 1 & a & a^2 & \ldots & a^{q-2} & a^{q-1} - 1 & 0 \end{array}$$

Dunque il quoziente è

$$g = x^{q-1} + ax^{q-2} + a^2x^{q-3} + \cdots + a^{q-2}x + (a^{q-1} - 1)$$

e se valutiamo in a otteniamo

$$g(a) = qa^{q-1} - 1 = -1 \neq 0,$$

dal momento che q è una potenza di p. Perciò a non è una radice di g e abbiamo la tesi. □

TEOREMA 2.2.5. *Se $q = p^n$ è potenza di un primo p, allora esiste, a meno di isomorfismi, un unico campo con q elementi.*

DIMOSTRAZIONE. Abbiamo già osservato l'unicità. Per quanto riguarda l'esistenza, ci basta considerare di nuovo il polinomio $f = x^q - x$ a coefficienti in $\mathbf{Z}/p\mathbf{Z}$. Sia K un campo di riducibilità completa di f, che sappiamo esistere. A priori non è detto che K abbia proprio q elementi; consideriamo dunque l'insieme F delle radici di f in K. Allora F è un sottocampo di K.

Infatti $1 = 1^q \in F$; se poi $a, b \in F$, abbiamo

$$(a + b)^q = (a + b)^{p^n} = \Phi^n(a + b) = \Phi^n(a) + \Phi^n(b) = a + b$$

e lo stesso ragionamento con la moltiplicazione prova che $ab \in F$. Inoltre $(a^{-1})^q = (a^q)^{-1} = a^{-1}$, se $a \neq 0$. Per l'opposto abbiamo $(-a)^q = -a^q = -a$, se p è dispari; se invece $p = 2$, abbiamo $-a = a$, e non c'è altro da dimostrare.

Per il lemma precedente F ha q elementi, perché le sue radici in K sono tutte distinte e ne ha q per l'ipotesi che K è un campo di riducibilità completa per f. Ne segue allora che $F = K$ ha q elementi. □

Il campo con $q = p^n$ elementi, unico a meno di isomorfismi, si denota con $GF(q)$. In particolare $GF(p) = \mathbf{Z}/p\mathbf{Z}$.

Come si costruisce esplicitamente un campo con $q = p^n$ elementi? Ad esempio, supponiamo di voler determinare $GF(9)$. Secondo il teorema precedente, dobbiamo calcolare un campo di riducibilità completa di $x^9 - x$ su $GF(3)$. Prima di tutto scriviamolo come prodotto di fattori irriducibili:

$$x^9 - x = x(x-1)(x-2)(x^2+1)(x^4+1).$$

Se esaminiamo $g = x^2 + 1$, notiamo che è irriducibile in $GF(3)$ e quindi l'anello quoziente $GF(3)[x]/g\ GF(3)[x]$ è un campo che ha dimensione 2 su $GF(3)$ e perciò è $GF(9)$. La sua base è $\{1, b\}$, dove b ha la proprietà che $b^2 = -1$; quindi gli elementi si scrivono come $\alpha + \beta b$, con l'addizione per componenti e la moltiplicazione

$$(\alpha + \beta b)(\gamma + \delta b) = \alpha\gamma + \alpha\delta b + \beta\gamma b + \beta\delta b^2 = (\alpha\gamma - \beta\delta) + (\alpha\delta + \beta\gamma)b.$$

Le radici di $x^2 + 1$ sono $b = 0 + 1b$ e $-b = 0 + 2b$. Cerchiamo le radici di $x^4 + 1$, che sappiamo devono esistere. Si ha

$$x^4 + 1 = x^4 - b^2 = (x^2 - b)(x^2 + b).$$

Affinché $\alpha + \beta b$ sia radice di $x^2 - b$ occorre che

$$\alpha^2 - \beta^2 = 0 \qquad \text{e} \qquad 2\alpha\beta = 1.$$

Abbiamo quindi le soluzioni $1 + 2b$ e $2 + b$. Si lascia come esercizio la verifica che gli altri due elementi $1 + b$ e $2 + 2b$ sono proprio le radici di $x^2 + b$.

Se prendiamo invece come polinomio irriducibile $x^2 - x - 1$ e indichiamo con c una sua radice, gli elementi di $GF(9)$ si scrivono come $\alpha + \beta c$ con moltiplicazione

$$(\alpha + \beta c)(\gamma + \delta c) = \alpha\gamma + \alpha\delta c + \beta\gamma c + \beta\delta c^2 = (\alpha\gamma + \beta\delta) + (\alpha\delta + \beta\gamma + \beta\delta)c,$$

perché c soddisfa l'identità $c^2 = 1 + c$. Questa scelta può essere più conveniente perché le potenze di c sono tutti gli elementi non nulli di $GF(9)$. Infatti

$$\begin{aligned}
c^2 &= 1 + c &&= 1 + c,\\
c^3 &= c + c^2 &&= 1 + 2c,\\
c^4 &= c + 2c^2 &&= 2,\\
c^5 &= 2c &&= 2c,\\
c^6 &= 2c^2 &&= 2 + 2c,\\
c^7 &= 2c + 2c^2 &&= 2 + c,\\
c^8 &= 2c + c^2 &&= 1.
\end{aligned}$$

Esercizio: determinare l'isomorfismo fra i campi costruiti nei due modi.

Non occorre costruire passo passo i campi $GF(p^n)$; infatti basta considerare un unico polinomio $f \in GF(p)[x]$ che abbia grado n e sia irriducibile. Un tale polinomio esiste sempre ed è determinabile in tempo finito: basta scrivere $n = i + j$ in tutti i modi possibili con $i, j \geq 1$ e calcolare tutti i prodotti dei polinomi di grado i e j in $GF(p)[x]$; quelli di grado n che non si ottengono così sono certamente irriducibili.

Vogliamo come altro esempio determinare $GF(8)$. Ci serve un polinomio $f \in GF(2)[x]$ di grado 3 e irriducibile. Una scelta possibile è $f = 1 + x + x^3$ e quindi $GF(8)$ ha come base su $GF(2)$ gli elementi 1, b e b^2, dove $1 + b + b^3 = 0$. Dunque la moltiplicazione è

$$\begin{aligned}&(\alpha_1 + \beta_1 b + \gamma_1 b^2)(\alpha_2 + \beta_2 b + \gamma_2 b^2)\\ &= (\alpha_1\alpha_2 + \beta_1\gamma_2 + \gamma_1\beta_2) + (\alpha_1\beta_2 + \beta_1\alpha_2 + \beta_1\gamma_2 + \gamma_1\beta_2 + \gamma_1\gamma_2)b + (\alpha_1\gamma_2 + \beta_1\beta_2 + \gamma_1\alpha_2 + \gamma_1\gamma_2)b^2\end{aligned}$$

2.3. Chiusura algebrica

Il teorema di Steinitz dice che ogni campo ammette una chiusura algebrica. La costruzione rigorosa della chiusura algebrica $GF(p^\infty)$ di $GF(p)$ richiederebbe troppo tempo, non perché sia intrinsecamente difficile, quanto perché coinvolge concetti abbastanza delicati.

Tuttavia non è necessario, per molti scopi, considerare la chiusura algebrica. Quello che basta sapere è come costruire, dato un numero finito di polinomi in $GF(p^n)[x]$, un campo che ne contenga tutte le radici. Ovviamente ci basterebbe considerare un solo polinomio f, il prodotto di quelli dati, e costruirne il campo di riducibilità completa.

Il problema è già risolto se nella decomposizione di f non compaiono fattori di grado > 1. Supponiamo allora che si trovi un fattore g di grado > 1.

(1) $F_0 = GF(p^n)$;
(2) $F_1 = F_1[x]/gF_1[x]$ è un campo in cui g ha una radice;
(3) si decompone f in fattori irriducibili in $F_1[x]$;
(4) se esiste ancora un fattore irriducibile di grado > 1 si eseguono di nuovo i passi precedenti ottenendo via via campi F_2, F_3, ..., F_l, fino a che non risultano più fattori di grado > 1.

Il procedimento ha ovviamente termine. In modo analogo si potrebbe procedere per costruire la chiusura algebrica di $GF(p)$.

(1) $F_0 = GF(p)$;
(2) si prende un polinomio irriducibile $f_1 \in F_0[x]$ di grado minimo > 1;
(3) $F_1 = F_0[x]/fF_0[x]$ è un'estensione di F_0 in cui f_1 ha una radice;
(4) si prende un polinomio irriducibile $f_2 \in F_1[x]$ di grado minimo > 1;
(5) ...

Il procedimento va ripetuto all'infinito: infatti nessun campo finito può essere algebricamente chiuso. La dimostrazione di questo fatto è analoga alla prova dell'infinità dei numeri primi.

Se $F = \{a_1, \ldots, a_q\}$ è un campo finito, si considera il polinomio

$$f = (x - a_1)(x - a_2)\ldots(x - a_q) + 1.$$

Allora nessun elemento di F è radice di f e quindi F non è algebricamente chiuso.

CAPITOLO 3

Curve algebriche

3.1. Polinomi in più indeterminate e curve

Se A è un anello commutativo, possiamo costruire l'anello dei polinomi in una indeterminata $A[x]$. A partire da questo possiamo ancora costruire l'anello dei polinomi in una indeterminata, chiamiamola questa volta y. Gli elementi di questo nuovo anello sono allora le espressioni

$$f_0 + f_1 y + f_2 y^2 + \cdots + f_n y^n,$$

dove $f_0, f_1, \ldots, f_n \in A[x]$. Un modo diverso e più informale di presentare questo anello è di considerare le *combinazioni lineari a coefficienti in A dei monomi $x^i y^j$* al variare di i e j nei numeri naturali. La moltiplicazione si esegue usando le proprietà di anello e la commutatività.

Ovviamente il procedimento si può ripetere induttivamente:

$$A[x_1, x_2] = A[x_1][x_2], \quad A[x_1, x_2, \ldots, x_{n-1}, x_n] = A[x_1, x_2, \ldots, x_{n-1}][x_n].$$

Chiameremo $A[x_1, x_2, \ldots, x_{n-1}, x_n]$ *anello dei polinomi in n indeterminate a coefficienti in A*. Anche questo è l'insieme delle combinazioni lineari a coefficienti in A dei monomi $x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$. Il *grado totale di un monomio* è la somma degli esponenti; il *grado totale di un polinomio* è il massimo dei gradi totali dei monomi con coefficienti non nulli. Il grado del polinomio nullo è $-\infty$.

Come per l'anello dei polinomi in una indeterminata vale una proprietà importante che riguarda la sostituzione.

PROPOSIZIONE 3.1.1. *Se A è un sottoanello dell'anello commutativo B e $b_1, b_2, \ldots, b_n \in B$, allora esiste un unico omomorfismo di anelli*

$$v_{b_1, b_2, \ldots, b_n} \colon A[x_1, x_2, \ldots, x_n] \to B$$

tale che

$$v_{b_1, b_2, \ldots, b_n}(x_i) = b_i \qquad (i = 1, 2, \ldots, n)$$

e $v_{b_1, b_2, \ldots, b_n}(a) = a$ per ogni $a \in A$.

Non dimostreremo il teorema seguente, che tuttavia è intuitivamente plausibile.

TEOREMA 3.1.2. *Se F è un campo, ogni polinomio non costante in $F[x_1, x_2, \ldots, x_n]$ può essere scritto in modo essenzialmente unico come prodotto di polinomi irriducibili.*

3.2. Curve algebriche

Se F è un campo e $n > 0$ denoteremo con F^n l'insieme delle n-uple di elementi di F; questo è lo *spazio affine* di dimensione n su F. Di particolare importanza è il caso di $n = 2$, per il quale parleremo del *piano affine* su F.

DEFINIZIONE 3.2.1. Dato un polinomio non nullo $f \in F[x, y]$, la *curva algebrica* definita da f è l'insieme

$$C(f) = \{\, (a, b) \in F^2 \mid f(a, b) = 0 \,\}.$$

Il grado totale di f si dice anche grado di $C(f)$. Una curva algebrica di grado due è una *conica*, una di grado tre è una *cubica*. Nel caso in cui il grado totale di f sia 1, $C(f)$ è una *retta*.

Useremo la solita terminologia geometrica: gli elementi del piano affine sono chiamati *punti*; si parla di retta o curva passante per un punto, eccetera. Cerchiamo allora di stabilire quanti punti possono avere in comune una curva algebrica e una retta. Il caso di due rette è facile: possono essere coincidenti, avere in comune un punto o nessuno; in questo caso si diranno *parallele*.

Consideriamo ora una retta $C(\alpha x + \beta y + \gamma)$. Esiste un altro modo di scriverne i punti: se (a, b) appartiene alla retta, avremo $\gamma = -\alpha a - \beta b$. Quindi i punti della retta sono tutti e soli quelli della forma

$$(a + \beta t, b - \alpha t)$$

al variare di $t \in F$ (esercizio). Chiameremo $x = a + \lambda t$, $y = b + \mu t$ le *equazioni parametriche* di una retta. Dobbiamo ovviamente avere λ e μ non entrambi nulli. Dalla forma parametrica otteniamo anche quella "polinomiale" come $\mu x - \lambda y + (\lambda b - \mu a)$.

Sia ora data una curva $C(f)$. Un punto in comune tra curva e retta dovrà allora essere della forma $(a + \lambda t, b + \mu t)$ con

$$f(a + \lambda t, b + \mu t) = 0.$$

L'espressione $f(a + \lambda t, b + \mu t)$ è un polinomio in t che ha grado non maggiore del grado totale di f. Ci sono due casi: il primo è quando esso è non nullo, in cui la retta e la curva hanno al massimo n punti in comune, dove n è il grado di f; il secondo è quando $f(a + \lambda t, b + \mu t)$ è il polinomio nullo.

Questo caso succede se e solo se il polinomio $\mu x - \lambda y + (\lambda b - \mu a)$ è un divisore di f. Distinguiamo anche qui due casi. Se $\lambda \neq 0$, possiamo scrivere la retta come $C(y - mx - q)$. Consideriamo poi $A = F[x]$ e in $A[y]$ eseguiamo la divisione di f per $y - (mx + q)$; questo è possibile perché $y - (mx + q)$ è monico. Possiamo allora scrivere

$$f = (y - (mx + q))g + r$$

dove $r \in A$, cioè è un polinomio nella sola x. Il fatto che $f(x, mx + q) = 0$ dice che $r = 0$.

Nel caso in cui $\lambda = 0$, la retta può essere scritta come $x - a$; se $f(a, y) = 0$, abbiamo che $x - c$ è un divisore di f (esercizio).

Quando tutti i punti di una retta fanno parte anche di una curva data, diremo che la retta è una *componente della curva*.

PROPOSIZIONE 3.2.2. *Una curva di grado n e una retta che non sia una sua componente possono avere al massimo n punti in comune.*

Saremo interessati principalmente a curve di grado > 1 che siano definite da un polinomio irriducibile. In questo caso una retta non può essere una componente della curva. Diremo allora che la curva è *F-irriducibile*.

Analizziamo un esempio molto importante, la curva in $\mathbf{R}^2$ definita da $f = x^3 - x^2 + y^2$. Non è difficile dimostrare che questo polinomio è irriducibile (esercizio). Il grafico di questa curva è illustrato nella figura 1.

Se consideriamo una retta passante per un punto della curva diverso dall'origine, questa interseca la curva in vari punti. Prendiamo ad esempio una retta $x = 1 + \lambda t$, $y = \mu t$:

$$f(1 + \lambda t, \mu t) = t(\lambda^3 t^2 + (2\lambda^2 + \mu^2)t + \lambda).$$

Questo polinomio è di grado 2 se $\lambda = 0$ e, in tal caso, ha la radice zero di molteplicità 2. Altrimenti ha grado 3 e non ha radici multiple se non quando $\mu = 0$; in effetti 0 non è radice del fattore di grado 2, il quale ha discriminante $\mu^2(4\lambda^2 + \mu^2)$.

Diverso è il caso delle rette passanti per l'origine, che hanno equazioni $x = \lambda t$, $y = \mu t$. Si ha infatti

$$f(\lambda t, \mu t) = t^2(\lambda^3 t + (\mu^2 - \lambda^2))$$

che ha 0 come radice almeno doppia. È addirittura tripla quando $\mu = \pm\lambda$. Il grado si abbassa quando $\lambda = 0$.

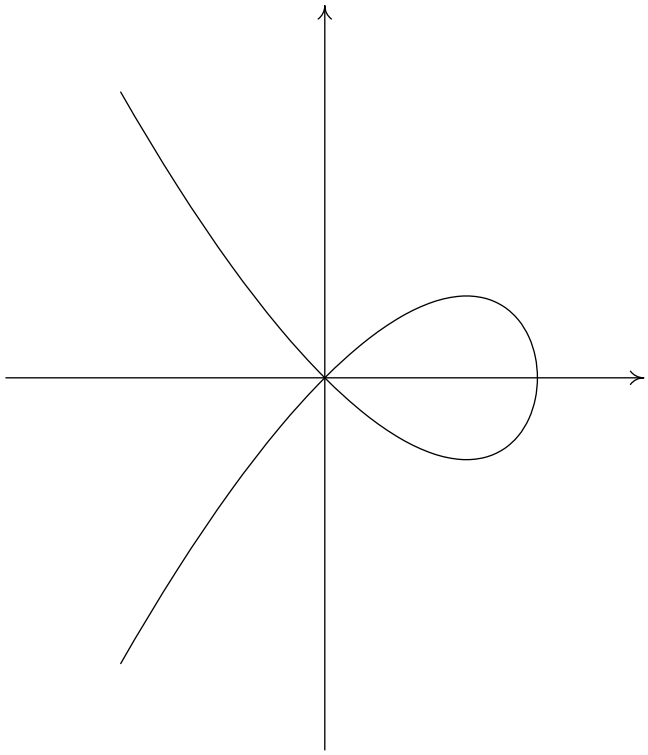

FIGURA 1. Folium di Cartesio

Dovrebbe essere evidente il motivo di questo comportamento: la curva "passa due volte" per l'origine. Le rette $x+y$ e $x-y$ sono "tangenti" ai due rami della curva. Nel punto $(1,0)$ la tangente è la retta verticale, per la quale la radice 0 è multipla.

DEFINIZIONE 3.2.3. Un punto (a,b) di una curva $C(f)$ si dice *multiplo* o *singolare* se, per ogni retta $x=a+\lambda t$, $y=b+\mu t$ passante per (a,b), il polinomio $f(a+\lambda t, b+\mu t)$ ha zero come radice multipla.

Esiste una semplice caratterizzazione dei punti singolari. Dato un polinomio $f \in F[x]$, definiamo la sua derivata tramite l'usuale regola nx^{n-1}: perciò, se $f = a_0 + a_1 x + \cdots + a_n x^n$, la derivata è $f' = a_1 + 2a_2 x + \cdots + n a_n x^n$. Le usuali regole di derivazione sono valide. Se invece $f \in F[x,y]$, possiamo definirne le derivate parziali eseguendo le derivate rispetto a x e a y. Le indicheremo con $D_1 f$ e $D_2 f$ rispettivamente.

TEOREMA 3.2.4. *Un punto (a,b) della curva $C(f)$ è multiplo se e solo se $D_1 f(a,b) = 0$ e $D_2 f(a,b) = 0$.*

DIMOSTRAZIONE. Proviamo a valutare il monomio $g = x^i y^j$ in $(a+\lambda t, b+\mu t)$: otteniamo

$$g(a+\lambda t, b+\mu t) = a^i b^j + i a^{i-1} b^j \lambda t + j a^i b^{j-1} \mu t + (\dots) t^2.$$

Diventa allora evidente che

$$f(a+\lambda t, b+\mu t) = f(a,b) + \big(\lambda D_1 f(a,b) + \mu D_2 f(a,b)\big) t + (\dots) t^2.$$

Se vogliamo che la radice 0 sia doppia per ogni retta, dobbiamo necessariamente avere $D_1 f(a,b) = D_2 f(a,b) = 0$. Ovviamente la condizione è anche sufficiente. □

Possiamo anche definire la molteplicità di un punto: ne daremo solo una versione operativa, che si può giustificare proseguendo nello sviluppo indicato nella dimostrazione precedente. Un punto (a, b) della curva $C(f)$ si dice di *molteplicità* m se tutte le derivate parziali di f fino all'ordine m si annullano in (a, b), ma una di ordine $m + 1$ non si annulla.

La verifica è facile quando il punto da considerare è l'origine $(0, 0)$: essa è un punto m-uplo per $C(f)$ se in f non compaiono monomi di grado totale minore di m.

Ad esempio, la curva di $f = x^5 - y(x^2 - y^2)(x^2 - 4y^2)$ ha un punto quadruplo nell'origine. Il suo grafico nel caso in cui $F = \mathbf{R}$ è indicato nella figura 2. Nella figura 3 è disegnato il grafico della curva di $x^5 - y(x^2 - y^2)^2$.

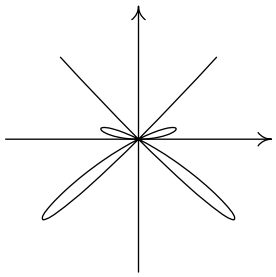

FIGURA 2. Curva con punto quadruplo $x^5 - y(x^2 - y^2)(x^2 - 4y^2)$

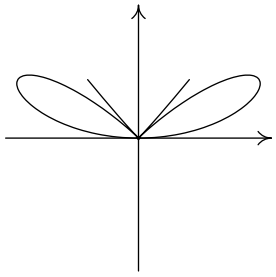

FIGURA 3. Curva con punto quadruplo $x^5 - y(x^2 - y^2)^2$

Se un punto (a, b) della curva $C(f)$ è semplice, esiste sempre una retta $x = a+\lambda t$, $y = b+\mu t$ tale che 0 sia una radice doppia del polinomio $f(a + \lambda t, b + \mu t)$. Infatti basta scrivere che $\lambda D_1 f(a, b) + \mu D_2 f(a, b) = 0$ e una soluzione (ed è l'unica come retta) si trova.

DEFINIZIONE 3.2.5. Una retta di equazioni parametriche $x = a + \lambda t$, $y = b + \mu t$ è *tangente* alla curva $C(f)$ nel punto m-uplo (a, b) se 0 è radice di molteplicità almeno $m + 1$ del polinomio $f(a + \lambda t, b + \mu t)$.

Non sempre è possibile invece determinare le tangenti in un punto multiplo: ad esempio, se $F = GF(3)$, la cubica $x^3 + x^2 + xy + 2y^2$ ha un punto doppio nell'origine; se consideriamo una retta $x = \lambda t$, $y = \mu t$, il polinomio che ne risulta è $\lambda^3 t^3 + (\lambda^2 + \lambda\mu + 2\mu^2)t^2$ e non esistono $\lambda, \mu \in GF(3)$, non entrambi nulli, tali che $\lambda^2 + \lambda\mu + 2\mu^2 = 0$. La situazione ovviamente non si presenta in un campo algebricamente chiuso.

Come ogni polinomio a coefficienti in F si può considerare come a coefficienti in ogni estensione di F, così per le curve.

DEFINIZIONE 3.2.6. Una curva $C(f)$ sul campo F si dice *irriducibile* se non esiste un'estensione algebrica K di F in cui f sia un polinomio irriducibile. Diremo che la curva è *non-singolare* se non esiste un'estensione algebrica K di F in cui la curva $C(f)$ ammette punti multipli.

È chiaro che come estensione si può prendere la chiusura algebrica di F ed eseguire i calcoli in essa. Ad esempio, la curva $x^2 + y^2$ sarebbe irriducibile in $\mathbf{R}$, ma non lo è, visto che nei complessi possiamo scrivere $x^2 + y^2 = (x + iy)(x - iy)$.

Come esempio, determiniamo quali curve $f = x^3 + \alpha x + \beta - y^2$ sono non-singolari. Abbiamo $D_1 f = 3x^2 + \alpha$ e $D_2 f = -2y$. Supponiamo che la caratteristica di F sia diversa da 2 e da 3. Un punto multiplo deve avere allora la forma $(t, 0)$, dove

$$\begin{cases} t^3 + \alpha t + \beta = 0 \\ 3t^2 + \alpha = 0 \end{cases}$$

Affinché il sistema ammetta soluzione in una opportuna estensione algebrica di F è necessario e sufficiente che $4a^3 + 27b^3 = 0$, condizione che vediamo dipendere solo da F.

3.3. Coordinate omogenee

Consideriamo di nuovo la cubica $f = x^3 + \alpha x + \beta - y^2$, che avrà particolare importanza nel seguito. Se la intersechiamo con una retta $x = a + \lambda t$, $y = b + \mu t$, otteniamo

$$f(a + \lambda t, b + \mu t) = \lambda^3 t^3 + (3\lambda^2 a - \mu^2)t^2 + (3\lambda a^2 + \lambda\alpha - 2\mu b)t + a^3 + \alpha a + \beta - b^2.$$

Vediamo allora che le rette con $\lambda \neq 0$ incontrano la curva in tre punti (con le loro molteplicità, ovviamente). Invece quelle con $\lambda = 0$ la incontrano al più due punti. Questa asimmetria è poco piacevole, ma è nella natura delle cose.

Ancora più spiacevole è il caso della cubica $x^2 y - y + 1$, il cui grafico è nella figura 4; ad esempio, le rette parallele all'asse y la incontrano solo in un punto.

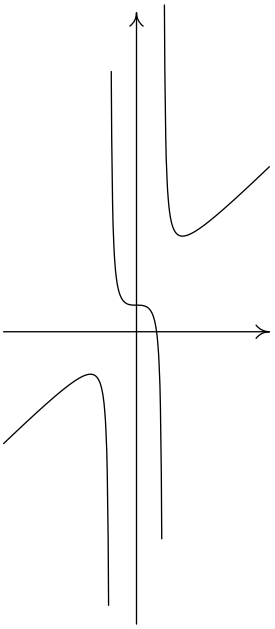

FIGURA 4. Curva $x^3 - (y - 1)x^2 + (y - 1)$

È possibile trovare qualcosa in comune fra questa cubica e il Folium di Cartesio? La risposta è che queste due curve sono *la stessa curva*, da due punti di vista diversi. Per vedere la somiglianza, però, è necessario introdurre il piano proiettivo.

DEFINIZIONE 3.3.1. Il *piano proiettivo* sul campo F è l'insieme $P^2(F)$ così determinato:

(1) nell'insieme $F^3 \setminus \{(0,0,0)\}$ delle terne ordinate di elementi di F esclusa $(0,0,0)$ si definisce la relazione di equivalenza

$$(a,b,c) \sim (a',b',c') \text{ se e solo se esiste } \rho \in F,\ \rho \neq 0 \text{ con } (a',b',c') = (\rho a, \rho b, \rho c).$$

(2) $P^2(F) = (F^3 \setminus \{(0,0,0)\})/\sim$.

La classe di equivalenza di (a,b,c) sarà indicata con $(a:b:c)$, un *punto* di $P^2(F)$.

È chiaro come generalizzare il tutto a "spazi" di dimensione maggiore, ma non lo useremo. Piuttosto cerchiamo di scoprire perché $P^2(F)$ si chiama "piano". Abbiamo punti e vogliamo definire le "rette".

Definizione 3.3.2. Dati $\alpha, \beta, \gamma \in F$ non tutti nulli, la *retta* $\alpha x_1 + \beta x_2 + \gamma x_3$ è l'insieme dei punti $(a:b:c) \in P^2(F)$ tali che

$$\alpha a + \beta b + \gamma c = 0.$$

Notiamo che questa è una buona definizione, perché non dipende dalla particolare terna che rappresenta il punto. Due rette sono uguali se hanno gli stessi punti.

Proposizione 3.3.3. *Due rette $\alpha x_1 + \beta x_2 + \gamma x_3$ e $\alpha' x_1 + \beta' x_2 + \gamma' x_3$ sono uguali se e solo se esiste $\rho \in F$, $\rho \neq 0$, tale che*

$$\alpha' = \rho\alpha, \quad \beta' = \rho\beta, \quad \gamma' = \rho\gamma.$$

Dimostrazione. Il punto $(a:b:c)$ appartiene alla retta $\alpha x_1 + \beta x_2 + \gamma x_3$ se e solo se il vettore $[a\ b\ c]^T$ è soluzione del sistema

$$\begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

Il sottospazio delle soluzioni di questo sistema ha dimensione 2, perché la matrice $[\alpha\ \beta\ \gamma]$ ha rango 1. Dire che le due rette hanno gli stessi punti equivale allora a dire che il sistema

$$\begin{bmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

ha come insieme delle soluzioni un sottospazio di dimensione 2, cioè che la matrice

$$\begin{bmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \end{bmatrix}$$

ha rango 1. Completare la dimostrazione per esercizio. □

Proposizione 3.3.4. *Due rette distinte in $P^2(F)$ hanno esattamente un punto in comune.*

Dimostrazione. Dire che le rette sono distinte significa che, usando le stesse notazioni di prima, la matrice

$$\begin{bmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \end{bmatrix}$$

ha rango 2. Il sistema associato ha allora come spazio delle soluzioni un sottospazio di dimensione 1, che quindi ha una base formata da un vettore non nullo $[a\ b\ c]^T$. Il punto $(a:b:c)$ è allora l'unico punto comune alle due rette. □

Proposizione 3.3.5. *Dati due punti distinti di $P^2(F)$, esiste una ed una sola retta che passa per i due punti.*

DIMOSTRAZIONE. Se i punti sono $(a : b : c)$ e $(a' : b' : c')$, una retta per essi deve essere della forma $\alpha x_1 + \beta x_2 + \gamma x_3$. Dunque $[\alpha\ \beta\ \gamma]^T$ deve essere soluzione del sistema che ha come matrice associata

$$\begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix}$$

che ha, per ipotesi, rango 2. L'insieme delle soluzioni è allora un sottospazio di dimensione 1, ciò che determina una retta. □

Il fatto che parliamo di punti e rette dovrebbe indurci a pensare di trattare oggetti che hanno qualcosa a che fare con gli enti geometrici tradizionali, almeno nel caso in cui $F = \mathbf{R}$.

Se consideriamo i punti $(a : b : c) \in P^2(F)$ tali che $c \neq 0$, abbiamo $(a : b : c) = (a/c : b/c : 1)$, quindi possiamo definire un'applicazione $j\colon F^2 \to P^2(F)$ con l'assegnazione $(a, b) \mapsto (a : b : 1)$ e questa è iniettiva. Decidiamo di chiamare *propri* i punti di $P^2(F)$ per i quali $c \neq 0$ e *impropri* gli altri. Identificheremo i punti propri con i corrispondenti punti di F^2 tramite j.

Supponiamo sia data una retta $\alpha x_1 + \beta x_2 + \gamma x_3$ in $P^2(F)$. I punti propri della retta sono quelli $(a : b : 1)$ per i quali $\alpha a + \beta b + \gamma = 0$, cioè costituiscono una retta di F^2. Su ogni retta $\alpha x_1 + \beta x_2 + \gamma x_3$ in cui uno fra α e β è non nullo esiste uno ed un solo punto improprio.

Sappiamo poi che la condizione affinché due rette del piano affine $\alpha x + \beta y + \gamma$ e $\alpha' x + \beta' y + \gamma'$ siano parallele è che il sistema che ha come matrice completa

$$\left[\begin{array}{cc|c} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \end{array}\right]$$

non abbia soluzioni, cioè, per il teorema di Rouché-Capelli, che la matrice completa abbia rango 2, mentre quella dei coefficienti abbia rango 1. È facile verificare allora che le rette del piano proiettivo $\alpha x_1 + \beta x_2 + \gamma x_3$ e $\alpha' x_1 + \beta' x_2 + \gamma' x_3$ hanno in comune uno ed un solo punto improprio.

Quello che succede è allora la realizzazione dell'intuizione di Desargues: abbiamo aggiunto ad ogni retta un punto improprio, in modo che due rette parallele abbiano in comune esattamente il loro punto improprio. La totalità dei punti impropri forma una retta (del piano proiettivo), che ha come polinomio x_3.

Mettiamoci ora nello spazio (affine) tridimensionale. Sia C il punto di coordinate $(0, 0, 1)$ e sia π il piano di equazione $y = -1$. Associamo ai punti di π un punto del piano xy nel modo seguente:

(1) dato il punto P di coordinate $(\xi, -1, \eta)$ si prende la retta r_P che passa per questo punto e C;
(2) si associa a P il punto $\tilde{P}$ di intersezione della retta r_P con il piano xy.

È facile vedere che le coordinate del punto $\tilde{P}$ sono

$$\left(\frac{\xi}{\eta - 1}, \frac{1}{\eta - 1}, 0\right).$$

Notiamo che non ad ogni punto di π corrisponde un punto del piano xy, infatti i punti della retta $y = -1$, $z = 1$ non hanno un corrispondente. Possiamo anche scrivere la corrispondenza "inversa": al punto Q di coordinate $(x, y, 0)$ associamo il punto $\tilde{Q}$ di coordinate

$$\left(\frac{x}{y}, -1, \frac{1}{y} + 1\right).$$

Non hanno corrispondente solo i punti della retta $y = 0$, $z = 0$. Se escludiamo i punti che non hanno corrispondente, le due applicazioni sono una l'inversa dell'altra, cioè $\tilde{\tilde{P}} = P$ e $\tilde{\tilde{Q}} = Q$.

Abbiamo ottenuto una corrispondenza fra punti di due piani. Un sistema di coordinate sul piano π è dato dalle coordinate "prima" e "terza". Vogliamo vedere quali punti di π sono

mandati in punti del Folium di Cartesio sul piano xy; saranno quelli di coordinate (locali) (ξ, η) per i quali

$$\left(\frac{\xi}{\eta-1}\right)^3 - \left(\frac{\xi}{\eta-1}\right)^2 + \left(\frac{1}{\eta-1}\right)^2 = 0$$

cioè quelli per cui $\xi^3 - (\eta-1)\xi^2 + (\eta-1) = 0$. L'immagine del Folium di Cartesio sul piano π è proprio la cubica che abbiamo descritto nella figura 4.

Proviamo a estendere la corrispondenza a una fra piani proiettivi. Al punto P di coordinate $(a : b : c)$ associamo il punto $\tilde{P}$ di coordinate $(a : c : b - c)$. Abbiamo in altre parole un'applicazione

$$\begin{array}{rccc} \Phi\colon & P^2(F) & \to & P^2(F) \\ & (a : b : c) & \mapsto & (a : c : b - c) \end{array}$$

la cui inversa è l'applicazione

$$\begin{array}{rccc} \Psi\colon & P^2(F) & \to & P^2(F) \\ & (a : b : c) & \mapsto & (a : b + c : b) \end{array}$$

Ora le due applicazioni sono definite ovunque e sono una l'inversa dell'altra. Sui punti propri coincidono con le funzioni definite parzialmente di cui sopra; in effetti un punto della retta propria $y - 1$, cioè un punto della forma $(a : 1 : 1)$ ha come corrispondente tramite Φ il punto $(a : 1 : 0)$ che è un punto improprio. Analogamente un punto $(a : 0 : 1)$ ha come corrispondente tramite Ψ il punto $(a : 1 : 0)$, che è improprio.

Possiamo anche scrivere le due applicazioni per mezzo di matrici:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ c \\ b - c \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ b + c \\ b \end{bmatrix}.$$

Le due matrici sono precisamente una inversa dell'altra.

Se abbiamo una matrice 3×3, chiamiamola A, possiamo associare ad essa una trasformazione del piano proiettivo, purché A sia non-singolare; infatti, in tal caso, $A\mathbf{v} \neq \mathbf{0}$ se $\mathbf{v} \neq 0$. Al vettore $\mathbf{v} = [a\ b\ c]^T$ possiamo associare il punto $\tilde{\mathbf{v}} = (a : b : c)$; questo modo di procedere definisce effettivamente un'applicazione $\tilde{A}\colon P^2(F) \to P^2(F)$. Infatti, se $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$, avremo $\mathbf{u} = \rho\mathbf{v}$ e quindi

$$A\mathbf{u} = A(\rho\mathbf{v}) = \rho(A\mathbf{v}).$$

Perciò

$$\tilde{A}(\tilde{\mathbf{u}}) = \widetilde{A\mathbf{u}} = \widetilde{\rho(A\mathbf{v})} = \widetilde{A\mathbf{v}} = \tilde{A}(\tilde{\mathbf{v}}).$$

L'applicazione $\tilde{A}$ è biiettiva, perché la sua inversa è $\widetilde{A^{-1}}$. Se poi $B = \rho A$, per $\rho \in F$, $\rho \neq 0$, abbiamo che $\tilde{B} = \tilde{A}$.

Definizione 3.3.6. Una proiettività del piano proiettivo su F è un'applicazione del tipo $\tilde{A}$, dove A è una matrice 3×3 non-singolare. L'insieme delle proiettività del piano proiettivo è un gruppo rispetto all'operazione di composizione, denotato con $PGL(2, F)$.

Come esercizio, cerchiamo di studiare meglio il gruppo delle proiettività. Partiamo dal gruppo $GL(3, F)$ delle matrici 3×3 non-singolari. Le matrici scalari

$$S_\rho = \begin{bmatrix} \rho & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{bmatrix}$$

con $\rho \neq 0$ formano un sottogruppo normale H. Sappiamo che $\rho A = S_\rho A$, quindi abbiamo proprio che

$$PGL(2, F) = GL(3, F)/H.$$

Esercizio: determinare l'ordine dei gruppi $PGL(2, F)$ quando F è il campo $GF(2)$, $GF(4)$, $GF(3)$ e $GF(9)$.

Vogliamo studiare anche i *punti uniti* di una proiettività: un punto $\tilde{\mathbf{v}}$ è unito per $\tilde{A}$ quando $\tilde{\mathbf{v}} = \tilde{A}(\tilde{\mathbf{v}})$. La condizione necessaria e sufficiente è che esista uno scalare non nullo ρ tale che

$$A\mathbf{v} = \rho\mathbf{v},$$

cioè che ρ sia un autovalore di A e che $\mathbf{v}$ ne sia un autovettore.

Supponiamo che ρ sia un autovalore di A di molteplicità geometrica 2; quindi esistono due autovettori $\mathbf{u}$ e $\mathbf{v}$ relativi a ρ linearmente indipendenti. La retta per i due punti $\tilde{\mathbf{u}}$ e $\tilde{\mathbf{v}}$ consiste dei punti della forma $\widetilde{\lambda\mathbf{u} + \mu\mathbf{v}}$, dove λ e μ sono scalari non entrambi nulli. I vettori $\lambda\mathbf{u} + \mu\mathbf{v}$ sono allora autovettori relativi a ρ, quindi sono tutti punti uniti.

Il linguaggio dei vettori permette di scrivere facilmente la retta passante per due punti. Infatti una retta r è data dal suo polinomio $\alpha x_1 + \beta x_2 + \gamma x_3$. Sia allora $\mathfrak{a}$ il vettore riga $[\alpha\ \beta\ \gamma]$. Di nuovo il vettore riga determina la retta a meno di multipli scalari non nulli. Il punto $\tilde{\mathbf{v}}$ appartiene a r se e solo se $\mathfrak{a}\mathbf{v} = 0$. Una retta è determinata da due suoi punti distinti: infatti, se $\tilde{\mathbf{u}}$ e $\tilde{\mathbf{v}}$ appartengono alla retta r, ogni vettore della forma $\lambda\mathbf{u} + \mu\mathbf{v}$ (con λ e μ non entrambi nulli) definisce un punto della retta r. Viceversa, dato $\mathbf{w}$ tale che $\tilde{\mathbf{w}} \in r$, abbiamo che $\mathbf{u}, \mathbf{v}, \mathbf{w} \in N(\mathfrak{a})$. Ma lo spazio nullo di $\mathfrak{a}$ ha dimensione 2, quindi i tre vettori sono linearmente dipendenti, cosicché esistono λ, μ, ν non tutti nulli con

$$\lambda\mathbf{u} + \mu\mathbf{v} + \nu\mathbf{w} = 0.$$

Ora necessariamente $\nu \neq 0$, visto che $\mathbf{u}$ e $\mathbf{v}$ sono linearmente indipendenti, quindi possiamo scrivere

$$\mathbf{w} = (-\lambda/\nu)\mathbf{u} + (-\mu/\nu)\mathbf{v}.$$

Dunque la retta per due punti distinti $\tilde{\mathbf{u}}$ e $\tilde{\mathbf{v}}$ è definita da un qualunque vettore riga non nullo $\mathfrak{a}$ che appartiene allo spazio nullo sinistro della matrice $[\mathbf{u}\ \mathbf{v}]$. Inoltre un punto appartiene alla retta se e solo se il vettore che lo definisce è combinazione lineare di due vettori che rappresentano punti distinti della retta.

Sia ora data una proiettività $\tilde{A}$ e siano $\tilde{\mathbf{u}}$ e $\tilde{\mathbf{v}}$ punti distinti. Vogliamo vedere che un punto $\tilde{\mathbf{w}}$ appartiene alla retta r per $\tilde{\mathbf{u}}$ e $\tilde{\mathbf{v}}$ se e solo se $\tilde{A}(\tilde{\mathbf{w}})$ appartiene alla retta per $\tilde{A}(\tilde{\mathbf{u}})$ e $\tilde{A}(\tilde{\mathbf{v}})$.

Se $\tilde{\mathbf{w}}$ appartiene alla retta r per $\tilde{\mathbf{u}}$ e $\tilde{\mathbf{v}}$, allora $\mathbf{w} = \lambda\mathbf{u} + \mu\mathbf{v}$. Allora $A\mathbf{w} = \lambda A\mathbf{u} + \mu A\mathbf{v}$, come volevamo. Il viceversa è analogo, usando A^{-1}.

Ci interessa scoprire qual è il vettore riga che definisce la retta trasformata. Se $\mathfrak{a}$ è il vettore riga della retta r e $\mathfrak{b}$ quello della retta trasformata, dobbiamo avere

$$\mathfrak{b}A\mathbf{u} = 0 = \mathfrak{b}A\mathbf{v}.$$

Dunque $\mathfrak{b}A$ è un vettore riga non nullo nello spazio nullo sinistro della matrice $[\mathbf{u}\ \mathbf{v}]$ e quindi è un multiplo scalare di $\mathfrak{a}$. Ne segue che possiamo anche prendere $\mathfrak{b}A = \mathfrak{a}$, cioè che $\mathfrak{b} = \mathfrak{a}A^{-1}$.

PROPOSIZIONE 3.3.7. *Una proiettività trasforma rette in rette e conserva l'incidenza.*

DIMOSTRAZIONE. La prima parte è già fatta: l'immagine diretta tramite $\tilde{A}$ di una retta è una retta. La seconda parte dice che, se due rette hanno in comune un punto, le loro immagini dirette hanno in comune l'immagine di quello stesso punto.

Infatti se $\tilde{\mathbf{v}}$ appartiene alle rette definite dai vettori riga $\mathfrak{a}$ e $\mathfrak{b}$, abbiamo $\mathfrak{a}\mathbf{v} = 0$ e $\mathfrak{b}\mathbf{v} = 0$. Ma allora

$$\mathfrak{a}A^{-1}A\mathbf{v} = \mathfrak{a}\mathbf{v} = 0$$

e analogamente per $\mathfrak{b}$. □

Vogliamo stabilire se in una proiettività $\tilde{A}$ esista una *retta unita*, cioè una retta r tale che $\tilde{A}(\tilde{\mathbf{v}}) \in r$, per ogni $\tilde{\mathbf{v}} \in r$. Ciò ovviamente significa che $\mathfrak{a}A^{-1} = \rho\mathfrak{a}$, per quanto visto prima, dove il vettore riga $\mathfrak{a}$ determina r. Prendendo le trasposte la cosa diventa chiara:

$$\mathfrak{a}A^{-1} = \rho\mathfrak{a} \iff \mathfrak{a} = \rho\mathfrak{a}A \iff A^T\mathfrak{a}^T = \rho^{-1}\mathfrak{a}^T,$$

cioè $\mathfrak{a}^T$ è un autovettore per A^T. Se il campo F è algebricamente chiuso, allora, ogni proiettività ammette un punto unito e una retta unita.

DEFINIZIONE 3.3.8. Una affinità del piano proiettivo su F è una proiettività $\tilde{A}$ tale che la retta impropria sia unita.

In altre parole $\tilde{A}$ è un'affinità se e solo se l'immagine di tutti i punti impropri $(a : b : 0)$ è ancora un punto improprio. In particolare ciò deve valere per i punti $\tilde{\mathbf{e}_1}$ e $\tilde{\mathbf{e}_2}$. La condizione sulla matrice $A = [a_{ij}]$ è dunque $a_{31} = a_{32} = 0$. Perciò $a_{33} \neq 0$, altrimenti A sarebbe singolare. Possiamo perciò prendere $a_{33} = 1$. Il punto di coordinate $(a : b : 1)$ viene allora mandato in quello definito dal vettore

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11}a + a_{12}b + a_{13} \\ a_{21}a + a_{22}b + a_{23} \\ 1 \end{bmatrix}$$

e quindi un'affinità è ciò a cui siamo abituati nel piano affine.

3.4. Polinomi omogenei

Come le rette affini possono essere viste come rette del piano proiettivo private del punto improprio, anche le curve algebriche affini possono essere viste come curve del piano proiettivo private di punti impropri. Ma quali?

Ad esempio, la cubica della figura 4 ha tre asintoti: due verticali, di equazioni $x - 1$ e $x + 1$ rispettivamente, e uno obliquo, di equazione $x - y + 1$. Sembra ragionevole dire che la cubica passa per i punti impropri di queste rette, che hanno coordinate $(0 : 1 : 0)$ quello delle prime due e $(1 : 1 : 0)$ per la terza. Di più: il fatto che queste rette siano asintoti dice intuitivamente che esse sono tangenti alla cubica nei punti impropri. Per il punto $(0 : 1 : 0)$ passano due tangenti, infatti questo punto è l'immagine del punto doppio del Folium di Cartesio. La retta impropria allora ha tre punti in comune con la cubica: due coincidenti (il punto doppio) e uno semplice.

Prendiamo il Folium di Cartesio: le rette che lo incontrano in due soli punti sono le rette parallele all'asse delle ordinate; sembra ragionevole assumere che il punto improprio da aggiungere sia allora $(0 : 1 : 0)$, che è il punto improprio di queste rette. Infatti le rette di quel tipo hanno la forma $x + \gamma z$.

Come si passa da una retta nel piano affine a quella corrispondente nel piano proiettivo? Se la prima è definita dal polinomio $\alpha x + \beta y + \gamma$, la seconda è definita da $\alpha x + \beta y + \gamma z$. Abbiamo semplicemente reso *omogeneo* il polinomio. Cerchiamo di dare una definizione più precisa.

DEFINIZIONE 3.4.1. Un polinomio $f \in F[x_1, \ldots, x_n]$ è *omogeneo* se tutti i monomi che compaiono con coefficiente diverso da zero in f hanno lo stesso grado totale. Per convenzione il polinomio nullo è considerato omogeneo di qualunque grado (questo serve per rendere l'insieme dei polinomi omogenei di grado n uno spazio vettoriale su F).

Indichiamo con D_i la derivata parziale rispetto alla i-esima indeterminata. Eulero dimostrò un importante fatto sui polinomi omogenei.

TEOREMA 3.4.2. *Sia $f \in F[x_1, \ldots, x_n]$ un polinomio omogeneo di grado k. Allora, come polinomi in $F[x_1, \ldots, x_n][t]$ si ha*

$$f(tx_1, \ldots, tx_n) = t^k f(x_1, \ldots, x_n);$$

inoltre $kf = \sum_{i=1}^{n} x_i D_i f$.

DIMOSTRAZIONE. La prima proprietà si verifica facilmente sui monomi e quella generale è una conseguenza immediata.

Non è troppo complicato verificare che la derivata rispetto a t del primo membro è

$$\sum_{i=1}^{n} x_i D_i f(tx_1, \ldots, tx_n)$$

(è come il *differenziale totale*). La derivata del secondo membro è invece $kt^{k-1}f(x_1, \ldots, x_n)$. Sostituendo allora 1 al posto di t, abbiamo la tesi. □

La prima condizione che abbiamo visto è anche sufficiente affinché un polinomio sia omogeneo. La dimostrazione è lasciata per esercizio. Ci limiteremo da ora in poi a polinomi in tre indeterminate, anche se le considerazioni si possono estendere facilmente a dimensione maggiore.

Dato un polinomio omogeneo $f \in F[x_1, x_2, x_3]$ di grado k, possiamo considerare gli elementi $(a_1, a_2, a_3) \in F^3$ tali che

$$f(a_1, a_2, a_3) = 0.$$

L'elemento $(0, 0, 0)$ soddisfa sempre questa uguaglianza; inoltre, se (a_1, a_2, a_3) la soddisfa e $\rho \in F$, abbiamo certo anche

$$f(\rho a_1, \rho a_2, \rho a_3) = \rho^k f(a_1, a_2, a_3) = 0.$$

Quindi è ben definito l'insieme dei punti del piano proiettivo

$$C_p(f) = \{\, (a_1 : a_2 : a_3) \in P^2(F) \mid f(a_1, a_2, a_3) = 0 \,\}$$

che chiameremo ancora la *curva* definita da f. Siccome i punti impropri sono per convenzione quelli in cui $a_3 = 0$, quelli propri sono i punti in cui si può prendere $a_3 = 1$. Ma allora i punti propri della curva sono quelli della forma $(a : b : 1)$ per i quali $f(a, b, 1) = 0$. Otteniamo dunque una curva algebrica affine eseguendo la sostituzione di x_1 con x, di x_2 con y e di x_3 con 1.

Viceversa, dato un polinomio $g \in F[x, y]$, non necessariamente omogeneo, ma non nullo, identifichiamo il suo grado totale k. Quindi, trasformiamo ogni monomio che vi compare, diciamo $x^i y^j$, in

$$x_1^i x_2^j x_3^{k-i-j}.$$

La cosa è possibile perché, per ipotesi, $i + j \leq k$. Otteniamo allora un polinomio omogeneo $f \in F[x_1, x_2, x_3]$ di grado k, che ha la proprietà che

$$f(x, y, 1) = g(x, y).$$

Inoltre un punto (a, b) del piano affine appartiene alla curva $C(g)$ se e solo se il punto $(a : b : 1)$ del piano proiettivo appartiene alla curva $C_p(f)$. In effetti abbiamo costruito un algoritmo per passare da polinomi in due indeterminate a polinomi omogenei in tre e viceversa. Il polinomio ottenuto da g si indicherà con $\tilde{g}$; abbiamo allora

$$g(x, y) = \tilde{g}(x, y, 1);$$

se poi $f \in F[x_1, x_2, x_3]$ e $g = f(x, y, 1)$, abbiamo

$$f = \tilde{g}.$$

Dunque la corrispondenza ottenuta è biunivoca. Si dimostri che $\widetilde{g_1 g_2} = \tilde{g}_1 \tilde{g}_2$.

LEMMA 3.4.3. *Se $f_1, f_2 \in F[x_1, x_2, x_3]$ e $f_1 f_2$ è omogeneo, allora f_1 e f_2 sono omogenei.*

DIMOSTRAZIONE. Sia k il grado di $f_1 f_2$. Possiamo scrivere $f_1 = f_{11} + f_{12} + \cdots + f_{1r}$ e $f_2 = f_{21} + f_{22} + \cdots + f_{2s}$ dove ciascun addendo f_{ij} è omogeneo di grado k_{ij}, con $k_{i1} < k_{i2} < \cdots$. Allora abbiamo $f_{11} f_{21}$ omogeneo di grado $k_{11} + k_{21} = k$; supponiamo $r > 1$. Allora $f_{12} f_{21}$ è omogeneo di grado $k_{12} + k_{21} = k$, ciò che implica $k_{12} = k_{11}$, assurdo. Analogamente non può essere $s > 1$. □

PROPOSIZIONE 3.4.4. *Dato $g \in F[x, y]$, g è irriducibile se e solo se $\tilde{g}$ è irriducibile.*

DIMOSTRAZIONE. ($\Rightarrow$) Supponiamo che $\tilde{g} = f_1 f_2$, con f_1 e f_2 omogenei. Allora $g(x,y) = \tilde{g}(x,y,1) = f_1(x,y,1)f_2(x,y,1)$, quindi possiamo supporre che $f_1(x,y,1)$ sia costante. Ciò significa che nei monomi di f_1 non compaiono x_1 e x_2, cioè che $f_1 = ax_3^k$. Dunque $\tilde{g} = ax_3^k f_2$, che va contro il modo di costruire $\tilde{g}$, a meno che $k = 1$. Perciò f_1 è costante.

($\Leftarrow$) Supponiamo $g = g_1 g_2$; allora $\tilde{g} = \tilde{g_1}\tilde{g_2}$ e quindi possiamo supporre che $\tilde{g_1}$ sia costante; ma allora g_1 è costante. □

Vogliamo ora definire che cosa intendiamo per molteplicità di intersezione di una curva nel piano proiettivo con una retta. Per semplicità parleremo solo di curve irriducibili, cioè definite da un polinomio omogeneo irriducibile di grado > 1.

Una retta è definita da due punti $(a_1 : a_2 : a_3)$ e $(b_1 : b_2 : b_3)$ tali che i vettori $u = [a_1\ a_2\ a_3]^T$ e $v = [b_1\ b_2\ b_3]^T$ siano linearmente indipendenti. Un punto di questa retta è della forma $\lambda u + \mu v$, con λ e μ non entrambi nulli. Possiamo allora sostituire in f, polinomio omogeneo che definisce la curva, x_i con $\lambda a_i + \mu b_i$ $(i = 1,2,3)$. Otteniamo allora un polinomio omogeneo in λ e μ di grado k uguale al grado di f:

$$c_0\lambda^k + c_1\lambda^{k-1}\mu + c_2\lambda^{k-2}\mu^2 + \cdots + c_{k-1}\lambda\mu^{k-1} + c_k\mu^k.$$

In generale, se $f \in F[x_1, x_2]$ è omogeneo, possiamo trovarne i fattori di grado 1, che saranno della forma

$$a_1x_1 + a_2x_2.$$

Due di questi fattori sono associati (cioè non sono fattori essenzialmente diversi) quando esiste un $\rho \neq 0$ tale che la moltiplicazione per ρ trasformi uno nell'altro. Una *radice* di questo polinomio è allora un'espressione della forma $(\lambda : \mu)$, dove $a_1\lambda + a_2\mu = 0$. Dovrebbe essere allora chiaro che cosa si intende per molteplicità di una radice. Una radice $(\lambda : \mu)$ ha molteplicità r se

$$(\mu x_1 - \lambda x_2)^r$$

divide il polinomio dato e la potenza successiva no.

TEOREMA 3.4.5. *Sia $f \in F[x_1, x_2, x_3]$ omogeneo di grado $k > 1$ e irriducibile. Se F è algebricamente chiuso, ogni retta proiettiva incontra la curva $C_p(f)$ in k punti, contati con la loro molteplicità.*

Il numero di intersezioni quando il campo non è algebricamente chiuso potrebbe essere minore. Siccome però nel seguito ci interesseranno le curve di grado 3 e rette che sappiamo incontrarle in due punti, il terzo ci deve essere per forza.

Un punto di una curva si dice un *flesso* se la tangente in quel punto ha molteplicità di intersezione almeno 3. Questo dice, ad esempio, che una conica non ha flessi. Invece una cubica ne ha sempre.

Se $f \in F[x_1, x_2, x_3]$ è omogeneo, il suo *Hessiano* è

$$H(f) = \det \begin{bmatrix} D_1D_1f & D_1D_2f & D_1D_3f \\ D_2D_1f & D_2D_2f & D_2D_3f \\ D_3D_1f & D_3D_2f & D_3D_3f \end{bmatrix}.$$

Questo è un polinomio omogeneo di grado $3(k-2)$, se k è il grado di f. Si dimostra che i flessi della curva definita da f sono fra i punti di intersezione della curva definita da f con la curva definita da $H(f)$ (fra questi ci sono anche i punti singolari). Se $k = 3$ il grado dell'Hessiano è ancora 3 (almeno quando la caratteristica non è 2).

Come esempio, prendiamo $f = x_1^3 - x_1^2x_3 + x_2^2x_3$, cioè il Folium di Cartesio in $\mathbf{C}$. Il suo Hessiano è

$$H(f) = \det \begin{bmatrix} 6x_1 - 2x_3 & 0 & -2x_1 \\ 0 & 2x_3 & 2x_2 \\ -2x_1 & 2x_2 & 0 \end{bmatrix} = 8(x_2^2x_3 - 3x_1x_2^2 - x_1^2x_3).$$

Si può verificare che i punti di intersezione sono $(0:1:0)$ e la coppia di punti complessi coniugati

$$\left(\frac{4}{3} : \pm\frac{4i}{3\sqrt{3}} : 1\right).$$