Appunti per Geometria e Algebra Computazionale

Corso di Laurea in Matematica, Università di Firenze, 2019/20

Giorgio Ottaviani

29 maggio 2020

Indice

1 Generalità sull'anello dei polinomi

Sia K un campo. Siamo interessati all'anello dei polinomi $K[x_1, \ldots, x_n]$. Come esempi consideriamo $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ (quest'ultimo campo con p primo, $p \gg 0$ è molto utilizzato in computer algebra e può simulare con maggiore efficienza un campo di caratteristica zero quando i coefficienti dei polinomi in gioco sono "piccoli"). Useremo la seguente proprietà, nota dai corsi di Algebra:

Teorema 1.1. *(Gauss) $K[x_1, \ldots, x_n]$ è un dominio a fattorizzazione unica (UFD) cioè ogni polinomio si decompone in modo unico come prodotto di fattori irriducibili.*

Ricordiamo che un anello A si dice noetheriano [1] se ogni suo ideale è finitamente generato. Questo equivale alla condizione della catena ascendente, cioè ogni catena ascendente di ideali $I_1 \subset I_2 \subset I_3 \subset \ldots$ è stazionaria nel senso che $\exists\ n$ tale che $I_n = I_{n+1} = I_{n+2} = \ldots$

In particolare ogni campo K è noetheriano perché gli unici suoi ideali sono 0 e K.

Teorema 1.2 (Teorema della base di Hilbert.(Basissatz))**.** *Sia R un anello.*

$$R \text{ è noetheriano} \Longrightarrow R[x] \text{ è noetheriano}$$

Dimostrazione. [*] (H. Sarges, 1976) Sia R un anello noetheriano e consideriamo (per assurdo) un ideale $I \subset R[x]$ non finitamente generato. Scegliamo $f_1 \in I$ di grado minimo. Scegliamo poi $f_2 \in I \setminus (f_1)$ ancora di grado minimo e procedendo in questo

[1] in ricordo di Emmy Noether (1882-1935)

[*] le dimostrazioni con asterisco possono essere prese dai corsi di Algebra

modo troviamo $f_k \in I \setminus (f_1, \ldots, f_{k-1})$ di grado minimo. Sia $n_k := deg\ f_k$ e sia $f_k = a_k x^{n_k} + \ldots$ Abbiamo ovviamente $n_1 \leq n_2 \leq \ldots$ e $(a_1) \subset (a_1, a_2) \subset \ldots$. Per ipotesi $\exists \quad p$ tale che $(a_1, \ldots, a_p) = (a_1, \ldots, a_{p+1})$ e quindi si può scrivere $a_{p+1} = \sum_{i=1}^{p} b_i a_i$ con $b_i \in R$. Poniamo $g := f_{p+1} - \sum_{i=1}^{p} x^{n_{p+1}-n_i} b_i f_i$. Quindi il termine di grado massimo di g è

$$a_{p+1} x^{n_{p+1}} - \sum_{i=1}^{p} b_i a_i x^{n_{p+1}} = 0$$

da cui $deg\ g < n_{p+1}$. D'altronde $g \in I$ e $g \notin (f_1, \ldots, f_p)$ (altrimenti $f_{p+1} \in (f_1, \ldots f_p)$). Questa è una contraddizione perché f_{p+1} era stato scelto come un polinomio di grado minimo in $I \setminus (f_1, \ldots, f_p)$. □

Corollario 1.3. *Sia R un anello.*

$$R \text{ è noetheriano} \Longrightarrow R[x_1, \ldots, x_n] \text{ è noetheriano}$$

Dimostrazione. Per induzione su n considerando che

$$R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n]$$

□

Corollario 1.4. *Sia K un campo. Allora $K[x_1, \ldots, x_n]$ è noetheriano.*

La seguente proposizione sarà utile in seguito.

Proposizione 1.5. *Se un ideale I di un anello noetheriano R è generato da $(f_i)_{i \in I}$ allora si può estrarre dagli f_i un numero finito di generatori $f_{i_1}, \ldots, f_{i_n}$*

Dimostrazione. * Scegliamo $f_{i_1} \in I$. Se $(f_{i_1}) \subset_{\neq} I$ allora scegliamo $f_{i_2} \in I$ tale che $(f_{i_1}) \subset_{\neq} (f_{i_1}, f_{i_2})$. Se $(f_{i_1}, f_{i_2}) \subset_{\neq} I$ allora scegliamo $f_{i_3} \in I$ tale che $(f_{i_1}, f_{i_2}) \subset_{\neq} (f_{i_1}, f_{i_2}, f_{i_3})$. Così procedendo si trova ad un certo punto un numero finito di generatori oppure si costruisce una catena ascendente di ideali non stazionaria contro l'ipotesi. □

Per $n = 1$ l'anello dei polinomi in una variabile gode di un'altra proprietà importante: è un dominio a ideali principali (PID). Infatti per ogni ideale I di $K[x]$ esiste $f \in K[x]$ tale che $I = (f)$. Quest'ultima proprietà permette di risolvere facilmente il seguente

Problema di appartenenza in $K[x]$. Dato un ideale I in $K[x]$, esiste un algoritmo per decidere se $g \in I$?

Infatti basta effettuare la divisione di g per il generatore f dell'ideale I. Abbiamo $g = qf + r$ con $deg\ r < deg\ f$. r si dice il resto. Segue che $g \in I$ se e solo se il resto della divisione di g per f è zero.

È naturale il seguente problema analogo

Problema di appartenenza in $K[x_1, \ldots, x_n]$. Dato un ideale I in $K[x_1, \ldots, x_n]$, esiste un algoritmo per decidere se $g \in I$?

Questo secondo problema è complicato dal fatto che l'ideale I non è necessariamente principale e quindi occorre eseguire una divisione per tutti i suoi generatori $f_1, \ldots, f_k$.

Vorremmo trovare un'espressione $g = q_1 f_1 + \ldots + q_k f_k + r$ e concludere che $g \in I$ se e solo se $r = 0$. Per fare questo occorre generalizzare l'algoritmo di divisione al caso di più polinomi. Prima di fare questo è allora opportuno approfondire su quali concetti si basa effettivamente il ben noto algoritmo di divisione per polinomi in una variabile.

L'algoritmo di divisione per polinomi in una variabile.

Dato un polinomio $f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0$ definiamo $LT(f) := a_n x^n$ (leading term). Ovviamente $deg\ (f) = deg(LT(f))$. Per effettuare la divisione di g per f controlliamo se $deg\ (LT(g)) < deg\ (LT(f))$, in tal caso scriveremo $LT(g) < LT(f)$. In caso affermativo il quoziente è zero ed il resto è uguale a g ($g = 0 \cdot f + g$). In caso negativo sommiamo $\frac{LT(g)}{LT(f)}$ al quoziente (che è inizialmente nullo), sostituiamo $g - \frac{LT(g)}{LT(f)} f$ al posto di g e continuiamo come sopra. Questo ciclo ha termine perché ad ogni passo $deg\ \left(g - \frac{LT(g)}{LT(f)} f\right) < deg\ g$ e quindi troviamo una successione strettamente decrescente di gradi che ad un certo punto diventano minori di $deg\ f$.

Il risultato può essere riassunto così:

Teorema 1.6. *Dati f, $g \in K[x]$, $g \neq 0$, esistono (unici) q, $r \in K[x]$ tali che $f = gq + r$ e deg $r <$ deg f oppure $r = 0$. In piú esiste un algoritmo che determina q, r.*

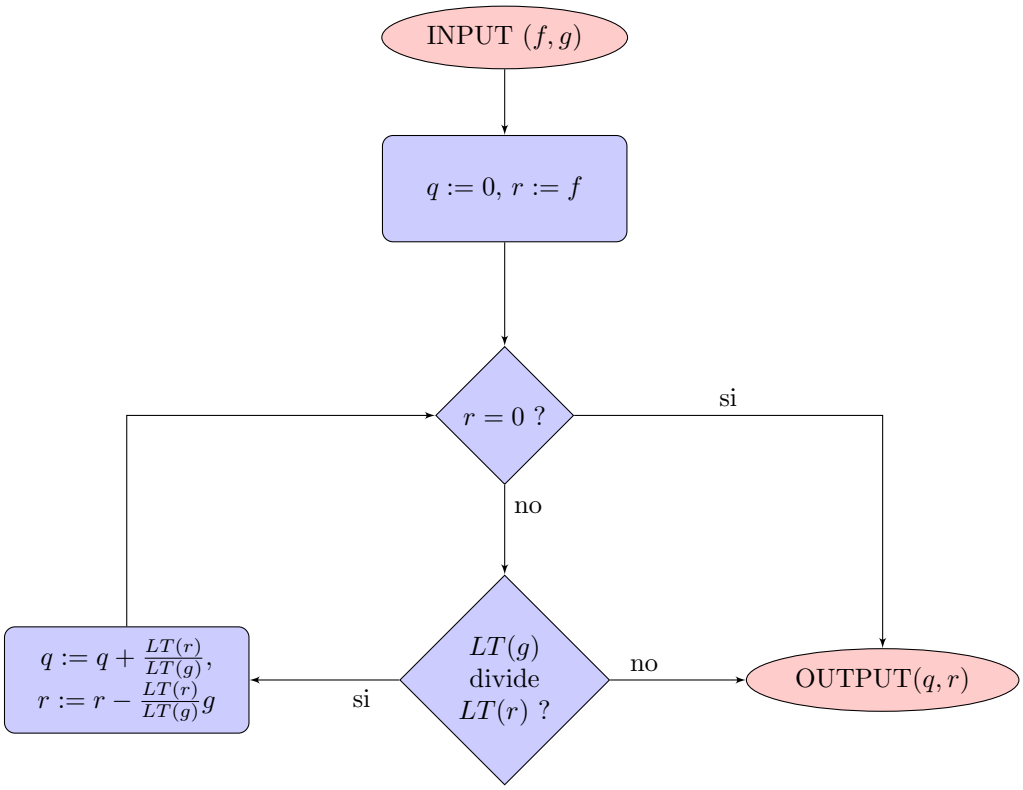
Il lettore è invitato a verificare questo ben noto algoritmo nel caso $f = x^4 + x + 1$, $g = 2x^2 + x + 1$. Si trova $q = \frac{1}{2}x^2 - \frac{1}{4}x - \frac{1}{8}$, $r = \frac{11}{8}x + \frac{9}{8}$. L'algoritmo può essere sintetizzato nel modo seguente:

Diagramma di flusso dell'algoritmo di divisione per polinomi in una variabile.

Dati $f, g \in K[x]$, si costruiscono (unici) $q, r \in K[x]$ tali che

$$f = gq + r,$$

dove $r = 0$ oppure $\deg r < \deg g$.



Dimostrazione L'equazione $f = gq+r$ vale con le prime definizioni di q, r, e continua a valere ad ogni passo dell'algoritmo. Si arriva all'output quando $r = 0$ oppure quando $LT(r) < LT(g)$. L'algoritmo termina perché la successione dei $LT(r)$ è decrescente ad ogni ciclo. Per provare l'unicità, consideriamo due scritture $f = gq + r = gq' + r'$, da cui $r - r' = g(q' - q)$ e se fosse $r - r' \neq 0$ avremmo $\deg(r - r') \geq \deg g$ che è una contraddizione, pertanto $r = r'$, da cui $q = q'$.

Osservazione Il fatto che ogni ideale di $K[x]$ è principale è una conseguenza dell' algoritmo di divisione. L'algoritmo euclideo per calcolare il MCD di due polinomi è basato sull' algoritmo di divisione e permette di calcolare il generatore di un ideale se è noto un insieme di generatori (necessariamente in numero finito per la noetherianità).

Esercizio 1.7.

- *i) Stabilire se* $x^4 + x^3 + x^2 + x + 1 \in \big(x^2 + x + 1\big)$.
- *ii) Stabilire se* $x^3 + 4x^2 + 3x - 6 \in \big(x^3 - 3x + 2, x^4 - 1, x^6 - 1\big)$.

L'algoritmo di divisione per polinomi in più variabili. Ideali monomiali. Ordini monomiali.

Uno dei fatti essenziali che assicura il successo dell'algoritmo di divisione per polinomi in una variabile è che dati due leading term uno dei due è sempre divisibile per l'altro. In altre parole la relazione "x^n divide x^m" è una relazione di ordine totale sui monomi in una variabile, che si identifica con la relazione di ordine usuale sull'insieme dei numeri naturali.

Inoltre una successione di monomi che è strettamente decrescente sui gradi termina dopo un numero finito di passi. Infatti l'ordine usuale sui numeri naturali è un buon ordinamento, cioè ogni sottoinsieme non vuoto ha un minimo.

Tra i monomi in più variabili la relazione di divisibilità definisce soltanto un ordine parziale. Ad esempio x non è divisibile per y e viceversa. Vogliamo definire alcuni ordini totali "ragionevoli" tra i monomi in più variabili. I monomi appartenenti all'anello $K[x_1, \ldots, x_n]$ sono in corrispondenza biunivoca con gli elementi di $\mathbf{Z}^n_{\geq 0}$, infatti possiamo identificare il monomio $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ con la n-pla $(\alpha_1, \ldots \alpha_n)$ dove α_i sono interi non negativi. Osserviamo che con queste notazioni $x^\alpha \cdot x^\beta = x^{\alpha+\beta}$. Poniamo $|\alpha| = \sum_i \alpha_i$.

Definizione 1.8. *Un ideale si dice monomiale se può essere generato da monomi.*

Come corollario della prop. 1.5 abbiamo che ogni ideale monomiale è generato da un numero finito di monomi (lemma di Dickson).

Con la notazione $< f_i, i \in I >$ intendiamo l'ideale generato dagli elementi f_i.

Lemma 1.9. *Sia $I =< x^\alpha, \alpha \in A >$ un ideale monomiale. Abbiamo che*

$$x^\beta \in I \Longleftrightarrow x^\alpha | x^\beta \text{ per qualche } \alpha \in A$$

Dimostrazione. $\Leftarrow$ è ovvia.

Per provare $\Rightarrow$ scriviamo $x^\beta = \sum h_i x^{\alpha(i)}$. A secondo membro ogni termine è divisibile per qualche $x^{\alpha(i)}$, pertanto tale proprietà deve rimanere vera anche a primo membro (dopo ever effettuato tutte le cancellazioni). □

È utile identificare l'insieme dei monomi multipli di x^α come un "ottante" n-dimensionale in $\mathbf{Z}^n_{\geq 0}$ con vertice α. Questa corrispondenza permette di visualizzare un ideale monomiale e allo stesso tempo di utilizzare tecniche di geometria discreta.

Definizione 1.10. *Un ordine monomiale in $K[x_1, \ldots, x_n]$ è una relazione $<$ su $\mathbf{Z}^n_{\geq 0}$ tale che:*

- *i) $<$ è un ordine totale*[2]
- *ii) $<$ è compatibile con la moltiplicazione, cioè $\forall \alpha, \beta, \gamma$ con $\alpha < \beta$ vale $\alpha + \gamma < \beta + \gamma$*
- *iii) $1 < x^\alpha \quad \forall \alpha \in \mathbf{Z}^n_{\geq 0}, \alpha \neq 0$.*

Scriveremo indifferentemente $\alpha > \beta$ oppure $x^\alpha > x^\beta$. Scriviamo $x^\alpha \geq x^\beta$ per indicare $x^\alpha > x^\beta$ oppure $x^\alpha = x^\beta$.

Lemma 1.11. *Se x^α divide x^β allora in un qualunque ordine monomiale $x^\alpha \leq x^\beta$. Pertanto ogni ordine monomiale è un raffinamento dell'ordine parziale definito dalla divisibilità.*

Dimostrazione. Se $\alpha = \beta$ non c'è niente da dimostrare. Per ipotesi $\beta = \alpha + \gamma$ con $\gamma \in \mathbf{Z}^n_{\geq 0}$, $\gamma \neq 0$. Dalla proprietà iii) della Definizione 1.10 abbiamo $1 < x^\gamma$ e quindi dalla proprietà ii) segue $x^\alpha < x^{\alpha+\gamma} = x^\beta$. □

Osservazione. Dal Lemma 1.9 segue che in $K[x]$ esiste un solo ordine monomiale che é quello usuale.

[2] cioè la relazione $<$ soddisfa la proprietà transitiva e $\forall \alpha, \beta \in \mathbf{Z}^n_{\geq 0}$ è vera esattamente una tra le formule $a < b$, $b < a$, $a = b$ ("tricotomia").

Proposizione 1.12. *Ogni ordine monomiale è un buon ordinamento, cioè ogni sottoinsieme non vuoto di monomi ammette un minimo.*

Dimostrazione. Sia $A \subset \mathbb{Z}^n_{\geq 0}$ un sottoinsieme non vuoto di monomi. Sia $I = \langle x^\alpha | \alpha \in A \rangle$ l'ideale generato dai monomi in A. Per la noetherianità dell'anello dei polinomi, esistono $\alpha(1) < \ldots < \alpha(s) \in A$ tali che I è generato da $x^{\alpha(1)}, \ldots, x^{\alpha(s)}$. Affermiamo che $x^{\alpha(1)}$ è un minimo. Infatti, per ogni $\alpha' \in A$ abbiamo $x^{\alpha'} \in I$ e quindi per il Lemma 1.9 esiste i con $1 \leq i \leq s$ tale che $x^{\alpha(i)} | x^{\alpha'}$, da cui $\alpha(1) < \alpha(i) < \alpha'$. □

È bene sapere subito che alcuni ordini monomiali sono preferibili ad altri dal punto di vista computazionale, secondo le applicazioni a cui si è interessati.

Ci sono tre ordini monomiali particolarmente importanti:

1. Lex ordine lessicografico: si definisce $\alpha >_{Lex} \beta$ se in $\alpha - \beta$ il primo coefficiente non nullo da sinistra è positivo. In particolare $x_1 > x_2 > \ldots > x_n$ ed un monomio di grado 10 in x_1 è maggiore di tutti i monomi di grado 9 in x_1 e minore di tutti i monomi di grado 11 in x_1. Se il grado in x_1 è uguale si guarda il grado in x_2 e così via.
2. GLex ordine lessicografico graduato: si definisce $\alpha >_{GLex} \beta$ se $\sum \alpha_i > \sum \beta_i$ oppure se $\sum \alpha_i = \sum \beta_i$ e in $\alpha - \beta$ il primo coefficiente non nullo da sinistra è positivo.
3. GRevLex ordine lessicografico inverso graduato: si definisce $\alpha >_{GRevLex} \beta$ se $\sum \alpha_i > \sum \beta_i$ oppure se $\sum \alpha_i = \sum \beta_i$ e in $\alpha - \beta$ il primo coefficiente non nullo da destra è negativo.

Vedremo che Lex è utile per eliminare variabili, (vedi il Cap.4) mentre GRevLex è ottimale per applicazioni più avanzate, come il calcolo delle sizigie.

Esercizio 1.13. *Provare che fino a due variabili, GLex e GRevLex coincidono, mentre in tre o più variabili sono ordini diversi.*

In generale, dato $\omega \in \mathbb{R}^n_{\geq 0}$ vettore di pesi, con coefficienti nonnegativi, si definisce $\alpha >_\omega \beta$ se $\alpha \cdot \omega > \beta \cdot \omega$ oppure se $\alpha \cdot \omega = \beta \cdot \omega$ e $\alpha >_{Lex} \beta$.

Un ordine monomiale si dice graduato se $x^\alpha > x^\beta$ quando $|\alpha| > |\beta|$. GLex e GRevLex sono graduati, mentre Lex non lo è.

Esempi.

$$x^2y >_{Lex} xy^2 \qquad x^2y >_{GRevLex} xy^2$$

$$x^2yz^2 >_{Lex} xy^3z \qquad x^2yz^2 <_{GRevLex} xy^3z$$

Definizione 1.14. *Sia fissato un ordine monomiale e sia $f \in K[x_1, \ldots, x_n]$. Il multigrado di f è la n-pla massima tra quelle corrispondenti ai termini di f con l'ordine prescelto e si indica con* MULTIDEG(f). *Con questa notazione* MULTIDEG(x^α) $= \alpha$. *$LT(f)$ é il termine di f corrispondente alla n-pla massima.*

Vediamo come possiamo impostare la divisione di f per $f_1, \ldots f_h$. Vogliamo scrivere $f = a_1f_1 + \ldots + a_hf_h + r$. Analogamente al caso dei polinomi in una variabile, chiediamo che $LT(r)$ non sia diviso da nessun $LT(f_i)$, vedremo però che questo è troppo poco.

Procediamo nel modo seguente:

– Poni $p := f$ (resto ausiliario)

– Dividi $LT(p)$ successivamente per $LT(f_1), \ldots, LT(f_h)$ e quando questo è possibile aggiungi $LT(p)/LT(f_i)$ all'i-esimo quoziente e sottrai $f_i \frac{LT(p)}{LT(f_i)}$ dal resto ausiliario p. Quando $LT(p)$ non è più divisibile per nessuno tra $LT(f_1), \ldots, LT(f_h)$ allora aggiungi $LT(p)$ al resto e continua con $p - LT(p)$ al posto di p.

L'algoritmo ha termine quando il resto ausiliario diventa nullo. Questo accade sempre perché ad ogni passo il multigrado del resto ausiliario p decresce strettamente e l'ordine monomiale scelto è un buon ordinamento.

Esempio 1.15. *Dividiamo in $K[x,y]$ con GLex $x^6y^3+2x^3y^2-y+1$ per xy^2-x e y^3-x. Il LT del dividendo è x^6y^3 che è divisibile per xy^2. Dividendo otteniamo x^5y come primo quoziente ed il resto ausiliario è $x^6y+2x^3y^2-y+1$ che ha x^6y come LT. x^6y non è divisibile per nessuno tra i LT dei divisori, pertanto si aggiunge x^6y al resto e si continua a dividere partendo da $2x^3y^2-y+1$. Continuando otteniamo il seguente schema:*

$$\begin{array}{l|l|l|l} x^6y^3+2x^3y^2-y+1 & xy^2-x & y^3-x & RESTO \\ & \overline{\qquad} & \overline{\qquad} & \\ & & & \end{array}$$

Eseguiamo la divisione, scrivendo i resti ausiliari e le operazioni svolte sotto il dividendo:

$$\begin{array}{lllll} x^6y^3+2x^3y^2-y+1 & |xy^2-x & |y^3-x & |RESTO \\ & |\overline{\qquad} & |\overline{\qquad} & \\ x^6y^3-x^6y & |x^5y & | & \\ \overline{\qquad\qquad\qquad\qquad} & & & \\ x^6y+2x^3y^2-y+1 & & & \\ 2x^3y^2-y+1 & \longrightarrow & & x^6y \end{array}$$

Andando avanti:

$$\begin{array}{lllll} x^6y^3+2x^3y^2-y+1 & |xy^2-x & |y^3-x & |RESTO \\ & |\overline{\qquad} & |\overline{\qquad} & \\ x^6y^3-x^6y & |x^5y+2x^2 & | & \\ \overline{\qquad\qquad\qquad\qquad} & & & \\ x^6y+2x^3y^2-y+1 & & & \\ 2x^3y^2-y+1 & \longrightarrow & & x^6y \\ 2x^3y^2-2x^3 & & & \\ \overline{\qquad\qquad\qquad\qquad} & & & \\ 2x^3-y+1 & & & \\ 0 & \longrightarrow & & 2x^3-y+1 \end{array}$$

da cui

$$x^6y^3+2x^3y^2-y+1=(x^5y+2x^2)(xy^2-x)+0(y^3-x)+x^6y+2x^3-y+1$$

Se scambiamo l'ordine dei divisori otteniamo

$$\begin{array}{lllll} x^6y^3+2x^3y^2-y+1 & |y^3-x & |xy^2-x & |RESTO \\ & |\overline{\qquad} & |\overline{\qquad} & \\ x^6y^3-x^7 & |x^6 & |2x^2 & \\ \overline{\qquad\qquad\qquad\qquad} & & & \\ x^7+2x^3y^2-y+1 & & & \\ 2x^3y^2-y+1 & \longrightarrow & & x^7 \\ 2x^3y^2-2x^3 & & & \\ \overline{\qquad\qquad\qquad\qquad} & & & \\ 2x^3-y+1 & & & \\ 0 & \longrightarrow & & 2x^3-y+1 \end{array}$$

da cui

$$x^6y^3 + 2x^3y^2 - y + 1 = 2x^2(xy^2 - x) + x^6(y^3 - x) + x^7 + 2x^3 - y + 1$$

Quindi scambiando l'ordine dei divisori sia i quozienti che il resto cambiano.

Esercizio 1.16. *Verificare che dividendo con Lex* $xy^2 - x$ *per* $xy + 1$ *e* $y^2 - 1$ *otteniamo*

$$xy^2 - x = y(xy + 1) + 0(y^2 - 1) - y - x$$

mentre scambiando l'ordine dei divisori:

$$xy^2 - x = 0(xy + 1) + x(y^2 - 1)$$

Osserviamo dall'esercizio precedente che la condizione

$$\text{resto della divisione di } f \text{ per } \{f_1, \ldots, f_h\} \quad = \quad 0$$

è una condizione sufficiente ma non necessaria affinché $f \in (f_1, \ldots, f_h)$. Questo è in contrasto con quanto accade per polinomi in una variabile. Ovvieremo a questo inconveniente definendo un insieme opportuno di generatori per l'ideale $(f_1, \ldots, f_h)$, "adattato" all'ordine monomiale scelto.

L'algoritmo di divisione si riassume nel seguente risultato

Teorema 1.17. *Sia fissato un ordine monomiale in* $K[x_1, \ldots, x_n]$*. Siano dati*

$$f, f_1, \ldots, f_h \in K[x_1, \ldots, x_n]$$

Allora esistono

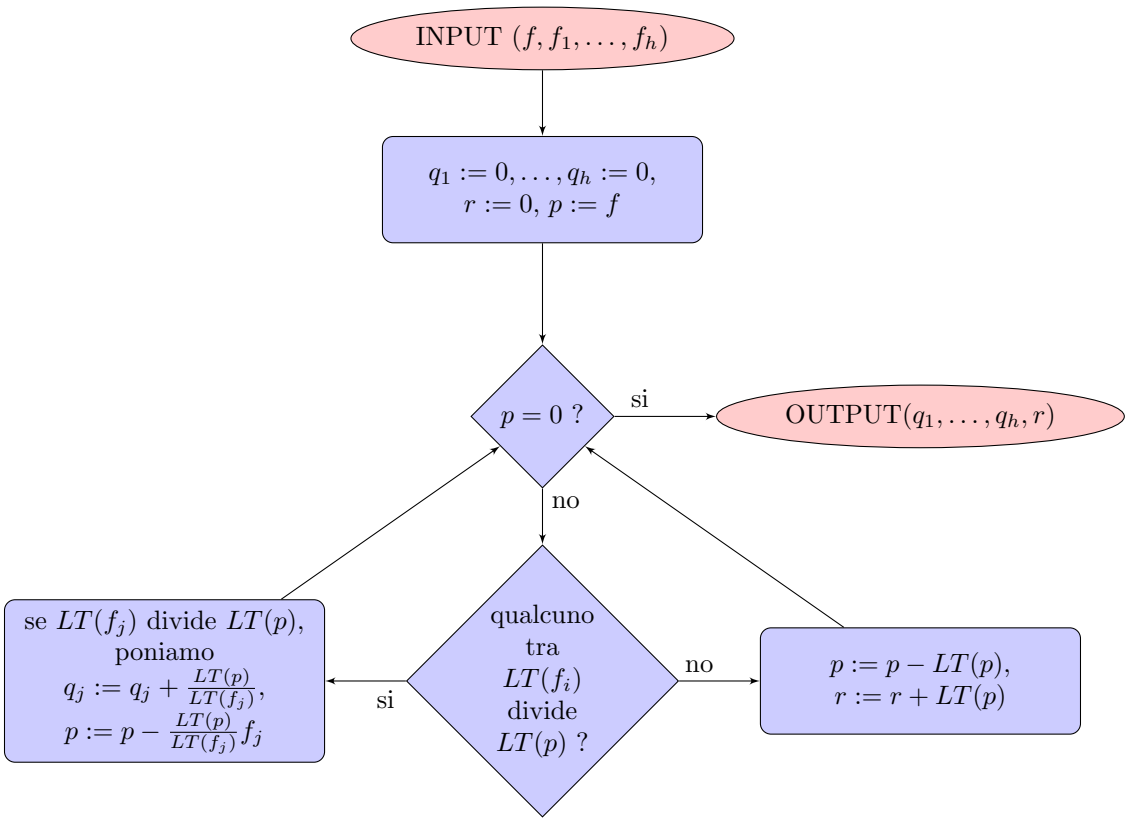
$$q_1, \ldots, q_h, r \in K[x_1, \ldots, x_n]$$

tali che

- *i)*$f = \sum q_i f_i + r$
- *ii) nessun termine di* r *è divisibile per qualcuno tra* $LT(f_1), \ldots, LT(f_h)$*.*
- *iii) se* $q_i f_i \neq 0$ *vale* $\mathrm{LT}(f) \geq \mathrm{LT}(q_i f_i)$*.*

In piú esiste un algoritmo che determina $q_1, \ldots, q_h, r$*.*

La dimostrazione segue dal seguente diagramma di flusso



Dimostrazione. Si introduce un resto ausiliario p, in modo che l'equazione $f = \sum_{i=1}^{h} f_i q_i + r + p$ vale con le posizioni iniziali di q_i, r, p e continua a valere ad ogni passo dell'algoritmo. Quando si arriva all'output, tutti i termini in r provengono da termini di p che non sono divisibili per nessuno dei $LT(f_i)$. Ad ogni passo, ciascun termine di q_i, moltiplicato per $LT(f_i)$, risulta $\leq LT(p)$ rispetto al passo precedente. Come conseguenza, ciascun termine di $q_i f_i$ risulta $\leq LT(f)$. L'algoritmo termina perché, sia nel ciclo sinistro che nel ciclo destro, la successione dei $LT(p)$ è decrescente, e ogni ordine monomiale è un buon ordinamento per la 1.12. Pertanto, al termine, avremo $p = 0$. □

2 Ideale dei leading term e basi di Gröbner

Il prossimo Lemma estende il Lemma 1.9 dai monomi ai polinomi.

Lemma 2.1. *Sia I un ideale monomiale. Sono equivalenti*

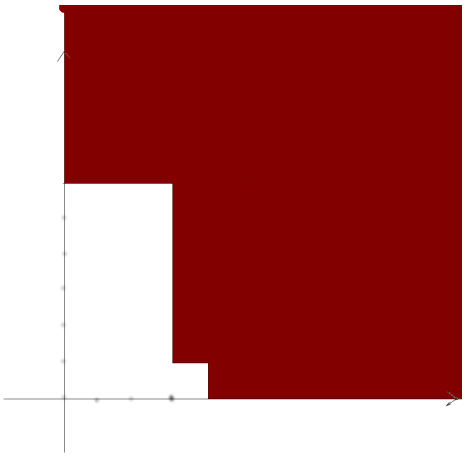
- *i) $f \in I$*
- *ii) ogni termine di f appartiene a I.*

Dimostrazione. ii) ⇒ i) è banale.

Per provare i) ⇒ ii) scriviamo $f = \sum_i f_i$ (ogni f_i è un termine)$= \sum_j g_j m_j$ (ogni m_j è un monomio in I). A secondo membro ogni termine appartiene a I, quindi questo è vero anche a primo membro che è ottenuto cancellando tra loro alcuni termini a secondo membro. □

Corollario 2.2. *Due ideali monomiali sono uguali se e solo se contengono gli stessi monomi.*

Il corollario precedente permette quindi di identificare gli ideali monomiali con dei sottoinsiemi di $\mathbf{Z}^n_{\geq 0}$. Ad esempio l' ideale monomiale (x^4, x^3y, y^6) in $K[x, y]$ corrisponde alla regione seguente:



Da queste rappresentazioni il lettore interessato può ricavare una dimostrazione diretta del lemma di Dickson indipendente dal teorema della base di Hilbert (si veda [CLO]).

Definizione 2.3. *Un insieme di monomi B si dice una base minimale per un ideale monomiale I se i monomi di B generano I e se nessun monomio di B divide qualche altro monomio di B.*

La seguente proposizione dovrebbe essere evidente dalle rappresentazioni grafiche descritte sopra.

Proposizione 2.4. *Sia I un ideale monomiale. Allora esiste una unica base minimale per I.*

Dimostrazione. L'esistenza di una base minimale segue subito prendendo un insieme di generatori ed eliminando i monomi divisi da qualcun altro. L'unicità è evidente dal lemma 1.9. □

Definizione 2.5. *Sia I un ideale di $K[x_1, \ldots, x_n]$ e sia fissato un ordine monomiale.*
$LT(I)$ è l'ideale (monomiale) generato da tutti i termini $LT(f)$ dove $f \in I$. In formula

$$LT(I) \colon = \langle LT(f) | f \in I \rangle$$

Il comando corrispondente in Macaulay2 è leadTerm(I)*, con l'algoritmo di Buchberger del §3 vedremo come calcolare $LT(I)$.*

Osservazione Se $I = (g_1, \ldots g_k)$ allora $LT(I) \quad \supset (LT(g_1), \ldots, LT(g_k))$ ma può valere l'inclusione stretta come mostra il seguente

Esempio 2.6. *Sia $I = \left(x^2 + y, x^2 - y\right) \subset K[x, y]$ con un qualunque ordine monomiale graduato. Allora $y \in I$ da cui $y \in LT(I)$ mentre $y \notin \left(LT(x^2 + y), LT(x^2 - y)\right) = \left(x^2\right)$*

L'osservazione precedente motiva la seguente

Definizione 2.7. *Un insieme $(g_1, \ldots, g_k)$ di elementi di I si dice una base di Gröbner per I se*

$$LT(I) = (LT(g_1), \ldots, LT(g_k))$$

Proposizione 2.8. *Ogni ideale di $K[x_1, \ldots, x_n]$, con un ordine monomiale fissato, ammette una base di Gröbner.*

Dimostrazione. È sufficiente estrarre dall'insieme $\{LT(f)|f \in I\}$ un numero finito di generatori per $LT(I)$. Questo è sempre possibile per noetherianità (si veda la prop.1.5). □

La dimostrazione precedente è non costruttiva. Buchberger sviluppó nel 1965 (nella sua tesi di dottorato) un algoritmo per calcolare effettivamente una base di Gröbner a partire da un insieme di generatori. Vedremo questo algoritmo nel Capitolo 3.

Teorema 2.9. *Una base di Gröbner per I genera I.*

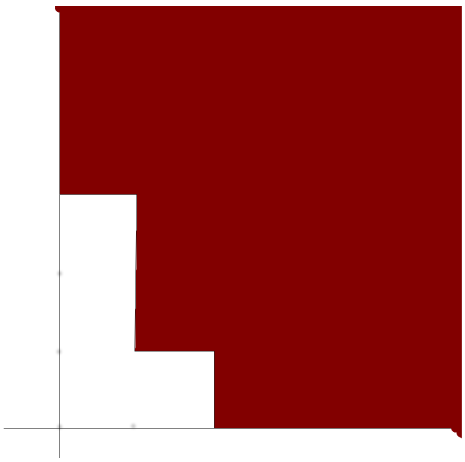
Dimostrazione. Sia $g_1, \ldots, g_k$ una base di Gröbner, pertanto $LT(I) = (LT(g_1), \ldots, LT(g_k))$. Se $f \in I$ allora per l'algoritmo di divisione possiamo scrivere $f = \sum a_i g_i + r$ da cui $r = f - \sum a_i g_i \in I$ ed in particolare $LT(r) \in (LT(g_1), \ldots, LT(g_k))$. Se fosse $r \neq 0$ allora dal teorema 1.17 $LT(r)$ non è divisibile per nessuno dei $LT(g_i)$ e questa è una contraddizione con il lemma 1.9 □

Esempio 2.10. *Sia $I = \left(x^2 + y^2, xy\right) \subset K[x, y]$ con l'ordine Lex. Una base di Gröbner per I è costituita da almeno tre elementi.*

Infatti osserviamo che tutti i monomi di grado 3 appartengono ad I (e quindi anche a $LT(I)$:

$$\begin{aligned} x^3 &= x(x^2 + y^2) - y(xy) \\ x^2y &= x(xy) \\ xy^2 &= y(xy) \\ y^3 &= y(x^2 + y^2) - x(xy). \end{aligned}$$

Ne segue che tutti i polinomi omogenei di grado 3 appartengono a $LT(I)$. Siccome ogni monomio di grado ≥ 3 é divisibile per un monomio di grado 3, segue che ogni monomio di grado ≥ 3 appartiene a I (e quindi anche a $LT(I)$. Anche x^2 e xy appartengono a $LT(I)$ e devono appartenere ad un qualunque insieme di generatori di $LT(I)$. Infine notiamo che $y^3 \notin (x^2, xy)$ e quindi sono necessari almeno tre elementi come asserito. La rappresentazione grafica di $LT(I)$ è la seguente:



Esercizio 2.11. *Trovare una base di Gröbner per I dell'esempio precedente (affronteremo di nuovo questo problema nell'esercizio 3.4 1).Risposta: $x^2 + y^2, xy, y^3$.*

L'utilità delle basi di Gröbner è subito illustrata dal seguente:

Teorema 2.12. *[Forma normale di un elemento rispetto a un ideale] Sia fissato un ordine monomiale e per l'ideale $I \subset K[x_1, \ldots, x_n]$. Sia $f \in K[x_1, \ldots, x_n]$. Allora esiste unico $r \in K[x_1, \ldots, x_n]$ tale che:*

- *i) nessun termine di r appartiene a $LT(I)$.*
- *ii) esiste $g \in I$ tale che $f = g + r$.*

In particolare r è il resto della divisione di f per una base di Gröbner di I. r si dice la forma normale di f rispetto a I e dipende solo dall'ordine monomiale fissato.

Dimostrazione. Sia $G = \{g_1, \ldots, g_t\}$ una base di Gröbner, dunque $LT(I) = \langle LT(g_1), \ldots, LT(g_t)\rangle$. Scegliamo r come il resto della divisione di f per G, pertanto r esiste (vedi teor. 1.17). Per provare l'unicità consideriamo $f = g' + r' = g'' + r''$. Allora $r'' - r' = g' - g'' \in I$ da cui $LT(r'' - r') \in LT(I) = (LT(g_1), \ldots, LT(g_k))$. Se $r'' - r' \neq 0$ allora $LT(r'' - r')$ sarebbe divisibile per qualche $LT(g_i)$ e questo è impossibile perché nessun termine di r' o di r'' è divisibile per qualche $LT(g_i)$. □

Esercizi.

1. Sia $I = (g_1, g_2, g_3) \subset \mathbb{R}[x, y, z]$ dove $g_1 = xy^2 - xz + y$, $g_2 = xy - z^2$, $g_3 = x - yz^4$. Utilizzando Lex, dare un esempio di $g \in I$ tale che $LT(g) \not\in < LT(g_1), LT(g_2), LT(g_3) >$.
2. Sia $G = \{x^4y^2 - z^5, x^3y^3 - 1, x^2y^4 - 2z\}$. Provare che G non è una base di Gröbner per $< G >$ rispetto a GRevLex.
3. Sia $I \subset K[x_1, \ldots, x_n]$ un ideale principale. Provare che un sottoinsieme di I è una base di Gröbner per I se e solo se contiene un generatore di I.
4. Sia $I = (f)$ un ideale principale. Provare che $LT(I)$ è principale ed è generato da $LT(f)$. In questo caso il problema di appartenenza si risolve mediante la divisione per f (perché ?).
5. Provare che g divide f se e solo se la divisione di f per g dà resto zero. Questo non dipende dall' ordine monomiale scelto.
6. Calcolare la forma normale di x^2y rispetto a $I = (x^2 + y^2, xy)$ effettuando la divisione rispetto alla base di Gröbner $\{x^2 + y^2, xy, y^3\}$. Notare che i quozienti dipendono dall'ordine dei tre elementi (si provi l'ordine $\{xy, x^2 + y^2, y^3\}$.

Corollario 2.13. *Quando si divide per una base di Gröbner l'algoritmo di divisione porta sempre allo stesso resto qualunque sia l'ordine dei divisori.*

Pertanto secondo il corollario precedente esempi come 1.15 non possono capitare se si divide per una base di Gröbner. Di più vale

Corollario 2.14. *$f \in I \iff$ il resto della divisione di f con una base di Gröbner di I è zero*

Dimostrazione. $\Leftarrow$ è ovvia
$\Rightarrow$ $f = f + 0$ nel teorema 2.12. □

Il corollario precedente risolve quindi il problema di appartenenza posto dopo la prop. 1.5 se si conosce una base di Gröbner.

Definizione 2.15. *Scriveremo $f \,\%\, I$ per indicare la forma normale di f rispetto a I. $f \,\%\, I$ dipende solo dall'ordine monomiale, ed in particolare*

$$f \in I \iff f \,\%\, I = 0$$

In Macaulay2 f % I è il resto della divisione di f per una base di Gröbner di I.

Esercizio 2.16. *Sia $f = x^4y + y^3$ e $I = (x^2 + y^2, xy)$. Fissato l'ordine Lex, calcolare f % I.*

Osservazione L'esercizio 3.4, 2) mostra che i quozienti non sono unici: l'unicità del resto nella divisione è il massimo che si riesce ad ottenere.

Definizione 2.17. *Siano $f, g \in K[x_1, \dots, x_n]$ e sia $x^\gamma = m.c.m.(\mathrm{LT}(f), \mathrm{LT}(g))$. Definiamo la S-coppia:*

$$S(f,g) := \frac{x^\gamma}{LT(f)} f - \frac{x^\gamma}{LT(g)} g$$

Notiamo che in $S(f,g)$ i termini di multigrado γ si cancellano mentre tutti gli altri termini hanno multigrado $< \gamma$. Pertanto

$$\text{MULTIDEG } S(f,g) < \gamma$$

Una "ostruzione" a che $\{f_1, \dots, f_h\}$ sia una base di Gröbner è

$$LT\left(S(f_i, f_j)\right) \notin (LT(f_1), \dots, LT(f_h))$$

Vedremo che questa è in sostanza l'unica ostruzione.

Lemma 2.18. *Supponiamo di avere una cancellazione tra i LT di un insieme di polinomi g_i. Cioè supponiamo di avere una combinazione $\sum_{i=1}^t c_i x^{\alpha_i} g_i$ con $c_i \in K$, $\alpha_i + \text{MULTIDEG } g_i = \delta$ (se $c_i \neq 0$) e $\text{MULTIDEG } (\sum c_i x^{\alpha_i} g_i) < \delta$. Allora, posto $x^{\gamma_{jk}} := m.c.m.(\mathrm{LT}(g_j), \mathrm{LT}(g_k))$ esistono c_{jk} tali che*

$$\sum_{i=1}^t c_i x^{\alpha_i} g_i = \sum_{j,k=1}^t c_{jk} x^{\delta - \gamma_{jk}} S(g_j, g_k)$$

In particolare ogni termine del secondo membro ha multigrado $< \delta$

Dimostrazione. Poniamo $LT(g_i) := d_i x^{\beta_i}$, quindi

$$\alpha_i + \beta_i = \delta \tag{2.1}$$

$$\sum_{i=1}^t c_i d_i = 0 \tag{2.2}$$

Adesso $x^{\delta-\gamma_{jk}} S(g_j, g_k) = x^{\delta-\gamma_{jk}} \left(\frac{x^{\gamma_{jk}} g_j}{d_j x^{\beta_j}} - \frac{x^{\gamma_{jk}} g_k}{d_k x^{\beta_k}} \right) = (\text{per } 2.1) = \frac{x^{\alpha_j} g_j}{d_j} - \frac{x^{\alpha_k} g_k}{d_k}$
Quindi $\sum_{i=1}^t c_i x^{\alpha_i} g_i = \sum_{i=1}^t c_i d_i \left(\frac{x^{\alpha_i} g_i}{d_i} \right) = (\text{ponendo } g_{t+1} = 0)$
$= \sum_{i=1}^t c_i d_i \left(\sum_{j=i}^t \left(\frac{x^{\alpha_j} g_j}{d_j} - \frac{x^{\alpha_{j+1}} g_{j+1}}{d_{j+1}} \right) \right) = (\text{scambiando le sommatorie})$
$= \sum_{j=1}^t \sum_{i=1}^j c_i d_i \left(\frac{x^{\alpha_j} g_j}{d_j} - \frac{x^{\alpha_{j+1}} g_{j+1}}{d_{j+1}} \right) = (\text{usando } 2.2)$
$\sum_{j=1}^{t-1} \sum_{i=1}^j c_i d_i x^{\delta - \gamma_{j,j+1}} S(g_j, g_{j+1})$ □

3 L'algoritmo di Buchberger

Teorema 3.1. *(Criterio di Buchberger, 1965) Sia I un ideale di $K[x_1, \dots, x_n]$ generato da $(g_1, \dots, g_t)$. Sia fissato un ordine monomiale.*

$$\begin{array}{c} (g_1, \dots, g_t) \\ \textit{è una base di Gröbner per } I \end{array} \iff \begin{array}{c} \textit{il resto della divisione} \\ \textit{di } S(g_i, g_j) \textit{ per } (g_1, \dots g_t) \\ \textit{è zero } \forall i \neq j \end{array}$$

La divisione di $S(g_i, g_j)$ per $(g_1, \dots g_t)$ può essere effettuata prendendo $g_1, \dots, g_t$ in un ordine qualunque; segue in particolare (vedi cor.2.13) che se il resto è zero in un ordine rimane zero in tutti gli altri ordini.

Dimostrazione. $\Rightarrow$ è ovvia da $S(g_i, g_j) \in I$ e dal cor.2.14

$\Leftarrow$ Sia $f \in I$, voglio provare che $LT(f) \in < LT(g_1), \dots, LT(g_t) >$. Per ipotesi abbiamo $f = \sum h_i g_i$ con MULTIDEG $(f) \leq max\,(\text{MULTIDEG}(h_i g_i))$. Sia δ il minimo tra tutte le espressioni $f = \sum h_i g_i$ di $max\,(\text{MULTIDEG}(h_i g_i))$, tale minimo esiste per la proprietà di buon ordinamento (Prop. 1.12). Se MULTIDEG $f = \delta$ abbiamo concluso. Sia per assurdo MULTIDEG $f < \delta$ e poniamo $m(i) :=$ MULTIDEG $(g_i h_i)$. Abbiamo:

$$f = \sum h_i g_i = (\text{dove il massimo MULTIDEG a secondo membro è } \delta)$$
$$= \sum_{m(i)=\delta} h_i g_i + \sum_{m(i)<\delta} h_i g_i = \sum_{m(i)=\delta} LT(h_i) g_i + \sum_{m(i)=\delta} (h_i - LT(h_i))\, g_i + \sum_{m(i)<\delta} h_i g_i$$

I monomi nella seconda e terza somma hanno MULTIDEG $< \delta$. Quindi nella prima somma abbiamo cancellazioni tra i LT dei g_i e possiamo applicare il lemma 2.18. Pertanto possiamo riscrivere la prima somma come combinazione dei $S(g_j, g_k)$. A loro volta questi ultimi polinomi possono essere scritti nella forma $S(g_j, g_k) = \sum a_{ijk} g_i$ per l'algoritmo di divisione (l'ipotesi è resto zero!). Siccome MULTIDEG $(a_{ijk} g_i) \leq$ MULTIDEG $S(g_j, g_k)$ (si veda il teor. 1.17), risostituendo nell' espressione iniziale abbiamo $f = \sum h'_i g_i$ con tutti i multigradi a secondo membro $< \delta$. Pertanto δ non è il minimo. Questa contraddizione conclude il ragionamento. $\square$

Esempio 3.2. *Proviamo che $\left(x - z^4, y - z^{10}\right)$ è una base di Gröbner secondo Lex utilizzando il criterio di Buchberger. Abbiamo la sola S-coppia $S(x - z^4, y - z^{10}) = \frac{xy}{x}(x - z^4) - \frac{xy}{y}(y - z^{10}) = yx - yz^4 - xy + xz^{10} = xz^{10} - yz^4$. La divisione porta a:*

$xz^{10} - yz^4$	$\vert x - z^4$	$\vert y - z^{10}$	$\vert RESTO$
	$\vert$———	$\vert$———	
$xz^{10} - z^{14}$	$\vert z^{10}$	$\vert z^4$	
———			
$-yz^4 + z^{14}$	$\longrightarrow$		0
$yz^4 - z^{14}$			
———			
0	$\longrightarrow$		0

Il resto è zero e quindi la condizione del criterio di Buchberger è verificata.

Osservazione critica. Nel corso della dimostrazione dell'implicazione $\Leftarrow$ del criterio di Buchberger abbiamo scritto $S(g_j, g_k) = \sum a_{ijk} g_i$ usando l'algoritmo di divisione. Perché non si é fatto uso direttamente della definizione di $S(g_j, g_k)$ che lo esprime come combinazione di g_j e g_k?. Il punto é che la disuguaglianza

$$\text{MULTIDEG } (a_{ijk} g_i) \leq \text{MULTIDEG } S(g_j, g_k)$$

non sarebbe stata soddisfatta!

La dimostrazione del Teorema 3.1 fornisce immediatamente il seguente rafforzamento del Criterio di Buchberger.

Corollario 3.3. *Sia I un ideale di $K[x_1, \dots, x_n]$ generato da $(g_1, \dots, g_t)$. Sia fissato un ordine monomiale.*

$$\begin{array}{c}(g_1, \dots, g_t) \\ \textit{è una base di Gröbner per } I\end{array} \iff \begin{array}{c}\textit{esistono polinomi } a_{ijk} \textit{ tali che} \\ S(g_j, g_k) = \sum_i a_{ijk} g_i \\ \text{LT } (a_{ijk} g_i) \leq \text{LT } S(g_j, g_k) \\ \forall i, j, k,\ i \neq j\end{array}$$

Esercizio 3.4. *Esercizi*

1. *Provare che* $G = \{x + z, y - z\}$ *è una base di Gröbner rispetto a Lex.*
2. *Dividere* xy *per* $x+z$, $y-z$ *(rispetto a Lex, si veda l'eserc. 1). Poi dividere* xy *per* $y-z$, $x+z$. *Nei due casi il resto è lo stesso, in accordo con il Teorema 2.12, ma i quozienti sono differenti. Quindi questo esercizio mostra che non si riesce ad avere l'unicità dei quozienti, in contrasto col caso dei polinomi in una sola variabile. Concludere che la forma normale di* xy *rispetto a* $I = (x + z, y - z)$ *è* $-z^2$.
3. *Si calcoli* $S(f, g)$ *rispetto a Lex nei seguenti casi:*
   (a) $f = 4x^2z - 7y^2, g = xyz^2 + 3xz^4$
   (b) $f = x^4y - z^2, g = 3xz^2 - y$
   (c) $f = x^7y^2z + 2ixyz, g = 2x^7y^2z + 4$
   (d) $f = xy + z^3, g = z^2 - 3z$
4. $S(f, g)$ *dipende dall'ordine monomiale scelto?*

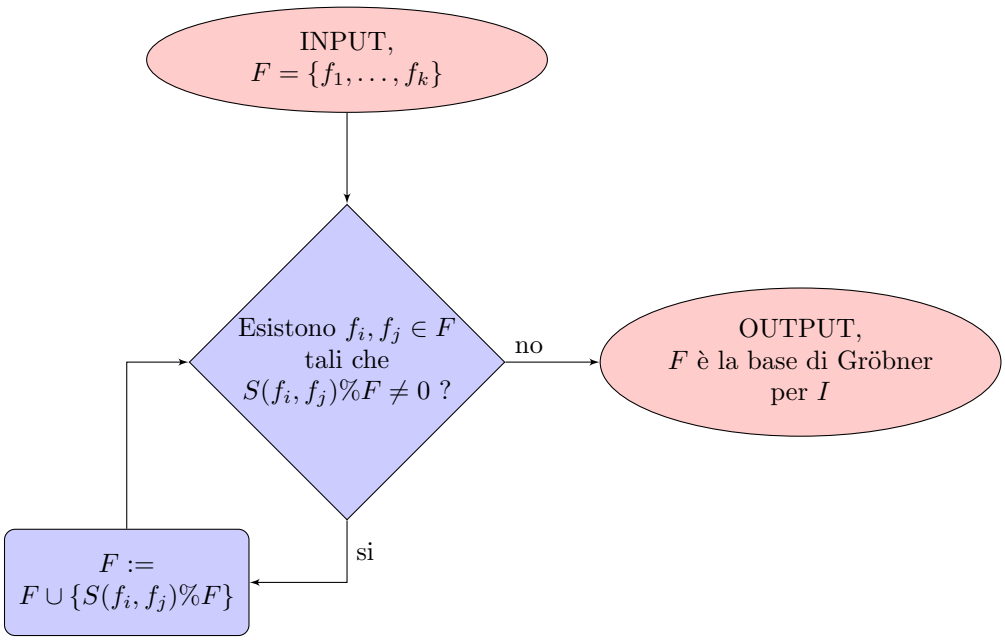
L'algoritmo di Buchberger

Il criterio di Buchberger suggerisce un algoritmo per costruire una base di Gröbner. Si considera $F = \{f_1, \ldots, f_k\}$ insieme di generatori di I. Indichiamo provvisoriamente con $f\%\ F$ il resto della divisione di f per gli elementi di F nell'ordine in cui sono scritti. Aggiungiamo ad F stesso tutti gli elementi $[S(f_i, f_j)\%\ F]$ e ripetiamo questa operazione col nuovo insieme F (più grande !). Continuando in questo modo si ottiene corrispondentemente una catena ascendente di ideali monomiali data ad ogni passo da $< LT(F) >$. Per Noetherianità la catena diventa stazionaria e questo vuol dire esattamente che dopo un certo numero di passi $[S(f_i, f_j)\%\ F] = 0 \quad \forall \quad i, j$ e quindi per il criterio di Buchberger quando l'algoritmo ha termine F è una base di Gröbner.

Formalmente abbiamo:

Diagramma di flusso dell'algoritmo di Buchberger.

Sia fissato un ordine monomiale. Dato $F = \{f_1, \ldots, f_k\}$, insieme di generatori di un ideale $I \subset K[x_1, \ldots, x_n]$, si costruisce $G = \{g_1, \ldots, g_s\}$ base di Gröbner per I. In particolare $LT(g_1), \ldots, LT(g_s)$ generano $LT(I)$ che viene cosìdeterminato.



Dimostrazione La notazione $S(f_i, f_j)\%F$ indica il resto della divisione della S-coppia $S(f_i, f_j)$ per tutti gli elementi di F. Il fatto che l'output è una base di Gröbner è garantito dal criterio di Buchberger, che è verificato esattamente quando la risposta al test di esistenza di (f_i, f_j) è negativa. Il ciclo non può essere percorso infinite volte, perché ad ogni passo l'ideale monomiale $< LT(F) >$ contiene strettamente l'ideale monomiale del passo precedente. Non si può ottenere una catena ascendente infinita di ideali, perché $K[x_1, \ldots, x_n]$ è noetheriano.

Osservazione importante L'algoritmo permette di avere delle espressioni esplicite $g_i = \sum_{j=1}^{k} a_{ij} f_j$, di ciascun elemento della base di Gröbner in funzione dei generatori. Infatti, ogni elemento che viene aggiunto ad F, ad ogni passo, si può esprimere come combinazione degli elementi già presenti in F. Per esempio, se $F = \{f_1, \ldots, f_p\}$, abbiamo dall'algoritmo di divisione $[S(f_i, f_j)\%F] = S(f_i, f_j) - \sum_{i=1}^{p} q_i f_i$ e anche $S(f_i, f_j)$ è combinazione di elementi di F.

Esercizi.

1. Si trovi una base di Gröbner per l'ideale $I = (x^2 + y^2, xy)$ rispetto a Lex utilizzando l'algoritmo di Buchberger.
2. Sia $A = (a_{ij})$ una matrice $n \times m$ a scala a coefficienti reali e sia $J \subset \mathbb{R}[x_1, \ldots, x_m]$ l'ideale generato dai polinomi $f_i = \sum_{j=1}^{m} a_{ij} x_j$ per $1 \leq i \leq n$. Provare che $LT(J)$ è generato dalle variabili dei pivot. Provare che i generatori di J formano una base di Gröbner rispetto all'ordine Lex dove $x_1 > x_2 > \ldots$. *Suggerimento: ogni monomio di $S(f_i, f_j)$ è divisibile per almeno una variabile di un pivot.*
3. Sia $A = (a_{ij})$ una matrice $n \times m$ a coefficienti reali e sia $J \subset \mathbb{R}[x_1, \ldots, x_m]$ l'ideale generato dai polinomi $\sum_{j=1}^{m} a_{ij} x_j$ per $1 \leq i \leq n$. Provare che $LT(J)$ è generato dalle variabili dei pivot di una riduzione a scala di A.
4. Sia $I = (f, g) \subset K[x]$ un ideale nell'anello dei polinomi nella variabile x. Provare che $LT(I)$ è generato da $LT(\mathrm{GCD}(f, g))$ e che l'algoritmo di Buchberger si riconduce essenzialmente all'algoritmo euclideo.

Come conseguenza abbiamo il seguente:

Algoritmo per la forma normale di un elemento f rispetto a un ideale I.

Questo algoritmo è una sorta di *divisione "intelligente"* per polinomi in più variabili. Fissiamo un ordine monomiale. Dati $f, f_1, \ldots, f_h \in K[x_1, \ldots, x_n]$, denotiamo con I l'ideale generato da $(f_1, \ldots, f_h)$, si costruiscono $q_1, \ldots, q_h, r \in K[x_1, \ldots, x_n]$ tali che

$$f = \sum_{i=1}^{h} f_i q_i + r, \quad \text{dove}$$

- nessun termine di r appartiene a $LT(I)$.

In particolare, r è la forma normale di f rispetto a I, denotata con $f\%I$, e implementata in *Macaulay2* con questa stessa sintassi. Notiamo che l'equivalenza

$$f \in I \Longleftrightarrow (f\%I = 0)$$

permette di risolvere in modo effettivo il *problema di appartenenza* di f ad I.

Passi dell'algoritmo

1. Si calcola una base di Gröbner $G = \{g_1, \ldots, g_s\}$ di I mediante l'algoritmo di Buchberger.
2. L'algoritmo permette anche di avere delle espressioni

$$g_i = \sum_{j=1}^{h} a_{ij} f_j \tag{3.1}$$

3. Si divide f per $\{g_1, \ldots, g_s\}$, ottenendo

$$f = \sum_{i=1}^{s} q_i' g_i + r \tag{3.2}$$

   dove nessun termine di r è divisibile per $LT(g_1), \ldots LT(g_s)$. Dato che G è una base di Gröbner, segue che nessun termine di r appartiene a $LT(I)$.

4. Si sostituiscono le espressioni (3.1) nella formula (3.2).

Questo algoritmo è implementato in *Macaulay2*, con il comando

```
(q,r)=quotientRemainder(matrix{{f}},matrix{{f_1,..., f_h}})
```

Nell'output, q è una matrice che contiene i quozienti $q_1, ...q_h$ e r è il resto.

Esercizi.

1. Determinare se $f = xy^3 - z^2 + y^5 - z^3$ appartiene all'ideale $I = (-x^3 + y, x^2y - z)$.Suggerimento: utilizzando GRevLex la base di Gröbner di I é costituita da 3 elementi, mentre utilizzando Lex o GLex i calcoli sono piú complessi.
2. Determinare se $f = x^3z - 2y^2$ appartiene all'ideale $I = (xz - y, xy + 2z^2, y - z)$.

4 Il teorema di eliminazione e l'intersezione di due ideali

4.1 Eliminazione

Definizione 4.1. *Sia I un ideale di $K[x_1, \ldots, x_n]$, si pone*

$$I_k := I \cap K[x_{k+1}, \ldots, x_n]$$

I_k contiene le "conseguenze" dei polinomi di I che coinvolgono solo le variabili $x_{k+1}, \ldots, x_n$.

Teorema 4.2. *(di eliminazione). Sia I un ideale di $K[x_1, \ldots, x_n]$. Sia G una base di Gröbner per I rispetto a Lex. Allora $G_k := G \cap K[x_{k+1}, \ldots, x_n]$ è una base di Gröbner per I_k.*

Dimostrazione. Riordiniamo gli elementi di $G = \{g_1, \ldots, g_m\}$ in modo che i primi r elementi $\{g_1, \ldots, g_r\}$ formino G_k. Facciamo vedere che $\{g_1, \ldots, g_r\}$ generano I_k. Dato $f \in I_k$, abbiamo che il resto della divisione di f per G è zero. Notiamo che $LT(g_{r+1}), \ldots, LT(g_m)$ contengono termini dove compare qualche $x_1, \ldots, x_k$ e quindi hanno multigrado maggiore (per Lex) di ogni monomio di f. Pertanto $g_{r+1}, \ldots g_m$ non entrano in gioco nella divisione di f per G e risulta $f = \sum_{i=1}^{r} a_i g_i$ come volevamo.

Usiamo il criterio di Buchberger per provare che $\{g_1, \ldots, g_r\}$ è una base di Gröbner per I_k. Se $1 \leq j, k \leq r$ abbiamo $S(g_j, g_k) \in I_k$. Per quanto visto sopra la divisione di $S(g_j, g_k)$ per G coincide con la divisione per G_k, quindi il resto della divisione è zero come volevamo. □

Esercizi.

1. Provare che se $I = (x - y, x^2 + y^3) \subset K[x, y]$ allora $I_1 = (y^3 + y^2)$.
2. Provare che se $= (-x^3 + y, x^2 y - z) \subset K[x, y, z]$ allora $I_1 = (y^5 - z^3)$ e $I_2 = 0$.

4.2 Intersezione di due ideali

Vediamo adesso un algoritmo che permette di calcolare i generatori di $(f_1, \ldots, f_r) \cap (g_1, \ldots, g_s)$. Questo problema non è banale perché nel caso $(f) \cap I$ contiene il problema di appartenenza "$f \in I$?". Infatti $f \in I \iff (f) \cap I = (f)$.

Siano I, J due ideali di $K[x_1, \ldots, x_n]$. Definiamo tI come l'ideale di $K[x_1, \ldots, x_n, t]$ generato da tf con $f \in I$. Analogamente si puó definire $(1 - t)J$. Vale la

Proposizione 4.3. *[L'intersezione tra ideali si riconduce ad una eliminazione]*

$$I \cap J = [tI + (1 - t)J] \cap K[x_1, \ldots, x_n]$$

Dimostrazione.

$\subset$ è ovvia scrivendo $f = tf + (1 - t)f$

$\supset$ Sia $f(x) \in [tI + (1 - t)J] \cap K[x_1, \ldots, x_n]$. Pertanto $f(x) = g(x, t) + h(x, t)$ con $g \in tI$ e $h \in (1 - t)J$. In particolare

$$g(x, 0) = 0 \text{ da cui } f(x) = h(x, 0) \in J$$

$$h(x, 1) = 0 \text{ da cui } f(x) = g(x, 1) \in I$$

Quindi $f \in I \cap J$ come volevamo. □

La proposizione precedente dà un algoritmo per calcolare l'intersezione di due ideali. Infatti se $I = (f_1, \ldots, f_r)$ e $J = (g_1, \ldots, g_s)$ allora si può trovare una base di Gröbner (e quindi un insieme di generatori) di $I \cap J$ eliminando t da

$$(tf_1, \ldots, tf_r, (1 - t)g_1, \ldots, (1 - t)g_s)$$

utilizzando il teorema 4.2

4.3 Massimo comun divisore e minimo comune multiplo tra polinomi

Il minimo comune multiplo tra due polinomi f e g si può trovare come il generatore dell'ideale intersezione $(f)\cap(g)$. Un algoritmo per calcolarlo segue quindi dall'algoritmo per l'intersezione della Prop. 4.3, che a sua volta segue dall'eliminazione.

Siccome $M.C.D.(f,g) = \frac{fg}{m.c.m.(f,g)}$, abbiamo anche un algoritmo per calcolare il MCD. MCD e mcm possono essere trovati in M2 con i comandi $\gcd(f_1,\ldots,f_k)$ e $\mathrm{lcm}(f_1,\ldots,f_k)$, applicabili anche a più di due polinomi.

5 Complementi sugli ideali di un anello commutativo

Sia A un anello commutativo con unità (ad esempio $A = K[x_1,\ldots,x_n]$).

Se I, J sono ideali di A allora $I \cup J$ non è un ideale in generale

$$(x, y \in (x)\cup(y) \text{ ma } x+y \notin (x)\cup(y))\,.$$

Invece

$$I + J$$

$$I \cap J$$

$$IJ := < ij | i \in I, j \in J >$$

$$I\colon J \quad := \{a \in A | aj \in I \ \forall j \in J\}$$

sono tutti ideali di A.

Osservazione Vale $IJ \subset I \cap J$ e l'inclusione può essere stretta (esempio: $I = J = (x)$).

Definizione 5.1. *Se $X \subset K^n$ è un sottoinsieme poniamo*

$$I(X) := \{f \in K[x_1,\ldots,x_n] | f(x) = 0 \quad \forall x \in X\}$$

È immediato verificare che I è un ideale di $K[x_1,\ldots,x_n]$.

$I(X)$ è un ideale che conserva alcune informazioni geometriche del sottoinsieme X. È facile verificare la proprietà di *scambio delle inclusioni*, cioè

$$X \subseteq Y \quad \Longrightarrow \quad I(Y) \subseteq I(X) \tag{5.1}$$

Poniamo

$$\sqrt{I} := \{f \in A | \exists n \in \mathbf{N} \text{ tale che} f^n \in I\}$$

$\sqrt{I}$ si dice il radicale di I.

Lemma 5.2. *$\sqrt{I}$ è un ideale.*

Dimostrazione. Siano $f \in \sqrt{I}, g \in K[x_1,\ldots,x_n]$. Pertanto $\exists \quad n \in \mathbb{N}$ tale che $f^n \in I$. Quindi $(fg)^n = f^n g^n \in I$, da cui $fg \in \sqrt{I}$. Siano adesso $f, g \in \sqrt{I}$. Possiamo supporre che $\exists \quad n \in \mathbb{N}$ tale che $f^n, g^n \in I$. Utilizzando lo sviluppo del binomio di Newton si verifica che $(f+g)^{2n} \in I$, da cui $f + g \in \sqrt{I}$. □

Abbiamo le ovvie inclusioni:

$$I \subset \sqrt{I} \tag{5.2}$$

$$I \subset J \Rightarrow \sqrt{I} \subset \sqrt{J} \tag{5.3}$$

Un ideale I si dice radicale se $I = \sqrt{(I)}$.

Esercizio 5.3. *Se $X \subset K^n$ provare che $\sqrt{I(X)} = I(X)$, ovvero che $I(X)$ è un ideale radicale.*

Esempi. Se $I = (x^2) \subset K[x]$ allora $\sqrt{I} = (x)$.
Se $I =< xy^2z, x^3w^5 >$ allora $\sqrt{I} =< xyz, xw >$
Se $I =< x^\alpha >_{\alpha \in A}$ è un ideale monomiale, allora $\sqrt{I} =< x^{\alpha'} >$ dove
$\alpha_i{}' = \begin{cases} 1 \text{ se } \alpha_i \neq 0 \\ 0 \text{ se } \alpha_i = 0 \end{cases}$

Esercizio 5.4. *Provare che*

$$\sqrt{IJ} = \sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$$

$$\sqrt{\sqrt{I}} = \sqrt{I}$$

Lemma 5.5. *Se I è un ideale primo, allora $\sqrt{I} = I$.*

Dimostrazione. Sia $f \in \sqrt{I}$ e sia n il minimo intero tale che $f^n \in I$. Considerando che $f^n = f \cdot f^{n-1}$ abbiamo $f \in I$ oppure $f^{n-1} \in I$. Se $n \geq 2$ questa è una contraddizione. Quindi $n = 1$ e $f \in I$. □

L'esempio seguente è particolarmente importante:

Esempio 5.6. *Sia $f\colon K \to K^3$ data da $f(t) = (t, t^2, t^3)$ e poniamo $V := Im\ f \subset K^3$. V si chiama la <u>cubica gobba</u>, ed è parametrizzata dalle espressioni $x = t$, $y = t^2$, $z = t^3$. Adesso vogliamo provare direttamente che se K è infinito allora*

$$I(V) = (y - x^2, z - x^3)$$

cioè che un qualunque polinomio che si annulla su V è combinazione di $y - x^2$ e $z - x^3$. Dato $f \in I(V)$, scegliendo l'ordine Lex con $z > y > x$ ed applicando l'algoritmo di divisione otteniamo $f = (y-x^2)q_1 + (z-x^3)q_2 + r$ dove nessun termine di r è divisibile per $LT(y-x^2) = y$ o per $LT(z - x^3) = z$. Pertanto $r = r(x)$ da cui si ricava

$$0 \equiv f(t, t^2, t^3) = 0 + 0 + r(t)$$

e quindi $r \equiv 0$ come volevamo.

Esercizio 5.7. *Sia $V \subset K^3$ la cubica gobba. Provare che $f = z^2 - x^4y \in I(V)$ e trovare esplicitamente una scrittura come combinazione lineare di $y - x^2$ e $z - x^3$. Ripetere l'esercizio con $f = z - xy$, $f = xz - y^2$ (si veda anche l'esercizio seguente).*

Esercizio 5.8. *Provare che $y - x^2$ e $z - x^3$ costituiscono una base di Gröbner di $I(V)$ secondo l'ordine Lex con $z > y > x$.*

6 Le varietà algebriche affini e la topologia di Zariski su K^n

Definizione 6.1. *Se I è un ideale di $K[x_1, \ldots, x_n]$ poniamo*

$$V(I) := \{a \in K^n | f(a) = 0 \quad \forall f \in I\}$$

$V(I)$ si dice una <u>varietà algebrica affine</u>

Osservazione Notiamo subito che se $I = (f_1, \ldots, f_r)$ allora

$$V(I) = \{a \in K^n | f_1(a) = \ldots f_r(a) = 0\}$$

cioè $V(I)$ coincide con il luogo degli zeri dei polinomi $f_1, \ldots, f_r$.

Esempi. Se I è un ideale principale generato da un polinomio f, scriviamo $V(I) = V(f)$. Queste varietà si chiamano ipersuperfici. Se $deg\ f = 1$ si tratta di varietà lineari, se $deg\ f = 2$ allora $V(f)$ si dice una quadrica. Per il teorema della base di Hilbert e l'osservazione precedente ogni varietà algebrica è intersezione di un numero finito di ipersuperfici. La cubica gobba dell'esempio 5.6 è una varietà algebrica, infatti coincide con $V(I)$ dove $I = (y - x^2, z - x^3)$ (la verifica di questo fatto è immediata).

Osservazione La cubica gobba C é una varietá algebrica affine. Infatti verifichiamo che $C = V(I)$ dove $I = (y - x^2, z - x^3)$. Se $p = (x, y, z) \in C$ allora $\exists t$ tale che $p = (t, t^2, t^3)$ e quindi $p \in V(I)$. Viceversa se $p = (x, y, z) \in V(I)$ allora $y - x^2 = 0$ e $z - x^3 = 0$. Pertanto posto $t := x$ abbiamo $p = (t, t^2, t^3)$.

Approfondiremo lo studio delle varietá descritte da equazioni parametriche nel capitolo 10.

Esercizio 6.2. *Se $I \subseteq J$ sono due ideali, provare che $V(J) \subseteq V(I)$. Questa proprietà di scambio delle inclusioni è analoga alla (5.1).*

Esercizio 6.3. *Provare che $V(I + J) = V(I) \cap V(J)$.*

Esercizio 6.4. *Provare che $V(I) = V(\sqrt{I})$.*

Esercizio 6.5. *Sia $V \subset \mathbb{R}^3$ la curva parametrizzata da (t, t^m, t^n) per $n, m \geq 2$. Provare che V è una varietà affine e calcolare $I(V)$.*

Lemma 6.6.

- *i)* $V((1)) = \emptyset$
- *ii)* $V(0) = K^n$
- *iii)* $V(f_1, \ldots, f_r) \cap V(g_1, \ldots, g_s) = V(f_1, \ldots, f_r, g_1, \ldots, g_s)$. *In generale* $V(I) \cap V(J) = V(I + J)$ *e* $\cap_{a \in \mathcal{A}} V(I_a) = V(\sum_{a \in \mathcal{A}} I_a)$
- *iv)* $V(f_1, \ldots, f_r) \cup V(g_1, \ldots, g_s) = V((\ldots, f_i g_j, \ldots))$. *In generale* $V(I) \cup V(J) = V(IJ)$.

Quindi le varietà algebriche affini soddisfano gli assiomi degli insiemi chiusi per una topologia su K^n.

Dimostrazione. i), ii), iii) seguono subito dalle definizioni. Per provare iv) notiamo che $I, J \supset IJ$ e quindi $V(I) \cup V(J) \subset V(IJ)$. Viceversa sia $a \in V(IJ)$. Se per assurdo $a \notin V(I)$ e $a \notin V(J)$ allora $\exists \quad i \in I$ tale che $i(a) \neq 0$ e $\exists \quad j \in J$ tale che $j(a) \neq 0$. Pertanto $ij(a) \neq 0$ che è una contraddizione. □

Definizione 6.7. *La topologia su K^n che ha per chiusi le varietà algebriche affini $V(I)$ si dice topologia di Zariski.*

Esempio 6.8. *I chiusi della topologia di Zariski in K sono gli insiemi finiti. Se K è infinito questa topologia è $T1$ ma non di Hausdorff.*

Per rendersi conto di quanta cautela occorre lavorando con la topologia di Zariski, il lettore può verificare che l'applicazione somma $s\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ definita da $s(x, y) = x + y$ non è continua se definiamo su $\mathbb{R}$ la topologia di Zariski e nel dominio di s la topologia prodotto. L'applicazione s diventa continua se consideriamo la topologia di Zariski anche sul dominio $\mathbb{R}^2$.

Proposizione 6.9.

$$V(I) \cup V(J) = V(I \cap J)$$

Dimostrazione.

"$\subset$" Abbiamo $I, J \supset I \cap J$ da cui $V(I) \cup V(J) \subset V(I \cap J)$

"$\supset$" Abbiamo $IJ \subset I \cap J$ da cui $V(IJ) \supset V(I \cap J)$. A questo punto è sufficiente utilizzare il lemma 6.6 iv). □

Esercizio 6.10. *Si trovi $I \subset \mathbb{R}[x, y, z]$ tale che $V(I) \subset \mathbb{R}^3$ consiste nell'unione del piano $\{z = 0\}$ con l'asse delle z.*

Proposizione 6.11. *Se J è un ideale di $K[x_1, \ldots, x_n]$ vale*

$$J \subset I(V(J))$$

Dimostrazione. Se $j \in J$ allora $j(x) = 0 \quad \forall x \in V(J)$. □

L'inclusione della Prop. 6.11 può essere stretta, come mostrano i due esempi:

$$J = (x^2) \subset \mathbb{R}[x] \quad \Rightarrow \quad I(V(J)) = I(\text{origine}) = (x) \supsetneq (x^2) = J \tag{6.1}$$

$$J = (x^2 + 1) \subset \mathbb{R}[x] \quad \Rightarrow \quad I(V(J)) = I(\emptyset) = (1) \supsetneq (x^2 + 1) = J \tag{6.2}$$

Il primo esempio (6.1) porta a considerare che:

Lemma 6.12. *Se J è un ideale di $K[x_1, \ldots, x_n]$ vale $\sqrt{J} \subset I(V(J))$.*

Dimostrazione. Dalla Prop. 6.11 $J \subset I(V(J))$. Basta applicare (5.3) ed il fatto che $\sqrt{I(V(J))} = I(V(J))$ (eserc. 5.3). □

Il secondo esempio (6.2) è di natura diversa da (6.1) ed è legato al fatto che $\mathbb{R}$ non è algebricamente chiuso. Infatti vale il

Teorema 6.13. *Teorema degli zeri di Hilbert (HilbertNullStellenSatz). Sia K un campo algebricamente chiuso e sia J un ideale di $K[x_1, \ldots, x_n]$. Allora*

$$\sqrt{J} = I(V(J))$$

Dimostreremo nel capitolo 8 il teorema degli zeri di Hilbert.

Proposizione 6.14. *Vale $X \subset V(I(X))$.*

Dimostrazione. Se $a \in X$ allora $f(a) = 0 \ \forall f \in I(X)$. □

Lemma 6.15. *Se W è una varietà algebrica affine allora $W = V(I(W))$*

Dimostrazione. Sia $W = V(J)$. Abbiamo $J \subset I(W)$ per la Prop. 6.11. Utilizzando l'esercizio 6.2 segue $W = V(J) \supset V(I(W))$. L'altra inclusione è stata vista nella Prop. 6.14. □

Più precisamente abbiamo la

Proposizione 6.16. *Se $S \subset K^n$ è un sottoinsieme allora $V(I(S)) = \overline{S}$ (chiusura secondo la topologia di Zariski)*

Dimostrazione. Abbiamo già visto nella Prop. 6.14 che $S \subset V(I(S))$ e quindi $\overline{S} \subset V(I(S))$ perché $V(I(S))$ è chiuso. Viceversa consideriamo che $I(\overline{S}) \subset I(S)$ e quindi $V(I(S)) \subset V(I(\overline{S})) = \overline{S}$ per il lemma 6.15. □

Definizione 6.17. *Una varietà algebrica affine $V \subset K^n$ si dice riducibile se $V = V_1 \cup V_2$ con V_i sottovarietà proprie. Altrimenti si dice irriducibile.*

Teorema 6.18. *Sia V una varietà algebrica affine*

$$V \text{è irriducibile} \iff I(V) \text{è primo}$$

Dimostrazione.

$\Rightarrow$ Sia $fg \in I(V)$ e poniamo $V_1 := V \cap V(f)$, $V_2 := V \cap V(g)$. Se $f \notin I(V)$ allora $V_1 \neq V$. Preso un qualunque $a \in V \setminus V_1$ abbiamo $f(a) \neq 0$ e quindi $g(a) = 0$, cioè $a \in V_2$. Quindi $V = V_1 \cup V_2$ e per l'ipotesi $V = V_2$ da cui $V \subset V(g)$ e $g \in I(V)$ Segue che $I(V)$ è primo.

$\Leftarrow$ Sia per assurdo $V = V_1 \cup V_2$ con V_i sottovarietà algebriche proprie. Pertanto esistono $f \in I(V_1) \setminus I(V)$ e $g \in I(V_2) \setminus I(V)$ da cui fg si annulla su $V_1 \cup V_2 = V$. Quindi $fg \in I(V)$ e per l'ipotesi $f \in I(V)$ oppure $g \in I(V)$, che è una contraddizione. □

Se per il momento diamo per buono il Teorema degli Zeri di Hilbert 6.13, otteniamo il

Corollario 6.19. *Sia K algebricamente chiuso. C'é una corrispondenza biunivoca naturale tra varietá algebriche ed ideali radicali di $K[x_1, \ldots, x_n]$ data da $W \mapsto I(W)$ con inversa $J \mapsto V(J)$. La corrispondenza porta varietá algebriche irriducibili in ideali primi e viceversa.*

Dimostrazione. La prima parte dell'enunciato segue direttamente dal teor.6.18. Se W é una varietá irriducibile allora $I(W)$ é primo per il teor.6.18 e $V(I(W)) = W$ per il lemma 6.15. Se J é un ideale primo $I(V(J)) = \sqrt{J} = J$ per il teorema 6.13 ed il lemma 6.1. Quindi $V(J)$ é irriducibile per il teorema 6.18. □

Osservazione Se I è primo allora $V(I)$ è irriducibile, usando il Teorema 6.18 e il Teorema degli Zeri. Il viceversa è vero se I è radicale e se $K = \overline{K}$. Su $K = \mathbb{R}$ il viceversa è falso, un controesempio è $I = \big(x \cdot (x^2 + 1)\big)$, che è radicale ma non primo, mentre $V(I)$ è irriducibile.

Esercizi.

- i) Sia f un monomio. Provare che $V(f)$ é dato dall'unione di sottovarietá lineari di codimensione 1.
- ii) Descrivere $V(I)$ dove $I = (xy, xz) \subset K[x, y, z]$
- iii) Sia I un ideale monomiale. Provare che $V(I)$ é dato dall'unione di sottovarietá lineari.

7 Il risultante

Teorema 7.1. *Sia R un UFD. Siano F, G polinomi in $R[x]$ di gradi rispettivamente $f, g > 0$.*

$$\begin{array}{c} F, G \quad \textit{hanno un fattore irriducibile} \\ \textit{di grado positivo in comune} \end{array} \iff \exists \quad \begin{array}{c} A, B \quad \textit{di gradi risp. } g-1, f-1 \\ \textit{tali che } AF + BG = 0 \end{array}$$

Dimostrazione. $\Longrightarrow$ Sia $F = af_1$, $G = ag_1$. Poniamo $A := g_1$, $B := -f_1$. Allora abbiamo $Af + Bg = g_1 af_1 - f_1 ag_1 = 0$. Se *deg* A, *deg* B sono minori di quanto è richiesto basta moltiplicare A e B per lo stesso fattore.

$\Longleftarrow$ Per ipotesi $AF = -BG$. Quindi ogni fattore irriducibile di G divide A oppure F. Siccome *deg* $A = g-1$ allora esiste un fattore irriducibile di G che divide F, come volevamo. □

L'introduzione del risultante è adesso semplice. Il problema è di trovare condizioni per cui due polinomi F, $G \in R[x]$ hanno un fattore in comune.

Si considera come incognite i coefficienti dei due polinomi:

$A := a_0x^{g-1} + \ldots + a_{g-2}x + a_{g-1}$, $B := b_0x^{f-1} + \ldots + b_{f-2}x + b_{f-1}$

e si pone la condizione

$$AF + BG = 0. \tag{7.1}$$

Questo è un sistema lineare con $f+g$ incognite e $f+g$ equazioni. Il determinante della matrice del sistema (7.1) *è per definizione il risultante di F e G.*

Posto $F := f_0x^f + \ldots$, $G := g_0x^g + \ldots$ il sistema (7.1) diventa:

$$\begin{array}{lllll}
a_0f_0+ & & b_0g_0 & & = 0 \quad \text{coeff. di } x^{f+g-1} \\
a_0f_1 + a_1f_0 & & b_0g_1 + b_1g_0 & & = 0 \quad \text{coeff. di } x^{f+g-2} \\
\vdots & & \vdots & \vdots & \\
 & a_{g-1}f_f+ & & b_{f-1}g_g & = 0 \qquad \text{coeff. di} x^0
\end{array}$$

e la sua matrice è

$$\begin{pmatrix}
f_0 & & & g_0 & & & \\
f_1 & f_0 & & g_1 & g_0 & & \\
\vdots & & \ddots & \vdots & & \ddots & \\
f_f & & f_0 & g_g & & & g_0 \\
 & \ddots & & & \ddots & & \\
 & & f_f & & & & g_g
\end{pmatrix}$$

Per comodità di scrittura si scrive il risultante come il determinante della matrice trasposta, cioè si pone:

Definizione 7.2.

$$Res(f,g,x) := \det \begin{vmatrix}
f_0 & f_1 & \cdots & f_f & & & & \\
 & f_0 & \ddots & & \ddots & & & \\
 & & \ddots & \ddots & & \ddots & & \\
 & & & f_0 & f_1 & \cdots & f_f & \\
g_0 & g_1 & \cdots & \cdots & g_g & & & \\
 & g_0 & \ddots & & & \ddots & & \\
 & & \ddots & \ddots & & & \ddots & \\
 & & & \ddots & \ddots & & & \ddots \\
 & & & & g_0 & g_1 & & g_g
\end{vmatrix}$$

Teorema 7.3. *Sia R un UFD.*

$$f, g \in R[x] \quad \textit{hanno un fattore in comune di grado} \geq 1 \iff Res(f,g,x) = 0$$

Dimostrazione.

$\Longrightarrow$ Se fosse $Res(f,g,x) \neq 0$ allora consideriamo il sistema (7.1) nel campo dei quozienti di R. Dalla teoria dei sistemi lineari l'unica soluzione di (7.1) è quella nulla.

$\Longleftarrow$ Nel campo dei quozienti esiste una soluzione di (7.1) non nulla. Moltiplicando per il denominatore comune si trova una soluzione a coefficienti in R. $\square$

Esempio 7.4. *Siano*

$$F = x^2 - 4x + 3$$
$$G = x^2 - 6x + 5$$

che hanno a comune il fattore $x - 1$. *Infatti*

$$Res(f, g, x) = \begin{vmatrix} 1 & -4 & 3 & \\ & 1 & -4 & 3 \\ 1 & -6 & 5 & \\ & 1 & -6 & 5 \end{vmatrix} = 0$$

Il risultante di due polinomi F e G rispetto alla variabile x è implementato in Macaulay2 con il comando resultant(F, G, x), mentre la matrice che appare in 7.2 può essere ricavata con sylvesterMatrix(F, G, x).

Esercizio 7.5. *Verificare se i polinomi* $x^5 + x + 1$ *e* $x^4 + x^3 + 1$ *hanno una radice in comune.*

Definizione 7.6. $Discr(f) := Res(f, f', x)$ *si dice il discriminante di* f. *È implementato in Macaulay2 col comando* discriminant(f, x).

Il discriminante del polinomio monico $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$ é nullo se e solo se f ha una radice doppia (nel proprio campo di spezzamento).

Esempi. Se $f = ax^2 + bx + c$ allora $Discr(f) = a(4ac - b^2)$. Se $f = x^3 + px + q$ allora $Discr(f) = -(4p^3 + 27q^2)$

Teorema 7.7. *Dati* $p, q \in R[x]$ *esistono due polinomi* $A, B \in R[x]$ *tali che* $Ap + Bq = Res(p, q, x)$

Dimostrazione. Se $Res(p, q, x) = 0$ la tesi è ovvia prendendo $A = q$ e $B = -p$. Se $Res(p, q, x) \neq 0$ scriviamo il sistema lineare (analogo di (7.1)) $\tilde{A}p + \tilde{B}q = 1$ con incognite $\tilde{A}$, $\tilde{B}$. Si trova un sistema lineare quadrato di ordine $deg\ p + deg\ q$ con termine noto $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ e determinante della matrice dei coefficienti $Res(p, q, x) \neq 0$. Risolvendo il sistema con la regola di Cramer nel campo dei quozienti di R troviamo $\tilde{A}$ e $\tilde{B}$ soluzioni che hanno a denominatore $Res(p, q, x)$. Posto $A = \tilde{A}Res(p, q, x)$ e $B = \tilde{B}Res(p, q, x)$ si ottiene la tesi. □

Teorema 7.8. *Siano* $f, g \in K[x_1, \ldots x_n]$ *di grado positivo in* x_1. *Allora*

- *i)* $Res(f, g, x_1) \in (f, g) \cap K[x_2, \ldots, x_n]$ *(che si dice il primo ideale di eliminazione)*
- *ii)* $Res(f, g, x_1) \equiv 0 \iff f, g$ *hanno un fattore a comune di grado positivo in* x_1

Dimostrazione. Consideriamo $f, g \in K[x_2, \ldots x_n][x_1]$. Allora per il teorema 7.7 esistono due polinomi $A, B \in K[x_2, \ldots x_n][x_1]$ tali che $Af + Bg = Res(f, g, x_1)$. Quindi $Res(f, g, x_1) \in (f, g) \cap K[x_2, \ldots, x_n]$.

La seconda parte è esattamente il teorema 7.3 □

Osservazione critica Il risultante di due polinomi in piú variabili rispetto ad x_1 calcolato per un valore assegnato delle variabili rimanenti puó essere diverso dal risultante che si ottiene sostituendo subito i valori assegnati. Infatti nel secondo caso i gradi possono diminuire. Questa osservazione é cruciale per la comprensione della dimostrazione del teorema di estensione del

prossimo capitolo. Ad esempio se $F(x_1,x_2)=(x_2-5)x_1+6$ e $G(x_1,x_2)=x_1^2x_2^2-(x_2-2)x_1+5x_2$ allora $Res(F(x_1,0),G(x_1,0),x_1)=-12$ e $Res(F,G,x_1)_{|\{x_2=0\}}=60$

Dal teorema 7.8 i) sorge spontanea la domanda se, dati $f,g\in K[x,y]$, il polinomio in y $Res(f,g,x)$ è il generatore dell'ideale principale $(f,g)\cap K[y]$. L' esempio seguente (e il successivo esempio 7.9) dà una risposta negativa.

Esempio. Siano $f=x^2+y^2-1$, $g=x^3+y^3-1$, ci si propone di calcolare $(f,g)\cap K[y]$ cioè il primo ideale di eliminazione di f e g. Il procedimento "intuitivo" per eliminare la x è il seguente:

$$x^2=1-y^2 \qquad \%I$$
$$x^3=1-y^3 \qquad \%I$$

e quindi $(1-y^2)^3-(1-y^3)^2\in I$. Infatti esplicitamente

$$y^2(y-1)^2(2y^2-4y+3)=(1-y^2)^3-(1-y^3)^2=\left[(1-y^2)^3-x^6\right]-\left[(1-y^3)^2-x^6\right].$$

Sviluppando la prima parentesi come differenza di 2 cubi e la seconda come differenza di 2 quadrati otteniamo che l'espressione precedente è uguale a:

$$\left[(1-y^2)-x^2\right]\left[(1-y^2)^2+x^2(1-y^2)+x^4\right]-\left[(1-y^3)-x^3\right]\left[(1-y^3)+x^3\right]=$$

$$=f\left[-(1-y^2)^2-x^2(1-y^2)-x^4\right]+g[1-y^3+x^3]$$

L'espressione precedente coincide con $Res(f,g,x)$ a meno del segno. $(1-y^2)^3-(1-y^3)^2$ appartiene al primo ideale di eliminazione di f e g ma non è il generatore. Infatti, posto $p(y):=y^2(y-1)(2y^2-4y+3)=\frac{(1-y^2)^3-(1-y^3)^2}{1-y}$ si trova

$$(f,g)\cap K[y]=(p(y)) \tag{7.2}$$

quando non è affatto evidente dalle espressioni precedenti che $p(y)\in(f,g)$. Notiamo che, a meno della molteplicità, le radici del generatore $p(y)$ dell'ideale di eliminazione sono le stesse di quelle di $Res(f,g,x)$, questo può essere falso in generale, si veda l'esempio 7.9.

Per provare (7.2) consideriamo l'ordine Lex con $x>y$ ed utilizziamo l'algoritmo di Buchberger. Abbiamo:

$$S(f,g)=xf-g=xy^2-x-y^3+1=:c(x,y)$$
$$S(f,c)\%\{f,g,c\}=(y^2f-xc)\%\{f,g,c\}=y^2f-xc-f-yc=$$
$$=xy-x+2y^4-y^2-y+1=:d(x,y)$$
$$S(c,d)\%\{f,g,c,d\}=(c-yd)\%\{f,g,c,d\}=c-yd-d=:-p(y)$$

Risostituendo:

$$p(y)=-c+(y+1)d=-c+(y+1)\left(y^2f-(x+y)c-f\right)=$$
$$=-(xf-g)+(y+1)\left(y^2f-(x+y)(xf-g)-f\right)=$$
$$=f\left[-x+(y+1)(y^2-x^2-xy-1)\right]+g\left[1+(1+y)(x+y)\right]$$

ed adesso è facile verificare che $p(y)$ genera $(f,g)\cap K[y]$.

Calcolando tutti i resti delle restanti S-coppie ed eliminando gli elementi superflui si determina una base di Gröbner ridotta per (f,g) che è data da $\{f(x,y),d(x,y),p(y)\}$. Si può quindi applicare anche il teor. di eliminazione 4.2.

Osservazione L'esempio precedente illustra il fatto generale che il calcolo con l'algoritmo di Buchberger di un base di Gröbner a partire da un insieme di generatori dà anche le espressioni degli elementi della base di Gröbner come combinazione dei generatori. Questo algoritmo è analogo all' algoritmo euclideo, con cui dati due interi a, b si determina $MCD(a,b) = d$ e si trovano contemporaneamente due interi x, y tali che $d = ax + by$. (si veda ad esempio [Chi], I, cap. 3)

Esempio 7.9. *Un altro esempio analogo al precedente (ma più semplice) si ha considerando $f = xy - 2$, $g = xy - 1$. In questo caso $Res(f, g, x) = y$ mentre $(f, g) \cap K[y] = (1)$. Notiamo che abbiamo anche $Res(f, g, y) = x$.*

Esercizio 7.10. *Sia $p = x + y - 1$, $q = x^2 + y^2/4 - 1$, poniamo $I = (p, q) \subset K[x, y]$. Provare che*

1. *$V(I)$ consiste dei due punti $(-3, 5, 8/5), (1, 0)$.*
2. *$Res(p, q, x) = y(y - 8/5)$ (a meno di scalari).*
3. *L'ideale di eliminazione I_1 è generato da $y(y - 8/5)$. Suggerimento. Sia $\alpha(y)$ il generatore dell'ideale di eliminazione. Per il Teorema 7.7 esistono A, B tali che $\alpha(y) = A(x, y)p(x, y) + B(x, y)q(x, y)$. Si sostituisca alle indeterminate rispettivamente i valori dei due punti di $V(I)$.*

8 Il teorema di estensione e la dimostrazione del Teorema degli Zeri (NullStellenSatz)

Notiamo che dati due qualunque polinomi f, $g \in K[x, y]$, i punti di coordinate (x, y) appartenenti a $V(f, g)$ hanno la seconda coordinata che annulla ogni polinomio $q(y)$ nell'ideale di eliminazione. È interessante chiedersi il viceversa, cioè ogni radice y_0 del polinomio generatore dell'ideale di eliminazione corrisponde a qualche $(x_0, y_0) \in V(f, g)$? Nell'esempio precedente la risposta è affermativa ma aumentando il numero delle variabili ci vogliono delle ipotesi opportune. Questo problema va sotto il nome di problema di estensione delle soluzioni.

Lemma 8.1. *Sia $I \subseteq K[x_1, \ldots, x_n]$ e I_1 il primo ideale di eliminazione. Sia π_1 la proiezione sulle ultime $(n - 1)$ indeterminate. Allora $\pi_1(V(I)) \subseteq V(I_1)$*

Dimostrazione. Sia $(a_2, \ldots, a_n) \in \pi_1(V(I))$. Pertanto esiste $a_1 \in K$ tale che $(a_1, a_2, \ldots, a_n) \in V(I)$. Se $f \in I_1$ abbiamo $f(a_1, a_2, \ldots, a_n) = f(a_2, \ldots, a_n) = 0$ da cui $(a_2, \ldots, a_n) \in V(I_1)$ come volevamo. □

L'uguaglianza può non valere nel Lemma 8.1, come è mostrato dal seguente esempio in tre variabili.

Esempio 8.2. *Siano $f := xy - 1, g := xz - 1 \in K[x, y, z]$ Eliminando la x troviamo $y - z = -yg + zf$ che è un generatore del primo ideale di eliminazione. Preso il punto di coordinate $(y, z) = (a, a)$ questo si estende a $(\frac{1}{a}, a, a) \in V(f, g)$ se $a \neq 0$ ma se $a = 0$ la soluzione non si estende! Il motivo è che il coefficiente di x si annulla per $(a, a) = (0, 0)$. Geometricamente la soluzione è andata all'infinito, vedremo infatti che nel proiettivo l'eliminazione è più semplice da trattare. In questo esempio, posto $I = (f, g)$, abbiamo $V(I) = \{(\frac{1}{a}, a, a) | a \neq 0\}$, $\pi(V(I)) = \{(a, a) | a \neq 0\} \subsetneq \{(a, a)\} = V(I_1)$.*

Teorema 8.3. *[Teorema di estensione, (Teorema fondamentale della teoria dell'eliminazione, caso affine)]*

Sia K un campo algebricamente chiuso. Siano

$$f_1 := g_1(x_2, \ldots, x_n)x_1^{N_1} + \ldots$$

$$\vdots$$

$$f_k := g_k(x_2, \ldots, x_n)x_1^{N_k} + \ldots$$

polinomi in $K[x_1, \ldots, x_n]$ e sia $I = (f_1, \ldots, f_k)$. Posto $I_1 := I \cap K[x_2, \ldots, x_n]$, sia $(a_2, \ldots, a_n) \in V(I_1)$ "soluzione parziale". Se $(a_2, \ldots, a_n) \notin V(g_1, \ldots, g_k)$ allora $\exists \quad a_1 \in K$ tale che $(a_1, a_2, \ldots, a_n) \in V(I)$

Dimostrazione. Possiamo assumere (rinumerando eventualmente $f_1, \ldots, f_k$) che $g_1(a_2, \ldots, a_n) \neq 0$. Introduciamo delle nuove indeterminate $u_2, \ldots, u_k$ e consideriamo

$$Res(f_1, u_2f_2 + \ldots + u_kf_k, x_1) = Af_1 + B(u_2f_2 + \ldots + u_kf_k) = \sum h_\alpha u^\alpha$$

$$\text{dove} \quad u^\alpha = u_2^{\alpha_2} \ldots u_k^{\alpha_k} \qquad A, B \in K[u_2, \ldots, u_k, x_1, \ldots, x_n] \qquad h_\alpha \in K[x_2, \ldots, x_n]$$

Sviluppando l'ultima uguaglianza si ottiene $h_\alpha \in I_1$ e quindi $h_\alpha(a_2, \ldots, a_n) = 0 \quad \forall \alpha$. Sostituendo eventualmente f_2 con $\tilde{f_2} := f_2 + x_1^N f_1$ ($N \gg 0$) possiamo supporre che $g_2(a_2, \ldots, a_n) \neq 0$ e che f_2 ha grado in x_1 maggiore di $f_3, \ldots, f_n$. Lavoriamo adesso in $K[x_1, u_2, \ldots, u_k]$ sostituendo $(x_2, \ldots, x_n) = (a_2, \ldots, a_n)$. Allora

$$Res(f_1, u_2f_2 + \ldots + u_kf_k, x_1)_{|(x_2,\ldots,x_n)=(a_2,\ldots,a_n)} = 0 \tag{8.1}$$

Siccome i leading term in x_1 di f_1 e di $u_2f_2 + \ldots + u_kf_k$ non si annullano quando sostituisco $(x_2, \ldots, x_n) = (a_2, \ldots, a_n)$ posso dire che l'espressione (8.1) coincide con il risultante tra $f_{1|(x_2,\ldots,x_n)=(a_2,\ldots,a_n)}$ e $u_2f_2 + \ldots + u_kf_{k|(x_2,\ldots,x_n)=(a_2,\ldots,a_n)}$ nell'anello $K[x_1, u_2, \ldots, u_k]$ (si veda l'oss. critica 7). Per il teorema 7.8 ii) segue che $f_1(x_1, a_2, \ldots, a_n)$ e $(u_2f_2 + \ldots + u_kf_k)_{|(x_2,\ldots,x_n)=(a_2,\ldots,a_n)}$ in $K[x_1, u_2, \ldots, u_k]$ hanno a comune un fattore F di grado positivo in x_1. Siccome $F|f_1(x_1, a_2, \ldots, a_n)$ abbiamo $F \in K[x_1]$ e quindi F divide anche $f_2(x_1, a_2, \ldots, a_n), \ldots, f_k(x_1, a_2, \ldots, a_n)$. Per ipotesi K è algebricamente chiuso, quindi esiste $a_1 \in K$ tale che $F(a_1) = 0$ da cui $f_i(a_1, a_2, \ldots, a_n) = 0$ e quindi $(a_1, a_2, \ldots, a_n) \in V(I)$ come volevamo. □

Vediamo ora come utilizzare il teorema di estensione per provare il teorema degli zeri di Hilbert enunciato in 6.13

Come passo intermedio, importante di per sé, si prova che il teorema degli zeri equivale ad una versione "debole" (qui il teorema di estensione ancora non interviene). La versione debole segue poi facilmente dal teorema di estensione.

Teorema 8.4. *Nullstellensatz debole. Sia K un campo algebricamente chiuso e I un ideale di $K[x_1, \ldots, x_n]$. Abbiamo*

$$V(I) = \emptyset \iff I = K[x_1, \ldots, x_n]$$

Proposizione 8.5.

$$\textit{Nullstellensatz} \iff \textit{Nullstellensatz debole}$$

Dimostrazione.

" $\Longrightarrow$ " Sia $V(I) = \emptyset$, allora per il Nullstellensatz $\sqrt{I} = (1)$ da cui $I = (1)$. Il viceversa è evidente.

" $\Longleftarrow$ " Sia $f \in I(V(f_1, \ldots, f_s))$. Vogliamo provare che esiste m tale che $f^m \in (f_1, \ldots, f_s)$. Sia $\tilde{I} := (f_1, \ldots, f_s, 1 - yf) \subset K[x_1, \ldots x_n, y]$ (*Rabinowitsch trick*). Affermiamo che $V(\tilde{I}) = \emptyset$. Se per assurdo esiste $P_0 := (a_1, \ldots, a_n, y_0) \in V(\tilde{I})$ allora in particolare $f_i(a_1, \ldots, a_n) = 0$, e quindi $(a_1, \ldots, a_n) \in V(f_1, \ldots, f_s)$ da cui $f(a_1, \ldots, a_n) = 0$. Pertanto $1 - yf$ vale 1 nel punto P_0 e questa è una contraddizione.

Per il Nullstellensatz debole segue $\tilde{I} = K[x_1, \ldots x_n, y]$, da cui

$$1 = \sum p_i(x_1, \ldots, x_n, y) f_i + q(x_1, \ldots, x_n, y)(1 - yf)$$

Sostituendo $y = 1/f$ abbiamo

$$1 = \sum p_i(x_1, \ldots, x_n, \frac{1}{f}) f_i$$

I termini della somma a secondo membro sono funzioni razionali aventi a denominatore qualche potenza di f. Raccogliendo sotto un unico denominatore si ottiene:

$$f^m = \sum \tilde{p}_i(x_1, \ldots, x_n) f_i$$

come volevamo. □

Esercizio 8.6. *Sia* $J = \left(x^2 + y^2 - 1, y - 1\right) \subset \mathbb{R}[x, y]$. *Trovare* $f \in I(V(J))$ *tale che* $f \notin J$.

Dimostrazione del Nullstellensatz debole 8.4

Se $n = 1$ il teorema è vero perché K é algebricamente chiuso, quindi ragioniamo per induzione su n. Se $V(I) = V((f_1, \ldots, f_k)) = \emptyset$ vogliamo provare che $(f_1, \ldots, f_k) = (1)$. Possiamo assumere che

$$f_i(x_1, \ldots, x_n) = c_i x_1^{N_i} + \ldots$$

con $c_i \neq 0$. Infatti consideriamo l'automorfismo ϕ di $K[x_1, \ldots, x_n]$ definito da

$$x_1 \mapsto x_1$$

$$x_2 \mapsto x_2 + a_2 x_1$$

$$\vdots$$

$$x_n \mapsto x_n + a_n x_1$$

con $a_2, \ldots, a_n$ da determinare (l'inversa di ϕ si ottiene cambiando i segni precedenti da $+$ a $-$. Abbiamo $V(\phi(f_1), \ldots, \phi(f_k)) = \emptyset$ perché $\phi(f)(x_1, \ldots, x_n) = f(\phi(x_1), \ldots, \phi(x_n))$ ed ovviamente:

$$I = (1) \Longleftrightarrow \phi(I) = (1)$$

Sia $\phi(x^\alpha) = g_\alpha(a_2, \ldots, a_n) x_1^{\sum \alpha_i} + \ldots$ (termini di grado inferiore in x_1), dove $g_\alpha = x_2^{\alpha_2} \ldots x_n^{\alpha_n}$, e quindi se $f_i = c_i(x_2, \ldots, x_n) x_1^{N_i} + \ldots$ (termini di grado inferiore in x_1) segue $\phi(f_i) = c_i(a_2, \ldots, a_n) x_1^{N_i} + \ldots$. Basta quindi scegliere $a_2, \ldots, a_n$ in modo che $c_i(a_2, \ldots, a_n) \neq 0$ per qualche i (ad esempio $i = 1$). Adesso possiamo applicare il teorema di estensione 8.3 Se fosse $V(I_1) \neq \emptyset$ avrei anche $V(I) \neq \emptyset$ che è una contraddizione. Quindi $V(I_1) = \emptyset$ e per l'ipotesi induttiva $1 \in I_1 \subset I$. □

La dimostrazione del teorema degli zeri 6.13 è così completa.

Lemma 8.7. *La base di Gröbner dell'ideale* (1) *contiene* $\{1\}$ *per ogni ordinamento monomiale.*

Dimostrazione. Dalle proprietà degli ordini monomiali, segue subito che $LT(1) = (1)$. Inoltre se un polinomio f ha 1 come leading term, siccome 1 è minore di qualunque monomio, segue che $f = 1$. Pertanto una base di Groebner per tutto l'anello (1) deve contenere l'elemento 1. □

Dal Nullstellensatz (debole) segue in particolare

Teorema 8.8. *[Algoritmo di consistenza.] Siano* $f_1, \ldots, f_s \in K[x_1, \ldots, x_n]$ *con* K *algebricamente chiuso. Vale:*

$$\begin{array}{c}\textit{Il sistema } f_i(x_1,\ldots,x_n) = 0 \\ \textit{ha una soluzione}\end{array} \iff \begin{array}{c}\textit{La base di Gröbner} \\ \textit{dell'ideale } (f_1,\ldots,f_s) \\ \textit{non contiene } \{1\}\end{array}$$

Teorema 8.9. *Sia* K *algebricamente chiuso. Gli ideali massimali di* $K[x_1, \ldots, x_n]$ *sono tutti e soli quelli della forma* $(x_1 - a_1, \ldots, x_n - a_n)$

Dimostrazione. Abbiamo $\frac{K[x_1,\ldots,x_n]}{(x_1-a_1,\ldots,x_n-a_n)} \simeq K$ e quindi $(x_1 - a_1, \ldots, x_n - a_n)$ è massimale e coincide con l'ideale $I(p)$ dei polinomi che si annullano in $p = (a_1, \ldots, a_n)$. Viceversa sia I un ideale massimale. Dal Nullstellensatz debole 8.4 abbiamo che esiste $(a_1, \ldots, a_n) \in V(I)$. Pertanto

$$I \subset I(V(I)) \subset I(a_1, \ldots a_n) = (x_1 - a_1, \ldots x_n - a_n)$$

e per la massimalità vale l'uguaglianza. □

Corollario 8.10. *Sia* K *algebricamente chiuso e sia* $V \subset K^n$ *una varietá algebrica affine. Allora gli ideali massimali di* $K[x_1, \ldots, x_n]/I(V)$ *sono tutti e soli quelli della forma* $(x_1 - a_1, \ldots, x_n - a_n)$ *con* $(a_1, \ldots, a_n) \in V$.

Dimostrazione. É sufficiente osservare che gli ideali di $K[x_1, \ldots, x_n]/I(V)$ sono in corrispondenza biunivoca con gli ideali di $K[x_1, \ldots, x_n]$ che contengono $I(V)$ e che $I(V) \subset (x_1 - a_1, \ldots x_n - a_n) \iff (a_1, \ldots, a_n) \in V$. □

Esercizi

1. Si consideri il sistema di equazioni

$$x^2 + 2y^2 = 3$$
$$x^2 + xy + y^2 = 3$$

   Se I è l'ideale generato da queste equazioni, si trovino generatori per $I \cap K[x]$ e $I \cap K[y]$. Si trovino tutte le soluzioni del sistema se $K = \mathbb{Q}, \mathbb{R}$ o $\mathbb{C}$.

2. Come nell'esercizio 1. per il sistema

$$x^2 + 2y^2 = 2$$
$$x^2 + xy + y^2 = 2$$

3. Trovare generatori per gli ideali di eliminazione I_1 e I_2 dove I è l'ideale generato da

$$x^2 + y^2 + z^2 = 4$$

$$x^2 + 2y^2 = 5$$

$$xz = 1$$

4. Si consideri il sistema di equazioni

$$x^5 + \frac{1}{x^5} = y$$

$$x + \frac{1}{x} = z$$

Sia I l'ideale in $\mathbb{C}[x, y, z]$ determinato da queste equazioni.

- a. Trovare una base per $I_1 \subset \mathbb{C}[y, z]$ e provare che $I_2 = 0$.
- b. Usare il teorema di estensione 8.3 per provare che ogni soluzione parziale $c \in V(I_2) = \mathbb{C}$ estende ad una soluzione $(x_0, y_0, c) \in V(I)$.
- c.Quali soluzioni parziali $(y, z) \in V(I_1) \subset \mathbb{R}^2$ si estendono a soluzioni in $V(I) \subset \mathbb{R}^3$? Confrontare la risposta con quanto affermato dal teorema di estensione.
- d. Guardando z come "parametro", risolvere il sistema con x, y funzioni razionali di z e trovare così una parametrizzazione di $V(I)$.

5. Siano f, $g \in \mathbb{C}[x, y]$. Questo esercizio è una guida per provare che

$V(f, g)$ è infinito $\iff$ f e g hanno un fattore a comune non costante in $\mathbb{C}[x, y]$

- a. Provare che se f è non costante allora $V(f)$ è infinito (ridursi al teorema fondamentale dell'algebra in una variabile) .
- b. Provare $\Longleftarrow$ utilizzando il punto a.
- c. Provare $\Longrightarrow$ mostrando che se f e g non hanno fattori non costanti a comune allora $Res(f, g, x)$ e $Res(f, g, y)$ sono entrambi non nulli.

6. Sia K algebricamente chiuso e siano $y_1, \ldots, y_k$ tutte le radici di $Res(f, g, x)$ e $x_1, \ldots, x_s$ tutte le radici di $Res(f, g, y)$. Provare che tutte le soluzioni del sistema $\begin{cases} f = 0 \\ g = 0 \end{cases}$ sono contenute tra le (x_i, y_j) per $i = 1, \ldots s$, $j = 1, \ldots, k$ (è sufficiente quindi eseguire un numero finito di verifiche per conoscere tutte le soluzioni)

Teorema 8.11. *Teorema di chiusura (Interpretazione geometrica dell'eliminazione).*
Sia $V = V(I) \subset K^n$ una varietà algebrica affine e sia K algebricamente chiuso. Sia $K^n \xrightarrow{\pi_t} K^{n-t}$ la proiezione sulle ultime $n - t$ coordinate. Allora

$$V(I_t) = \overline{\pi_t(V)}$$

Dimostrazione. Intanto notiamo, come nel Lemma 8.1, che $\pi_t(V) \subset V(I_t)$. Infatti sia $(a_1, \ldots, a_n) \in V$ e quindi $(a_{t+1}, \ldots, a_n) \in \pi_t(V)$. Se $f \in I_t$ abbiamo $f(a_{t+1}, \ldots, a_n) = f(a_1, \ldots, a_n) = 0$ e quindi

$$(a_{t+1}, \ldots, a_n) \in V(I_t)$$

Pertanto $V(I_t) \supset \overline{\pi_t(V)}$.

Per l'inclusione opposta è sufficiente provare che

$$I(\pi_t(V)) \subset \sqrt{I_t} \tag{8.2}$$

Se (8.2) è vera allora abbiamo $V(I_t) = V(\sqrt{I_t}) \subset V(I(\pi_t(V))) = \overline{\pi_t(V)}$ (vedi la prop. 6.16) come volevamo.

Per provare (8.2) prendiamo $f \in I(\pi_t(V)) \subset K[x_{t+1}, \ldots, x_n] \subset K[x_1, \ldots, x_n]$. Quindi se $(a_{t+1}, \ldots, a_n) \in \pi_t(V)$ abbiamo $f(a_{t+1}, \ldots, a_n) = 0$. Pertanto $f \in I(V)$ e per il Nullstellensatz $f \in \sqrt{I}$, cioè $\exists\ n$ tale che $f^n \in I \cap K[x_{t+1}, \ldots, x_n] = I_t$, cioè $f \in \sqrt{I_t}$. □

Osservazione Per verificare che l'ipotesi K algebricamente chiuso è necessaria nel teorema di chiusura è sufficiente considerare $I = (x^2 + y^2, 2x^2 + y^2 + 1) \subset \mathbb{R}[x, y]$. Eliminando la x abbiamo $V(I_1) = \{-1, 1\} \subset \mathbb{R}$ mentre $\pi_1(V) = \emptyset$

Possiamo risolvere adesso facilmente il problema di appartenenza di un elemento al radicale $\sqrt{I}$ di un ideale I di $K[x_1, \ldots, x_n]$.

Teorema 8.12. *Sia $I = (f_1, \ldots, f_s)$ un ideale di $K[x_1, \ldots, x_n]$. Allora*

$$f \in \sqrt{I} \Longleftrightarrow 1 \in (f_1, \ldots, f_s, 1 - yf) \subset K[x_1, \ldots, x_n, y]$$

Dimostrazione.

- "$\Longleftarrow$" è stata essenzialmente già vista nella dimostrazione della prop. 8.5 Se abbiamo $1 = \sum p_i(x, y) f_i + g(x, y)(1 - yf)$ ponendo $y = \frac{1}{f}$ e semplificando i denominatori si ottiene la tesi.
- "$\Longrightarrow$" Sia $f^m \in I \subset \tilde{I} := (f_1, \ldots, f_s, 1 - yf)$. Per definizione abbiamo anche $1 - yf \in \tilde{I}$. Pertanto

$$1 = y^m f^m + (1 - y^m f^m) = y^m f^m + (1 - yf)(1 + yf + \ldots + y^{m-1} f^{m-1}) \in \tilde{I}$$

□

Seguendo il Lemma 8.7 ed il Teorema 8.12 ricaviamo un algoritmo per decidere se $f \in \sqrt{I}$ che equivale alla inclusione $V(I) \subset V(f)$.

Il seguente diagramma schematizza come si possono trovare alcune soluzioni di un sistema generico di m equazioni polinomiali in n variabili (con $m \geq n$) se ad ogni passo sono verificate le ipotesi del teorema di estensione.

sistema in n variab.	$\xrightarrow{\text{elim.}}$	sistema in $(n-1)$ variab.	$\xrightarrow{\text{elim.}}$	$\ldots$	$\xrightarrow{\text{elim.}}$	sistema in 1 variab. $\downarrow$
soluz. in n variab.	$\xleftarrow{\text{estens.}}$	$\ldots$	$\xleftarrow{\text{estens.}}$	soluz. in 2 variab.	$\xleftarrow{\text{estens.}}$	soluz. in 1 variab. (magari numerica)

9 Colorabilità di un grafo via basi di Gröbner

Un grafo $G = (V, E)$ è formato da un insieme finito di vertici V e un insieme di lati E che consistono in coppie non ordinate di vertici. Nel contesto della colorabilità, si considerano solo grafi con lati che uniscono vertici *distinti*.

Un grafo si dice k-colorabile se esiste una funzione $f\colon V \to \{0, \ldots, k-1\}$ tale che assume valori distinti per ogni coppia di vertici uniti da un lato.

Ogni grafo è k-colorabile per $k \geq |V|$ (colorando ogni vertice con un colore diverso), la questione interessante è riuscire a k-colorare un grafo con valori di k più piccoli. L'indice cromatico di un grafo G è il minimo k tale che G è k-colorabile.

Lemma 9.1. *Sia $k \geq 2$ e sia $\xi = e^{2\pi\sqrt{-1}/k} \in \overline{\mathbb{Q}}$ una radice primitiva k-esima dell'unità. Per ogni $j = 0, \ldots, k-1$, il polinomio $f_j(x) = \frac{x^k-1}{x-\xi^j}$ di grado $k-1$ ha per radici $x = \xi^i$ per $i = 0, \ldots, k-1$, $i \neq j$.*

Ogni grafo G con n vertici $\{v_1, \ldots, v_n\}$ ammette (per ogni intero $k \geq 2$) il suo ideale k-cromatico $I_k(G) \subseteq \mathbb{Q}[x_1, \ldots, x_n]$ generato dai seguenti polinomi:

- $v_i = x_i^k - 1$ per $i = 1, \ldots, k$ *(polinomi dei vertici)*
- $e_{ij} = \frac{x_i^k - x_j^k}{x_i - x_j} = x_i^{k-1} + x_i^{k-2}x_j + \ldots + x_j^{k-1}$ per ogni coppia (v_i, v_j) di vertici che sono uniti da un lato *(polinomi dei lati).*

Teorema 9.2. *Un grafo è k-colorabile se e solo se l'ideale $I_k(G)$ è diverso da* (1).

Dimostrazione. Se un grafo è k-colorabile mediante una funzione $f\colon V \to \{0, \ldots, k-1\}$ consideriamo la radice k-esima dell'unità $\xi = e^{2\pi\sqrt{-1}/k} \in \overline{\mathbb{Q}}$ e il punto $\left(\xi^{f(1)}, \ldots, \xi^{f(n)}\right) \in V(I_k(G)) \subseteq \overline{\mathbb{Q}}^n$, da cui $I_k(G) \neq (1)$. Viceversa, se $I_k(G) \neq (1)$, per il NullStellenSatz esiste $(a_1, \ldots, a_n) \in V(I_k(G))$. Vale $a_i^k = 1$, pertanto possiamo definire $f\colon V \to \{0, \ldots, k-1\}$ dalla condizione $a_i = \xi^{f(v_i)}$, e ogni altra equazione $e_{ij}(\xi^{f(v_i)}, \xi^{f(v_j)}) = 0$ garantisce che $f(v_i) \neq f(v_j)$ dal Lemma 9.1 (considerando l'equazione $e_{ij}(x, \xi^{f(v_j)}) = 0$). □

Il Teorema 9.2 permette di verificare la k-colorabilità di un grafo G e di determinare il suo indice cromatico mediante il calcolo della base di Gröbner del suo ideale cromatico $I_k(G)$.

Il grado di $I_k(G)$ (ottenibile col comando degree$I_k(G)$), diviso per $k!$, fornisce il numero di k-colorazioni diverse di G.

Una osservazione interessante è che ogni soluzione del gioco *"Sudoku"* corrisponde alla 9-colorabilità di un grafo con 81 vertici x_{ij} per $1 \leq i, j \leq 9$ dove tutti i vertici con stessa riga, o con stessa colonna, sono connessi da un lato, e infine anche i vertici di uno stesso blocco 3×3 sono connessi. Per inserire nell'ideale 9-cromatico i vertici che hanno già una colorazione, si considera che il campo di spezzamento del polinomio ciclotomico $\Phi_9(t) = t^6 + t^3 + 1$ è $\mathbb{Q}[t]/(t^6 + t^3 + 1)$, questo campo coincide col campo ottenuto aggiungendo a $\mathbb{Q}$ una radice primitiva nona dell'unità. Si può lavorare in M2 con il seguente anello

```
K2=toField(QQ[t]/ideal(t^6+t^3+1))
R=K2[x_(0,0)..x_(8,8)]
```

Se la casella x_{ij} è riempita con il colore k, dove $k \in \{1, \ldots, 9\}$, si aggiunge all'ideale cromatico il generatore $x_{ij} - t^{k-1}$, e questo per tutte le condizioni iniziali.

Esercizio 9.3. *Calcolare l'indice cromatico dei grafi dati dalle province della Toscana, del Lazio e della Sardegna, dove due province sono unite da un lato quando sono confinanti.*

10 Parametrizzazioni, varietà razionali e unirazionali

Riflessioni sulle definizioni di varietà algebrica e di varietà differenziabile:
Una varietà differenziabile nonsingolare X di dimensione k in $\mathbf{R}^n$ può essere introdotta in uno dei due modi equivalenti:
i) PARAMETRICO $\forall x \in X$ esiste un intorno aperto $x \in U \subset X$ (con la topologia indotta), un aperto $V \subset \mathbf{R}^k$ ed un'applicazione suriettiva C^∞ $F\colon V \to U \subset \mathbb{R}^n$ di rango k.
ii) IMPLICITO $\forall x \in X$ esiste un intorno aperto $x \in W \subset \mathbf{R}^n$ ed una funzione C^∞ $G\colon W \to \mathbf{R}^{n-k}$ di rango $n-k$ tale che $X \cap W = \{y \in W | G(y) = 0\}$.

La condizione i) dà localmente una parametrizzazione della varietà. La condizione ii) dà invece equazioni implicite. L'equivalenza delle condizioni i) e ii) segue essenzialmente dal teorema della funzione implicita.

Se al posto di C^∞ leggiamo "analitico reale" allora le condizioni i) e ii) sono ancora equivalenti e definiscono una varietà analitica reale.

Le varietà algebriche nascono sostituendo le funzioni C^∞ con le funzioni razionali (con denominatore mai nullo dove sono definite). La condizione ii) definisce allora un aperto di una varietà algebrica affine. In questo caso la condizione i) implica la ii), ed il procedimento con cui si ottiene la funzione G va sotto il nome di eliminazione dei parametri.

La condizione ii) non implica la i). Questo è mostrato dal seguente

Esempio 10.1. *(Le curve di Fermat). Sia C la curva di $\mathbf{R}^2$ data dall'equazione*

$$x^n + y^n - 1 = 0$$

per $n \geq 3$. Allora non esistono funzioni razionali $x = x(t)$, $y = y(t)$ tali che $(x(t), y(t)) \in C$ per t in un aperto di $\mathbb{R}$.

Supponiamo che esistano $x(t) = \frac{p(t)}{r(t)}$, $y(t) = \frac{q(t)}{r(t)}$ come nell'enunciato con p, q, r polinomi senza fattori comuni.

Abbiamo

$$p^n(t) + q^n(t) - r^n(t) = 0$$

da cui p, q, r sono primi tra loro a due a due. Derivando rispetto a t otteniamo la relazione

$$np^{n-1}p' + nq^{n-1}q' - nr^{n-1}r' = 0$$

Le due relazioni precedenti si riassumono nella forma matriciale:

$$\begin{pmatrix} p & q & -r \\ p' & q' & -r' \end{pmatrix} \begin{pmatrix} p^{n-1} \\ q^{n-1} \\ r^{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Quindi il vettore $(p^{n-1}, q^{n-1}, r^{n-1})$ risulta proporzionale al vettore dato dai tre minori 2×2 (a segni alterni) della matrice 2×3 precedente. Si ottengono facilmente le due relazioni

$$\begin{vmatrix} q & -r \\ q' & -r' \end{vmatrix} q^{n-1} = -\begin{vmatrix} p & -r \\ p' & -r' \end{vmatrix} p^{n-1}$$

$$\begin{vmatrix} q & -r \\ q' & -r' \end{vmatrix} r^{n-1} = \begin{vmatrix} p & q \\ p' & q' \end{vmatrix} p^{n-1}$$

da cui le seguenti condizioni di divisibilità:

$$p^{n-1} \Big| \begin{vmatrix} q & -r \\ q' & -r' \end{vmatrix}$$

$$q^{n-1} \Big| \begin{vmatrix} p & -r \\ p' & -r' \end{vmatrix}$$

$$r^{n-1} \Big| \begin{vmatrix} p & q \\ p' & q' \end{vmatrix}$$

Poniamo deg $p = P$, deg $q = Q$, deg $r = R$.

Le tre relazioni precedenti forniscono

$$(n-1)P \leq Q + R - 1$$

$$(n-1)Q \leq P + R - 1$$
$$(n-1)R \leq P + Q - 1$$

e sommando

$$(n-1)(P+Q+R) \leq 2(P+Q+R) - 3$$

che è una contraddizione se $n \geq 3$.

L'esempio delle curve di Fermat puó essere ripetuto parola per parola in ogni campo K con car K che non divide n. Infatti la derivata di un polinomio puó essere definita formalmente in un campo qualunque.

Osservazione. Se $n = 2$ allora

$$x = \frac{1-t^2}{1+t^2} \qquad y = \frac{2t}{1+t^2} \tag{10.1}$$

è una parametrizzazione razionale della conica $x^2 + y^2 - 1 = 0$.

Se $t = \frac{p}{q}$ si ricava (eliminando i denominatori) che $(p^2 - q^2, 2pq, p^2 + q^2)$ sono terne pitagoriche.

Se $n = 1$ si trova la retta $x + y - 1 = 0$ che ammette la parametrizzazione

$$x = t \qquad y = -t + 1$$

Parametrizzazioni polinomiali

Sia $F\colon K^m \to K^n$ una funzione polinomiale definita da $F = (f_1, \ldots, f_n)$ con f_i polinomi in $t_1, \ldots, t_m$.

Abbiamo così una parametrizzazione polinomiale di $Im\ F = F(K^m)$. Il teorema seguente mostra che si possono sempre trovare con un procedimento di eliminazione delle equazioni per la chiusura di Zariski di $F(K^m)$.

Teorema 10.2. *Sia K un campo algebricamente chiuso. Sia $F = (f_1, \ldots, f_n)\colon K^m \to K^n$ una funzione polinomiale. Sia $I = (x_1 - f_1(t_1, \ldots, t_m), \ldots, x_n - f_n(t_1, \ldots, t_m)) \subset K[t_1, \ldots, t_m, x_1, \ldots, x_n]$ e sia $I_m = I \cap K[x_1, \ldots, x_n]$ l'm-esimo ideale di eliminazione. Allora*

$$\overline{F(K^m)} = V(I_m)$$

Dimostrazione. Consideriamo il diagramma

$$\begin{array}{ccccc} V(I) & \subset & K^{n+m} & & \\ & i \nearrow & & \searrow \pi_m & \\ K^m & & \longrightarrow^F & & K^n \end{array} \tag{10.2}$$

dove $i(t_1, \ldots, t_m) := (t_1, \ldots, t_m, f_1(t_1, \ldots, t_m), \ldots, f_n(t_1, \ldots, t_m))$.

È facile verificare che $i(K^m) = V(I)$, che si dice il *grafico* di F. Pertanto $F(K^m) = \pi_m(i(K^m)) = \pi_m(V(I))$ e quindi dal teorema di chiusura 8.11 $\overline{F(K^m)} = V(I_m)$ come volevamo. □

Il teorema 10.2 è vero se K è un qualunque campo infinito, anche non algebricamente chiuso (ad esempio $\mathbb{R}$). Naturalmente non è possibile applicare il teorema di chiusura e quindi la dimostrazione va modificata nel modo seguente. Sia $K \subset \overline{K}$ (chiusura algebrica). Come sopra abbiamo $F(K^m) = \pi_m(V(I)) \subset V(I_m)$ e quindi $\overline{F(K^m)} \subset V(I_m)$. Sia adesso $V_K(g_1, \ldots, g_s) = \overline{F(K^m)}$. Quindi $g_i \circ F \equiv 0$ come polinomi in $t_1, \ldots, t_m$. Il principio di identità dei polinomi vale su qualunque campo infinito, quindi g_i si annullano su $F(\overline{K}^m)$, da cui $V_{\overline{K}}(g_1, \ldots, g_s) \supset \overline{F(\overline{K}^m)} = V_{\overline{K}}(I_m) \supset V_K(I_m)$ (l' uguaglianza per il teorema 10.2) . Quindi

$$\overline{F(K^m)} = V_K(g_1, \ldots, g_s) \supset V_K(I_m)$$

da cui $V_K(I_m) \subset \overline{F(K^m)}$ come volevamo.

Il teorema 10.2 mostra che eseguendo l'eliminazione da $(x_1 - f_1, \ldots, x_n - f_n)$ si trova la più piccola varietà contenente $F(K^m)$. In generale $F(K^m)$ può essere strettamente contenuto nella sua chiusura (anche se K é algebricamente chiuso). Ad esempio si prenda $F\colon K^2 \to K^2$ definita da $F(x, y) = (xy, y)$. $Im\ F$ é uguale a $\{y \neq 0\} \cup \{(0, 0)\}$ e $\overline{Im\ F} = K^2$. $Im\ F$ é un tipico esempio di insieme costruibile, come vedremo nel §16.

Invece con le notazioni del teorema 10.2 $F(K)$ (corrispondente a $m = 1$) é chiuso. Cioé vale la

Proposizione 10.3. *Sia K algebricamente chiuso. Sia $F = (f_1, \ldots, f_n)\colon K \to K^n$ una funzione polinomiale. Sia $I = (x_1 - f_1(t), \ldots, x_n - f_n(t)) \subset K[t, x_1, \ldots, x_n]$. Allora*

$$F(K) = \overline{F(K)} = V(I_1)$$

Dimostrazione. Seguendo la dimostrazione del teorema 10.2 abbiamo $F(K) = \pi_1(V(I))$. Adesso se $(x_1, \ldots, x_n) \in V(I_1)$ dal teorema di estensione 8.3 esiste t tale che $(t, x_1, \ldots, x_n) \in V(I)$ e quindi $(x_1, \ldots, x_n) \in \pi_1(V(I))$. Pertanto $V(I_1) = \pi_1(V(I)) = F(K)$ e quindi $F(K)$ é chiuso. □

Esercizio 10.4. *Trovare un esempio di una parametrizzazione polinomiale F dove K non é algebricamente chiuso e $F(K)$ non é chiuso.*

Suggerimento: $F(x) = x^2$ su $\mathbb{R}$.

Parametrizzazioni razionali

Esempio 10.5. *Consideriamo $x = \frac{u^2}{v}$, $y = \frac{v^2}{u}$, $z = u$. Si vede subito che l'immagine di questa parametrizzazione F (che è definita per $u \neq 0$, $v \neq 0$) è contenuta in $x^2y - z^3 = 0$ Eliminando i denominatori abbiamo l'ideale*

$$I = (vx - u^2, uy - v^2, z - u) \subset K[u, v, x, y, z]$$

Eliminando u, v si trova $I_2 = \big(z(x^2y - z^3)\big)$. Pertanto in questo caso

$$V(I_2) = V(x^2y - z^3) \cup V(z) \supset_{\neq} V(x^2y - z^3) \supset \overline{F(K^2)}$$

Questo esempio mostra che il teorema 10.2 non puó essere generalizzato al caso di parametrizzazioni razionali semplicemente eliminando i denominatori.

Nel caso generale abbiamo

$$\begin{cases} x_1 = \frac{f_1(t_1, \ldots, t_m)}{g_1(t_1, \ldots, t_m)} \\ \qquad \vdots \\ x_n = \frac{f_n(t_1, \ldots, t_m)}{g_n(t_1, \ldots, t_m)} \end{cases} \tag{10.3}$$

dove f_i, g_i sono polinomi. In questo caso si dice che abbiamo una parametrizzazione razionale, che è definita dove $g_i \neq 0$. Precisamente, posto $W = V(g_1 \cdots g_n) \subset K^m$ le 10.2 definiscono

$$F\colon K^m \setminus W \to K^n \tag{10.4}$$

dove $F = (\frac{f_1}{g_1} \ldots \frac{f_n}{g_n})$. Consideriamo il diagramma (analogo a 10.2)

$$\begin{array}{ccccc} V(I) & \subset & K^{n+m} & & \\ & i \nearrow & & \searrow \pi_m & \\ K^m \setminus W & & \longrightarrow^F & & K^n \end{array}$$

Se $I := (g_1x_1 - f_1, \ldots, g_nx_n - f_n)$ allora l'esempio 10.5 mostra che può essere

$$i(K^m \setminus W) \subset_{\neq} V(I)$$

e quindi non si può ripetere la costruzione del teor. 10.2

Per trattare questo problema l'approccio giusto sta nel definire $g := g_1g_2 \cdots g_n$ e considerare il diagramma

$$\begin{array}{ccccc} W = V(J) & \subset & K^{n+m+1} & & \\ & j \nearrow & & \searrow \pi_{m+1} & \\ K^m \setminus W & & \longrightarrow^F & & K^n \end{array}$$

dove $j(t_1, \ldots, t_m) = (\frac{1}{g(t_1,\ldots,t_m)}, t_1, \ldots, t_m, \frac{f_1}{g_1}, \ldots, \frac{f_n}{g_n})$ e dove

$$J := (g_1x_1 - f_1, \ldots, g_nx_n - f_n, 1 - gy) \subset K[y, t_1, \ldots, t_m, x_1, \ldots, x_n]$$

Lemma 10.6.

$$j(K^m \setminus W) = V(J)$$

Dimostrazione.

- "$\subset$" È ovvia dalle definizioni
- "$\supset$" Se $(y, t_1, \ldots, t_m, x_1, \ldots, x_n) \in V(J)$ allora abbiamo $g(t_1, \ldots, t_m)y = 1$ da cui

$$g_i(t_1, \ldots, t_m) \neq 0$$

e quindi $x_i = \frac{f_i}{g_i}$

□

Teorema 10.7. *Se K è un campo infinito allora*

$$\overline{F(K^m \setminus W)} = V(J_{m+1})$$

La dimostrazione è analoga al teorema 10.2 ed all'osservazione successiva e viene lasciata come esercizio.

Ovviamente $F(K^m \setminus W)$ puó essere strettamente contenuto nella sua chiusura, anche se $m = 1$. Si veda ad esempio le equazioni (10.1) che definiscono tutta la circonferenza meno il punto $(-1, 0)$.

Esercizio 10.8. *Sia $W = V(xy) \subset K^2$ e sia $F\colon K^2 \setminus W \to K^3$ definita da $F(x, y) = (\frac{x}{y^2}, \frac{y}{x^2}, x)$. Calcolare $\overline{F(K^2 \setminus W)}$.*

Definizione 10.9. *Una varietà $V \subset K^n$ immagine di una parametrizzazione razionale si dice unirazionale. In altri termini, se V è unirazionale deve esistere F come in (10.4) e*

$$\overline{F(K^m \setminus W)} = V.$$

Se $K = \mathbb{C}$ una varietá unirazionale per cui esiste una parametrizzazione razionale F iniettiva si dice razionale.

Ogni varietá razionale é anche unirazionale. L'esempio 10.1 mostra che le curve di Fermat non sono unirazionali se $n \geq 3$ mentre sono razionali se $n \leq 2$. Un classico teorema di Lüroth afferma che le curve unirazionali sono razionali.

Il teorema 10.7 dà un algoritmo per trovare le equazioni di varietà unirazionali.

Nel caso di curve che ammettono una parametrizzazione razionale con un solo parametro l'algoritmo precedente può essere semplificato. Il seguente lemma mostra in sostanza che il fenomeno visto con l'esempio 10.5 non si può ripetere nel caso di curve.

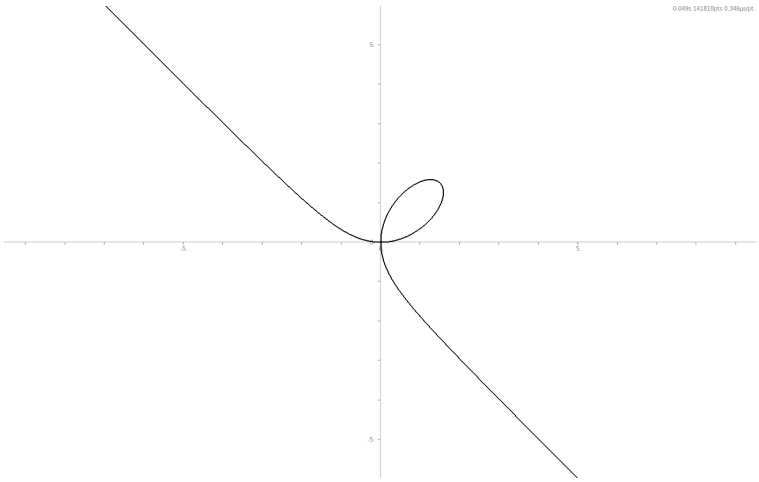


Figura 10.1: Il Folium di Cartesio

Lemma 10.10. *Sia K un campo infinito. Consideriamo le equazioni $x_i = \frac{f_i(t)}{g_i(t)}$ per $i = 1, \dots, n$ con $f_i, g_i \in K[t]$ polinomi primi tra loro. Sia $W := V(g_1 \cdots g_n) \subset K$, sia $F\colon K \setminus W \to K^n$ data da $F(t) := (\frac{f_1(t)}{g_1(t)}, \dots, \frac{f_n(t)}{g_n(t)})$, sia $I = (\dots g_i(t)x_i - f_i(t), \dots) \subset K[t, x_1, \dots, x_n]$ e sia $i\colon K \to K^{n+1}$ definita da $i(t) := (t, \frac{f_1(t)}{g_1(t)}, \dots, \frac{f_n(t)}{g_n(t)})$. Allora*

- *i) $V(I) = i(K \setminus W)$*
- *ii) $\overline{F(K \setminus W)} = V(I_1)$*

Dimostrazione. Si considera il diagramma

$$\begin{array}{ccccc} V(I) & \subset & K^{n+1} & & \\ & i \nearrow & & \searrow \pi_1 & \\ K \setminus W & & \longrightarrow^F & & K^n \end{array}$$

Per provare i) abbiamo che l'inclusione $\supset$ è banale. Viceversa sia $(\tilde{t}, \tilde{x}_1, \dots, \tilde{x}_n) \in V(I)$. Allora $g_i(\tilde{t})\tilde{x}_i - f_i(\tilde{t}) = 0 \quad \forall i$. Se $g_i(\tilde{t}) = 0$ allora anche $f_i(\tilde{t}) = 0$ e quindi f_i, g_i sarebbero divisibili per $t - \tilde{t}$ in contrasto con l'ipotesi. Pertanto $\tilde{t} \notin W$ e quindi si ricava $\tilde{x}_i = \frac{f_i(\tilde{t})}{g_i(\tilde{t})}$. Adesso ii) segue come nel teorema 10.2 □

Esempio 10.11. *Il Folium di Cartesio (si veda la Figura 10.1). Si considera la parametrizzazione razionale*

$$\begin{array}{rll} x & = \frac{3t}{1+t^3} & ((10.7)a) \\ y & = \frac{3t^2}{1+t^3} & ((10.7)b) \end{array} \tag{10.5}$$

Con la tecnica del Lemma 10.10 , per trovare equazioni implicite, occorre eliminare la t dall'ideale

$$\left((1+t^3)x - 3t, (1+t^3)y - 3t^2\right) \tag{10.6}$$

e si ottiene l'equazione

$$x^3 + y^3 - 3xy = 0 \tag{10.7}$$

Per il teorema 10.10 l'equazione rappresenta la chiusura del luogo descritto dalla parametrizzazione. Quindi $\overline{F(K \setminus W)} = V(I_1)$

Viceversa ogni punto in $\mathbb{C}^2$ che soddisfa l' equazione (10.7) proviene dalla parametrizzazione. Infatti il punto $(x, y) = (0, 0)$ viene ottenuto per $t = 0$, mentre per gli altri valori di x e y uno dei coefficienti di t^3 in 10.6 è $\neq 0$ e quindi dal teorema di estensione 8.3 si ha la tesi. Quindi in questo caso $F(K \setminus W) = \overline{F(K \setminus W)} = V(I_1)$.

Esercizio 10.12. *Provare che ogni $(x_0, y_0) \in \mathbb{R}^2$ appartenente al Folium di Cartesio $x^3+y^3-3xy=0$ proviene da un $t_0 \in \mathbb{R}$ secondo la parametrizzazione (*(10.5)*).*

Suggerimento: $\mathbb{R}$ non é algebricamente chiuso e quindi non si puó applicare il teorema 8.3 Posto $f=(1+t^3)x-3t$, $g=(1+t^3)y-3t^2$, possiamo supporre $x_0 \neq 0, y_0 \neq 0$ ed abbiamo $Res(f(x_0,y_0,t), g(x_0,y_0,t),t) = 0$. Pertanto $f(x_0,y_0,t)$ e $g(x_0,y_0,t)$ hanno un fattore a comune in $\mathbb{R}[t]$. Se questo fattore ha grado due allora si puó scrivere $t^3x_0-3t+x_0=(t^2+bt+c)(tx_0-\alpha)$ e $t^3y_0-3t^2+y_0=(t^2+bt+c)(ty_0-\beta)$ da cui $-\alpha c=x_0$, $-\beta c=y_0$, $\beta x_0=\alpha y_0$ e quindi ...

11 Ideali omogenei e varietà proiettive

Consideriamo lo spazio proiettivo $\mathbb{P}^n(K)$ sul campo K con coordinate omogenee $(x_0,\ldots,x_n)$. Se $f(x_0,\ldots,x_n)=f(x)$ è un polinomio qualunque non è possibile definire il luogo degli zeri $\{x \in \mathbb{P}^n(K) | f(x)=0\}$. Invece se f è un polinomio omogeneo, cioè se tutti i suoi termini hanno lo stesso grado allora posto $d = deg\ f$ vale la relazione $f(\lambda x_0, \lambda x_1, \ldots, \lambda x_n) = \lambda^d f(x_0, x_1, \ldots, x_n)$. Pertanto $f(\lambda x)=0$ se e solo se $f(x)=0$. Quindi é ben definito in $\mathbb{P}^n(K)$ il luogo degli zeri di un polinomio omogeneo. Abbiamo la:

Definizione 11.1. *Se f è un polinomio omogeneo in $K[x_0,\ldots,x_n]$ allora*

$$V(f) = \{x \in \mathbb{P}^n(K) | f(x)=0\}$$

si dice una ipersuperficie proiettiva.

Definizione 11.2. *Se $f_1, \ldots f_p$ sono polinomi omogenei in $K[x_0,\ldots,x_n]$ allora*

$$V(f_1,\ldots f_p) = \{x \in \mathbb{P}^n(K) | f_1(x) = \ldots = f_p(x) = 0\}$$

si dice una varietà (algebrica) proiettiva.

Notiamo che ogni polinomio è somma delle sue componenti omogenee, cioé se $f \in K[x_0,\ldots,x_n]$ abbiamo $f=\sum_{i=0}^{d} f_i$ dove f_i é omogeneo di grado d.

Vediamo adesso che in termini algebrici le varietà proiettive corrispondono agli ideali omogenei di $K[x_0,\ldots,x_n]$.

Definizione 11.3. *Un ideale I di $K[x_0,\ldots,x_n]$ si dice omogeneo se è generato da polinomi omogenei.*

Lemma 11.4. *Sia I un ideale*

$$I \text{ è omogeneo} \iff \text{se } g \in I \text{ anche le sue componenti omogenee } \quad g_i \in I \quad \forall i.$$

Dimostrazione.

$\Longrightarrow$ Siano $\{f_1,\ldots,f_p\}$ generatori omogenei di I. Allora esistono $a_j \in K[x_0,\ldots,x_n]$ tali che

$$g = \sum g_i = \sum a_j f_j \tag{11.1}$$

Posto $a_j = \sum a_{jk}$ (componenti omogenee), sostituiamo queste espressioni in (11.1) ed otteniamo $\sum g_i = \sum a_{jk} f_j$. Eguagliando tra loro i termini di ogni grado si ottiene che ogni g_i è combinazione dei f_j.

$\Longleftarrow$ Basta prendere le componenti omogenee di un insieme di generatori e si ottiene ancora un insieme di generatori. □

Conviene guardare in questo contesto a $S = K[x_0, \ldots, x_n]$ come ad un anello graduato, cioè posto $S_d = \{f \in S | deg\ f = d\}$ si ha $S = \oplus_{d \geq 0} S_d$ (somma diretta di spazi vettoriali) con la struttura moltiplicativa che soddisfa a $S_d \cdot S_{d'} \subset S_{d+d'}$. Allora il lemma 11.4 si traduce nel fatto che I è omogeneo se e solo se $I = \oplus_{d \geq 0}(I \cap S_d)$.

Se I è un ideale omogeneo generato da $f_1, \ldots, f_p$ allora $V(I) = V(f_1, \ldots, f_p) \subset \mathbb{P}^n(K)$ è una varietà proiettiva ed ogni varietá proiettiva ha questa forma. Precisamente si puó definire

$$V(I) = \{x \in \mathbb{P}^n | f(x) = 0 \quad \forall f \text{ omogeneo } \in I\}$$

Analogamente a quanto visto nel caso affine, le varietà proiettive sono gli insiemi chiusi per la topologia di Zariski su $\mathbb{P}^n(K)$ (la dimostrazione che soddisfano agli assiomi per gli insiemi chiusi é analoga a quella vista nella 6).

$\mathbb{P}^n(K)$ è ricoperto da $n+1$ aperti affini standard $U_i := \{x \in P^n | x_i \neq 0\}$ ciascuno dei quali è isomorfo a K^n. Se $V = V(f_1, \ldots, f_p)$ è una varietà proiettiva allora $V \cap U_i$ è la varietà affine $V(g_1, \ldots, g_p) \subset U_i$ dove $g_j(y_0, \ldots, \hat{y_i}, \ldots, y_n) = f_j(y_0, \ldots, 1, \ldots, y_n)$

Proposizione 11.5. *Se I è un ideale omogeneo allora $\sqrt{I}$ è un ideale omogeneo.*

Dimostrazione. Sia f un elemento di $\sqrt{I}$. Per il lemma 11.4 è sufficiente provare che tutte le sue componenti omogenee appartengono ancora a $\sqrt{I}$.

Infatti sia $f = \sum f_i$ e sia f_{max} la componente omogenea di grado massimo. Abbiamo per definizione che $\exists\ n$ tale che $f^n \in I$. Siccome I è omogeneo per il lemma 11.4 $(f^n)_{max} \in I$. È facile verificare che $(f^n)_{max} = (f_{max})^n$ e quindi $f_{max} \in \sqrt{I}$. Si può ripetere il ragionamento per $f - f_{max}$ provando così che tutte le componenti omogenee di f appartengono a $\sqrt{I}$. □

Se $V = V(I) \subset \mathbb{P}^n(K)$ è una varietà proiettiva allora è definito il cono affine $C_V = \{x \in K^{n+1} | f(x) = 0 \quad \forall f \in I\}$. C_V è un cono per l'origine perché se $x \in C_V$ allora $\lambda x \in C_V$ $\forall \lambda \in K$. In particolare C_V è un insieme finito se e solo se C_V coincide con l'origine. Questo può essere espresso nel modo seguente:

$$V \subset P^n(K) \text{ è vuota} \iff C_V \text{ è finito}$$

Se V è una varietà proiettiva allora

$$I(V) := \{f \in K[x_0, \ldots, x_n] | f(a_0, \ldots, a_n) = 0 \quad \forall \text{n-pla di coord. omogenee } (a_0, \ldots, a_n) \in V\}$$

L'ideale $I(V)$ coincide con l'ideale $I(C_V)$ definito nel caso affine.

Lemma 11.6. *Se K è un campo infinito e $V \subset \mathbb{P}^n(K)$ è una varietà proiettiva allora $I(V)$ è un ideale omogeneo.*

Dimostrazione. Sia $f \in I(V)$. Allora per ogni $(a_0, \ldots, a_n) \in V$ e per ogni $\lambda \in K$ abbiamo $f(\lambda a_0, \ldots, \lambda a_n) = 0$. Se $f = \sum f_i$ (decomposizione nelle componenti omogenee) l'equazione precedente diventa $\sum \lambda^i f_i(a_0, \ldots, a_n) = 0 \quad \forall \lambda \in K$. Dal principio di identità dei polinomi segue $f_i(a_0, \ldots, a_n) = 0\ \forall i$ e quindi $f_i \in I(V)$ □

Se $V \subset \mathbb{P}^n(K)$ é una varietá proiettiva, l'anello graduato $K[x_0, \ldots, x_n]/I(V)$ si dice l'anello delle coordinate omogeneo di V.

Nella parte restante di questa sezione studieremo le relazioni tra ideali omogenei e varietà proiettive. Occorre osservare subito che il Nullstellensatz debole non si estende parola per parola al caso proiettivo. Infatti $K[x_0, \ldots, x_n]$ contiene l'ideale massimale omogeneo $\mathcal{M} = (x_0, \ldots, x_n)$ per cui $V(\mathcal{M}) = \emptyset$ (mentre per il Nullstellensatz debole affine se $I \neq K[x_0, \ldots, x_n]$ allora $V(I) \neq \emptyset$). Considerazioni analoghe possono essere fatte per il Nullstellensatz forte.

Fortunatamente i casi patologici sono tutti della stessa natura di questo. L'ideale $\mathcal{M}$ assume un ruolo speciale nella teoria delle varietà proiettive ed è chiamato col nome di ideale massimale irrilevante.

Queste considerazioni portano a denotare con simboli diversi gli ideali associati a varietà proiettive o affini. Quando non é chiaro dal contesto scriviamo $V_a(I)$ per denotare la varietà affine associata a I e $I_a(V)$ per denotare l'ideale (generalmente non omogeneo) associato alla varietà affine V. Notiamo che se I è omogeneo e $V(I) \neq \emptyset$ è facile verificare che $V_a(I) = C_{V(I)}$, notando che se $V(I) = \emptyset$ tale uguaglianza è falsa perché $V_a(I)$ contiene l'origine se $1 \notin I$.

Si ricava la seguente versione proiettiva del Nullstellensatz debole:

Teorema 11.7. *(Nullstellensatz debole proiettivo). Sia $V(I) \subset \mathbb{P}^n(K)$ una varietà proiettiva con K algebricamente chiuso. Allora*

$$V(I) = \emptyset \Longleftrightarrow \exists r : (x_0, \ldots, x_n)^r = \mathcal{M}^r \subset I$$

Dimostrazione.

$\Longrightarrow$ Se $V(I) = \emptyset$ allora ci sono due casi: $V_a(I) = \emptyset$ oppure $V_a(I) = \{0\}$. Nel primo caso $I = (1)$ dal Nullstellensatz debole affine e la tesi è verificata con $r = 0$. Nel secondo caso $V_a(I) = C_V$ coincide con l'origine O. Quindi per il Nullstellensatz forte affine $\sqrt{I} = I_a(V_a(I)) = I_a(O) = \mathcal{M}$. In particolare esiste N tale che $x_i^N \in I \quad \forall i$ e quindi $\mathcal{M}^{N(n+1)} \subset I$.

$\Longleftarrow$ Per ipotesi $x_i^r \in I$, quindi $C_V = V_a(I)$ coincide con l'origine e pertanto $V(I) = \emptyset$. □

Teorema 11.8. *(Nullstellensatz proiettivo) Sia I un ideale omogeneo in $K[x_0, \ldots, x_n]$ con K algebricamente chiuso. Se $\mathbb{P}^n(K) \supset V(I) \neq \emptyset$ allora*

$$I(V(I)) = \sqrt{I}$$

Dimostrazione. Nell'ipotesi del Teorema è facile verificare che $V_a(I) = C_{V(I)}$, notando che se $V(I) = \emptyset$ tale uguaglianza è falsa perché $V_a(I)$ contiene l'origine se $1 \notin I$.

Usando la versione affine del NullStellenSatz (Teorema 6.13) abbiamo $\sqrt{I} = I_a(V_a(I)) = I(C_{V(I)}) = I(V(I)$. □

Osservazione Il lettore interessato può ripercorrere i passi dell'algoritmo di divisione e verificare che quando si divide un polinomio omogeneo f per dei polinomi omogenei $f_1, \ldots, f_r$ si ottiene $f = \sum a_i f_i + r$ dove i quozienti a_i ed il resto r sono ancora polinomi omogenei. In particolare se $r \neq 0$ allora $deg\ r = deg\ f$. Se f e g sono omogenei allora la S-coppia $S(f, g)$ è omogenea. Analizzando l'algoritmo di Buchberger, segue che un ideale omogeneo ha una base di base di Groebner formata da polinomi omogenei.

Esercizio 11.9. *Sia $V \subset \mathbb{P}^n$ una varietà proiettiva. Provare che $I(V)$ è primo se e solo se V è irriducibile.*

Vediamo adesso la relazione tra una varietà affine e la sua chiusura proiettiva. Lo spazio affine K^n con coordinate $(x_1, \ldots, x_n)$ può essere completato con un "iperpiano all'∞" ed immerso come aperto in $\mathbb{P}^n(K)$ con coordinate omogenee $(x_0, \ldots, x_n)$.

Definizione 11.10. *Sia g un polinomio in $K[x_1, \ldots, x_n]$ di grado d. Allora*

$$g^h := x_0^d g(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0})$$

è un polinomio omogeneo in $K[x_0, \ldots, x_n]$ ancora di grado d, che si dice l'omogeneizzato di g.

Esempio 11.11. *Se $g = x_1 + x_2x_3 + 5x_2^2x_3$ allora $g^h = x_0^2x_1 + x_0x_2x_3 + 5x_2^2x_3$.*

Lemma 11.12.

i) $g^h(1, x_1, \ldots, x_n) = g(x_1, \ldots, x_n)$ *cioè deomogeneizzando si riottiene* g

ii) Sia $F(x_0, \ldots, x_n)$ *un polinomio omogeneo e sia* x_0^e *la massima potenza di* x_0 *che divide* F. *Se* $f = F(1, x_1, \ldots, x_n)$ *è la deomogeneizzazione di* F *allora* $F = x_0^e \cdot f^h$.

iii) $(g^h)^m = (g^m)^h$ *per ogni* $m \in \mathbb{N}$, *dove* h *corrisponde all'omogeneizzazione.*

Dimostrazione. i) e iii) sono evidenti dalle definizioni. Per provare ii) osserviamo che f ha grado $d - e$ e quindi $f^h = x_0^{d-e} F(1, \frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}) = x_0^{-e} F(x_0, x_1, \ldots, x_n)$. □

Definizione 11.13. *Sia* I *un ideale in* $K[x_1, \ldots, x_n]$. *Definiamo*

$$I^h := < f^h | f \in I \ > \subset K[x_0, \ldots, x_n]$$

I^h è un ideale omogeneo.

Osservazione Se $f_1, \ldots, f_p$ generano I allora può essere $\left(f_1^h, \ldots, f_p^h\right) \subset_{\neq} I^h$. Ad esempio consideriamo $I(C) = (f_1, f_2) = (x_2 - x_1^2, x_3 - x_1^3) \subset \mathbb{R}[x_1, x_2, x_3]$ ideale della cubica gobba (si veda l'esempio 5.6). Allora $(f_1^h, f_2^h) = (x_2 x_0 - x_1^2, x_3 x_0^2 - x_1^3)$. Abbiamo $(f_2 - x_1 f_1)^h = (x_3 - x_1 x_2)^h = x_0 x_3 - x_1 x_2 \in I^h$ ma l'ultimo polinomio non è combinazione di f_1^h e f_2^h (basta guardare i gradi). Si può provare (ad esempio utilizzando il prossimo Teorema 11.14) che $I(C)^h$ è generato dai minori 2×2 della matrice

$$\begin{pmatrix} x_0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \end{pmatrix}.$$

Teorema 11.14. *Sia* I *un ideale di* $K[x_1, \ldots, x_n]$ *e sia* $G = \{g_1, \ldots, g_s\}$ *una base di Gröbner di* I *rispetto ad un ordine monomiale graduato. Allora* $G^h = \{g_1^h, \ldots, g_s^h\}$ *genera* I^h.

Dimostrazione. Proveremo l'enunciato più forte che G^h è una base di Gröbner per I^h rispetto ad un conveniente ordine monomiale in $K[x_0, \ldots, x_n]$ che andiamo a definire. Ogni monomio in $K[x_0, \ldots, x_n]$ si può scrivere come $x_1^{\alpha_1} \ldots x_n^{\alpha_n} x_0^d = x^\alpha x_0^d$. Allora possiamo estendere l'ordine graduato $>$ ad un ordine $>_h$ in $K[x_0, \ldots, x_n]$ dato da

$$x^\alpha x_0^d >_h x^\beta x_0^e \iff x^\alpha > x^\beta \text{oppure } x^\alpha = x^\beta \text{e } d > e$$

È facile verificare che $>_h$ è un ordine monomiale e che $\mathrm{LT}_{>_h}(f^h) = \mathrm{LT}_>(f) \quad \forall f \in K[x_1, \ldots, x_n]$ (infatti x_0 è minore rispetto a $>_h$ di qualunque polinomio in $x_1, \ldots, x_n$ e quindi $\mathrm{LT}_{>_h}(f^h)$ non contiene x_0 ed è un monomio che appariva già in f).

Dobbiamo provare che $< LT_{>_h}(I^h) >$ è generato da $LT_{>_h}(G^h)$. Infatti sia $F \in I^h$. Abbiamo $F = \sum a_j f_j^h$ con $f_j \in I$. Facciamo vedere che il deomogeneizzato di F appartiene a I. Infatti

$$f := F(1, x_1, \ldots, x_n) = \sum a'_j f_j^h(1, x_1, \ldots, x_n) = \sum a'_j f_j \in I$$

(l'ultima uguaglianza per il lemma 11.12, dove $a'_j(x_1, \ldots, x_n) = a_j(1, x_1, \ldots, x_n)$) Ancora dal lemma 11.12 abbiamo $F = x_0^e \cdot f^h$ e quindi

$$\mathrm{LT}_{>_h}(F) = x_0^e \cdot \mathrm{LT}_{>_h}(f^h) = x_0^e \cdot \mathrm{LT}_>(f)$$

Siccome G è una base di Gröbner per I, $\mathrm{LT}_>(f)$ è un multiplo di qualche $\mathrm{LT}_>(g_i) = \mathrm{LT}_{>_h}(g_i^h)$ e quindi anche $\mathrm{LT}_{>_h}(F)$ è un multiplo di qualche $\mathrm{LT}_{>_h}(g_i^h)$ □

Confrontando il teorema 11.14 con l'osservazione che lo precede e con l'esempio 5.6 il lettore può notare che $(x_2 - x_1^2, x_3 - x_1^3)$ è una base di Gröbner per l'ideale della cubica gobba $I(C)$ con Lex ma non può esserlo per un ordine graduato.

I comandi di M2 relativi all'omogeneizzazione sono homogenize(F, x) dove F é un polinomio e homogenize(gens gb I, x) dove I é un ideale (quest'ultimo comando é basato sul teor. 11.14).

Definizione 11.15. *Se $W \subset K^n$ allora la chiusura proiettiva $\overline{W}$ è la chiusura di W nella topologia di Zariski di $\mathbb{P}^n(K)$.*

Teorema 11.16. *Sia $W \subset K^n$ una varietà affine. Allora per la sua chiusura proiettiva*

$$\overline{W} = V(I_a(W)^h)$$

Dimostrazione.

$\subset$ Proviamo che $W \subset V(I_a(W)^h)$. Infatti se $x \in W$ e $f \in I_a(W)$ abbiamo $f(x) = 0$ da cui $f^h(1, x) = 0$ e quindi tutti i polinomi di $I_a(W)^h$ si annullano su x. La tesi segue prendendo la chiusura di ambo i membri.

$\supset$ Sia $\overline{W} = V(F_1, \ldots, F_s)$. Ogni F_i si annulla su $\overline{W}$ e quindi $f_i = F_i(1, x_1, \ldots, x_n)$ si annulla su W. Pertanto $f_i \in I_a(W)$ da cui $f_i^h \in I_a(W)^h$. Abbiamo $F_i =$ (vedi 11.12) $x_0^{e_i} f_i^h \in I_a(W)^h$ e quindi $(F_1, \ldots, F_s) \subset I_a(W)^h$ e prendendo V di ambo i membri si ha la tesi. □

Teorema 11.17. *Sia K un campo algebricamente chiuso e sia $I \subset K[x_1, \ldots, x_n]$ un ideale. Allora*

$$\overline{V_a(I)} = V(I^h) \subset \mathbb{P}^n(K)$$

Dimostrazione.

$\subset$ $V_a(I) \subset$ (dall'analoga inclusione del teorema precedente) $V(I_a(V_a(I))^h) =$ (dal Nullstellensatz) $V((\sqrt{I})^h) \subset V(I^h)$

$\supset$ Sia $\overline{V_a(I)} = V(F_1, \ldots, F_s)$. Come nell'analoga inclusione del teorema precedente abbiamo $(F_1, \ldots, F_s) \subset I_a(V_a(I))^h$. Per il Nullstellensatz l'ultimo ideale è uguale a $(\sqrt{I})^h$ ed è facile verificare che a sua volta questo ideale è incluso in $\sqrt{I^h}$ (occorre usare il lemma 11.12 iii). Pertanto $(F_1, \ldots, F_s) \subset \sqrt{I^h}$ e prendendo V di ambo i membri si ottiene $\overline{V_a(I)} \supset V(\sqrt{I^h}) = V(I^h)$ □

Dai teoremi 11.17 e 11.14 si ottiene un algoritmo per calcolare la chiusura proiettiva di una varietà affine su un campo algebricamente chiuso. Si procede nel modo seguente:

Se $W = V(I)$ è una varietà affine si calcola una base di Gröbner G di I rispetto ad un ordine graduato. Allora $\overline{W}$ è definita in $\mathbb{P}^n(K)$ da G^h.

Osservazione Il teorema 11.17 è falso su $\mathbb{R}$. Ad esempio se $I = (x_1^2 + x_2^4) \subset \mathbb{R}[x, y]$ allora $V_a(I) \subset \mathbb{R}^2$ consiste solo dell'origine e quindi la sua chiusura proiettiva è data dal punto di coordinate omogenee $(1, 0, 0) \subset \mathbb{P}^2(\mathbb{R})$. D'altronde $V(I^h) = V(x_0^2 x_1^2 + x_2^4) = (1, 0, 0) \cup (0, 1, 0)$.

12 Curve algebriche piane.

In questa sezione supponiamo K algebricamente chiuso. Studiamo brevemente le varietà date da una singola equazione in K^2 (curve piane affini) o in $P^2(K)$ (curve piane proiettive).

Proposizione 12.1. *Sia $f \in K[x, y]$ e sia $C = V(f) \subset K^2$ una curva affine. Allora $V(f^h) \subset \mathbb{P}^2(K)$ è la chiusura proiettiva di C.*

Dimostrazione. Dalla definizione segue che se $I = (f)$ allora $I^h = (f^h)$. La tesi segue allora dal teorema 11.17 . □

Lemma 12.2. *Sia $f \in K[x, y]$. $V(f) \subset K^2$ è irriducibile se e solo se f è potenza di un irriducibile.*

Dimostrazione. Per il teorema 6.18 $V(f)$ è irriducibile se e solo se $I(V(f))$ è primo. Per il teorema degli zeri di Hilbert 6.13 $I(V(f)) = \sqrt{(f)}$. Se $f = f_1^{a_1} \cdots f_k^{a_k}$ è la decomposizione di f in polinomi irriducibili segue facilmente che $\sqrt{(f)} = (f_1 \cdots f_k)$. Quindi $V(f)$ è irriducibile se e solo se $(f_1 \cdots f_k)$ è primo. Abbiamo che $(f_1 \cdots f_k)$ è primo se solo se $f_1 \cdots f_k$ è irriducibile, cioè se e solo se $k = 1$ da cui la tesi. □

Se $f = f_1^{a_1} \cdots f_k^{a_k}$ è la decomposizione di f in fattori irriducibili allora $V(f)$ è unione delle sue componenti irriducibili $V(f_i)$. Se poniamo

$$f_{rid} := f_1 \cdots f_k$$

allora evidentemente $V(f) = V(f_{rid})$ e $\sqrt{(f)} = (f_{rid})$.

Definizione 12.3. *Un polinomio f si dice ridotto se $f = f_{rid}$ cioè se f è irriducibile oppure contiene i suoi fattori irriducibili con molteplicità 1.*

Lemma 12.4. *Sia $f \in K[x_1, \ldots, x_n]$ e sia car $K = 0$. Allora*

$$f_{rid} = \frac{f}{MCD(f, f_{x_1}, \ldots, f_{x_n})}$$

Dimostrazione. Sia $f = f_1^{a_1} \cdots f_k^{a_k}$ la decomposizione di f in fattori irriducibili. È sufficiente provare che

$$MCD(f, f_{x_1}, \ldots, f_{x_n}) = f_1^{a_1 - 1} \cdots f_k^{a_k - 1}$$

Abbiamo

$$f_{x_i} = \sum_{j=1}^{k} a_j \frac{\partial f_j}{\partial x_i} f_1^{a_1} \cdots f_j^{a_j - 1} \cdots f_k^{a_k} = f_1^{a_1 - 1} \cdots f_k^{a_k - 1} \sum_{j=1}^{k} a_j \frac{\partial f_j}{\partial x_i} f_1 \cdots \hat{f}_j \cdots f_k \quad (12.1)$$

Quindi $f_1^{a_1 - 1} \cdots f_k^{a_k - 1}$ divide f e tutte le sue derivate prime f_{x_i}. È facile verificare da (12.1) che $\forall j \quad f_j^{a_j}$ non divide tutte le derivate prime di f e quindi segue la tesi. □

Nel resto di questo paragrafo consideriamo sempre curve $C = V(f)$ con f polinomio ridotto. Il lemma 12.4 mostra che possiamo sempre ricondurci a questo caso (almeno se car $K = 0$)

Lemma 12.5. *Sia $f \in K[x, y]$ di grado totale d. Allora l'intersezione di $V(f)$ con una retta che non è una sua componente irriducibile consiste al più di d punti.*

Dimostrazione. La retta può essere parametrizzata da $x = x_0 + at$, $y = y_0 + bt$. Sostituendo abbiamo l'equazione $f(x_0 + at, y_0 + bt) = 0$ che è un polinomio non nullo di grado $\leq d$ nella variabile t e quindi ha al più d radici. □

Definizione 12.6. *Sia $P = (x_0, y_0)$ un punto di intersezione tra una retta L ed una curva $C = V(f) \subset K^2$. Sia L parametrizzata da $x = x_0 + at$, $y = y_0 + bt$, (in modo che P corrisponde al valore $t = 0$). La molteplicità della radice $t = 0$ del polinomio $p(t) = f(x_0 + at, y_0 + bt)$ si dice molteplicità di intersezione di L e C nel punto P.*

Studiamo ora come varia la molteplicità di intersezione tra una curva C ed una retta L in un punto $P \in C$ al variare di L tra le rette passanti per P. Se $P = (x_0, y_0)$ abbiamo che $f(x_0, y_0) = 0$ e L è parametrizzata da $x = x_0 + at$, $y = y_0 + bt$ al variare di $(a, b) \in \mathbb{P}^1$.

Consideriamo lo sviluppo di Taylor

$$f(x_0+at, y_0+bt) = f(x_0, y_0)+[f_x(x_0, y_0)a + f_y(x_0, y_0)b]\, t+\ldots (\text{termini di grado superiore in } t) \tag{12.2}$$

Ne segue che

- i) Se $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ allora tutte le rette per P hanno molteplicità di intersezione ≥ 2 con C in P .
- ii) Se $(f_x(x_0, y_0), f_y(x_0, y_0)) \neq (0, 0)$ allora la molteplicità di intersezione della retta L è 1 se $[f_x(x_0, y_0)a + f_y(x_0, y_0)b] \neq 0$ mentre è ≥ 2 se $[f_x(x_0, y_0)a + f_y(x_0, y_0)b] = 0$.

Notiamo che l'equazione $[f_x(x_0, y_0)a + f_y(x_0, y_0)b] = 0$ è risolta da $a = f_y(x_0, y_0)$ e $b = -f_x(x_0, y_0)$. Eliminando il parametro la retta corrispondente ha equazione

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0$$

Questa osservazione motiva le seguenti

Definizione 12.7. *Un punto $P \in C = V(f)$ curva piana affine si dice singolare se $f_x(P) = f_y(P) = 0$. Altrimenti P si dice nonsingolare. Una curva si dice nonsingolare (o liscia) se tutti i suoi punti sono nonsingolari.*

Definizione 12.8. *In un punto $P = (x_0, y_0) \in C = V(f)$ nonsingolare la retta di equazione*

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0$$

si dice la retta tangente.

Algoritmo per la nonsingolarità di una curva.

Sia K algebricamente chiuso. Una curva $C = V(f)$ (ridotta) è nonsingolare se e solo se $V(f, f_x, f_y) = \emptyset$, ovvero (dal teorema degli zeri di Hilbert) se e solo se $(f, f_x, f_y) = (1)$. Quest'ultima condizione può essere verificata calcolando una base di Gröbner dell'ideale (f, f_x, f_y).

Definizione 12.9. *Un punto $p \in V(f)$ si dice di molteplicità r se tutte le derivate parziali di f di ordine $\leq r - 1$ si annullano in P e se esiste una derivata parziale di ordine r non nulla in P. Un punto di molteplicità $2, 3, \ldots$ si dice anche doppio, triplo, ...*

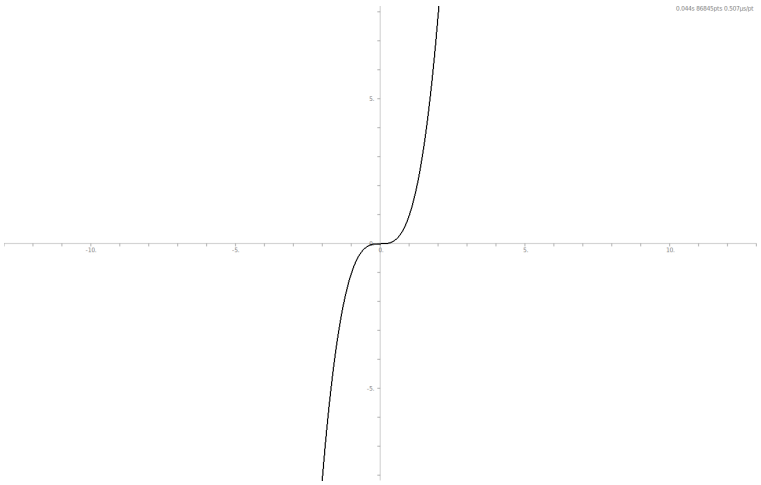


Figura 12.2: $y - x^3$

I punti nonsingolari corrispondono ai punti di molteplicità 1. Calcolando i termini successivi dello sviluppo di Taylor (12.2), otteniamo che in un punto P di molteplicità r tutte le rette per P incontrano C con molteplicità di intersezione in $P = r$ escluso un numero finito, dato dalle soluzioni in (a, b) di

$$\sum_{i=0}^{r} \binom{r}{i} \frac{\partial f}{\partial^{r-i} x \partial^i y}(P) a^{r-i} b^i = 0 \tag{12.3}$$

Le rette appena menzionate possono essere considerate come rette tangenti in P.

Definizione 12.10. *Un punto singolare di molteplicità $r(\geq 2)$ si dice <u>ordinario</u> se ha esattamente r tangenti distinte, cioè se il discriminante del polinomio in (12.3) è non nullo.*

Esempio 12.11. *La curva $V(y^2-x^2(x+1))$ ha un punto doppio ordinario nell'origine. La curva $V(y^2-x^3)$ ha un punto doppio non ordinario nell'origine, che si dice cuspide. La curva $V(x^4-x^2+y^2)$ (il cui grafico reale è un "otto") ha un punto doppio ordinario nell'origine.*

Definizione 12.12. *Un punto $P \in V(f) = C$ nonsingolare si dice un flesso se la tangente in P ha molteplicità di intersezione ≥ 3 con C in P.*

Esempio 12.13.

i) L'origine è un flesso per la curva $y - x^3 = 0$, (Figura 12.2).

ii) La curva

$$2x^4 - 3x^2y + y^2 - 2y^3 + y^4 = 0$$

(Figura 12.3) ha molteplicità 2 nell'origine. La molteplicità di intersezione della tangente (generalizzata) $y = 0$ nell'origine è 4. L'origine si dice un <u>tacnodo</u>.

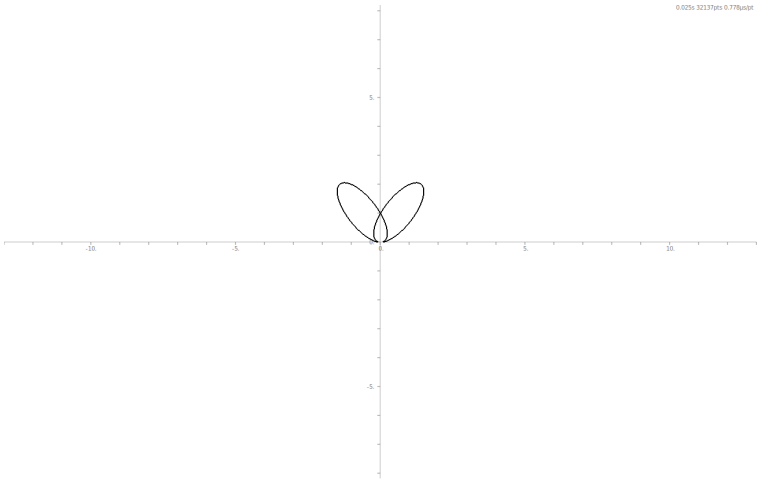


Figura 12.3: $2x^4 - 3x^2y + y^2 - 2y^3 + y^4$

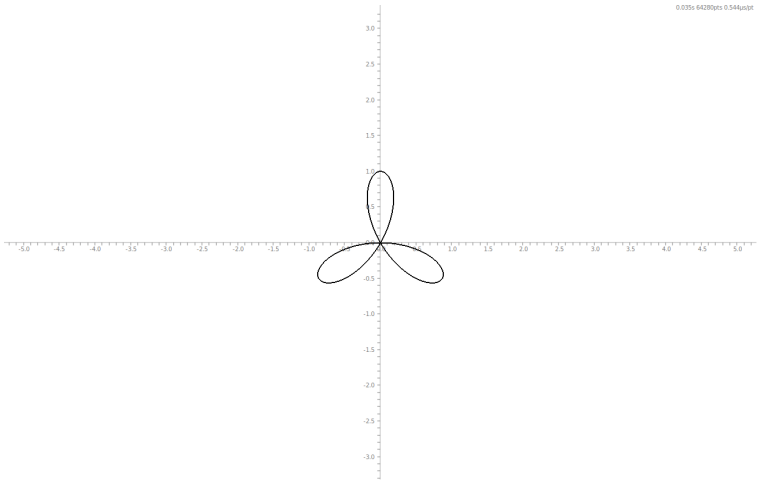


Figura 12.4: $(x^2 + y^2)^2 + 3x^2y - y^3$

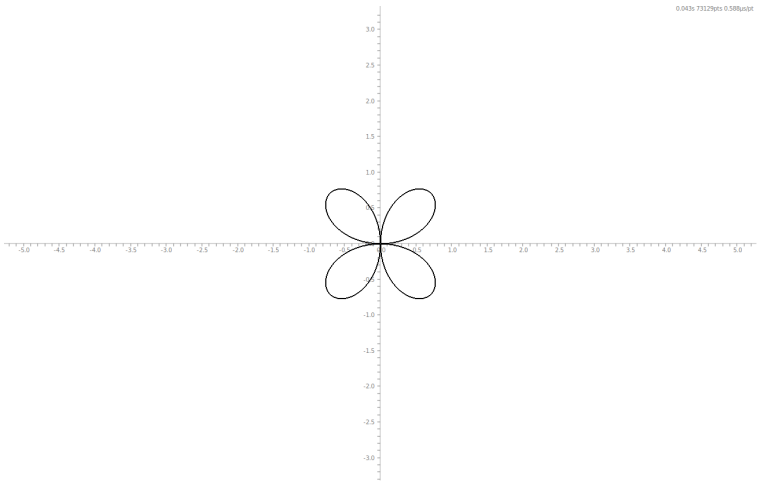


Figura 12.5: $(x^2+y^2)^3 - 4x^2y^2$

iii) La curva

$$(x^2+y^2)^2 + 3x^2y - y^3 = 0$$

ha un punto triplo ordinario nell'origine, (Figura 12.4).

iv) La curva

$$(x^2+y^2)^3 - 4x^2y^2 = 0$$

("quadrifoglio") ha un punto quadruplo non ordinario nell'origine, (Figura 12.5).

Proposizione 12.14. *Un punto P nonsingolare per $V(f)$ è un flesso se e solo se*

$$det \begin{vmatrix} 0 & f_x & f_y \\ f_x & f_{xx} & f_{xy} \\ f_y & f_{xy} & f_{yy} \end{vmatrix} = 0 \tag{12.4}$$

Dimostrazione. Sostituendo $(a,b) = (-f_y, f_x)$ all'equazione (12.3) $af_{xx} + 2abf_{xy} + b^2f_{yy} = 0$ si ottiene esattamente la (12.4). □

Osservazione Una curva differenziabile in $\mathbb{R}^2$ di equazione implicita $f(x,y) = 0$ ha curvatura nei punti nonsingolari uguale a

$$k = \pm \frac{det \begin{vmatrix} 0 & f_x & f_y \\ f_x & f_{xx} & f_{xy} \\ f_y & f_{xy} & f_{yy} \end{vmatrix}}{(f_x^2 + f_y^2)^{3/2}}$$

Pertanto i punti di flesso corrispondono ai punti di curvatura nulla.

12.1 Il Teorema di Bezout

Teorema 12.15 (Forma debole del Teorema di Bezout.)**.** *Siano $C = V(f)$ e $D = V(g)$ due curve senza componenti comuni con f, g di gradi rispettivamente d, e. Allora $C \cap D$ consiste al più in de punti.*

Dimostrazione. Se uno dei due polinomi f, g dipende solo da x o solo da y allora la curva corrispondente è data da una unione di rette, ed in questo caso il risultato è evidente. Altrimenti le coordinate dei punti di intersezione devono soddisfare le due equazioni non nulle $Res(f, g, x) = 0$ (in y) e $Res(f, g, y) = 0$ (in x) e quindi i punti di intersezione sono in numero finito che denotiamo con $\{P_1, \ldots, P_k\}$. Sia L_{ij} la retta per P_i e P_j. Per valutare il numero dei punti conviene scegliere un nuovo sistema di coordinate (con un cambiamento affine) dove gli assi coordinati non sono paralleli alle rette L_{ij} per $1 \leq i < j \leq k$.

Un cambiamento di coordinate affine non altera i gradi dei polinomi, che chiamiamo ancora con f e g. Adesso per ogni radice x_i del polinomio $Res(f, g, y)$ esiste al più un valore y_i tale che $(x_i, y_i) \in C \cap D$. Quindi è sufficiente osservare che il grado di $Res(f, g, y)$ è $\leq de$.

Posto $f = \sum_{i=0}^{d} a_i(x) y^{d-i}$, $g = \sum_{j=0}^{e} b_j(x) y^{e-j}$ con $deg\ a_i \leq i$ e $deg\ b_j \leq j$, con una trasformazione affine posso supporre $a_0 \neq 0$, $b_0 \neq 0$, calcoliamo

$$\det \begin{vmatrix} a_0 & a_1 & \ldots & a_d & & & & & \\ & a_0 & \ddots & & \ddots & & & & \\ & & \ddots & \ddots & & \ddots & & & \\ & & & a_0 & a_1 & \ldots & a_d & & \\ b_0 & b_1 & \ldots & \ldots & b_e & & & & \\ & b_0 & \ddots & & & \ddots & & & \\ & & \ddots & \ddots & & & \ddots & & \\ & & & \ddots & \ddots & & & \ddots & \\ & & & & b_0 & b_1 & & & b_e \end{vmatrix}$$

Il determinante precedente è una somma di tanti monomi in $a_i(x)$ e $b_j(x)$. Affermiamo che ciascuno di questi monomi ha grado de. Infatti, assumendo come convenzione $a_i = 0$ per $i < 0, d < i$ ed analogamente per b_j, abbiamo che la matrice ha coefficienti

$$c_{ij} = \begin{cases} a_{j-i} & \text{per } 0 \leq i \leq e-1 \\ b_{j+e-i} & \text{per } e \leq i \leq d+e-1 \end{cases}$$

e ogni addendo del determinante ha la forma

$$\pm \prod_{\sigma(j) \leq d-1} c_{\sigma(j)j} \prod_{\sigma(j) \geq d} c_{\sigma(j)j}$$

che ha grado (quando è non nullo)

$$\leq \sum_{\sigma(j) \leq e-1} (j - \sigma(j)) + \sum_{\sigma(j) \geq e} (j + d - \sigma(j)) = \sum_j j - \sum_j \sigma(j) + ed = ed.$$

□

Sia ora $F \in K[x,y,z]$ un polinomio omogeneo. Consideriamo la curva proiettiva $V(F) \subset \mathbb{P}^2$ che ha parte affine definita da $f(x,y) = F(x,y,1)$.

Possiamo definire un punto $P \in V(F)$ di molteplicità r se per un aperto affine standard $K^2 \subset \mathbb{P}^2$ contenente P abbiamo che P è di molteplicità r per la curva affine $V(F) \cap K^2$. Analogamente possiamo definire la molteplicità di intersezione con una retta e la retta tangente. I punti singolari sono quelli di molteplicità ≥ 2.

Lemma 12.16. *Un punto P è singolare per la curva $V(F) \subset \mathbb{P}^2$ se e solo se $F_x(P) = F_y(P) = F_z(P) = 0$.*

Dimostrazione. La relazione di Eulero $(deg\ F)F = xF_x + yF_y + zF_z$ mostra che le equazioni $F_x(P) = F_y(P) = F_z(P) = 0$ implicano $F(P) = 0$. Sia $P = (x_0, y_0, z_0)$. Possiamo supporre (a meno di cambiare nome alle coordinate) che $z_0 \neq 0$. Allora poniamo $f(x,y) = F(x,y,1)$, da cui $f_x = F_x$, $f_y = F_y$ e la tesi segue dalla definizione. □

Esercizio 12.17. *Esercizi.*

- *a) Verificare che la curva $x^2 + y^2 - 1 = 0$ è nonsingolare e la sua chiusura proiettiva rimane nonsingolare.*
- *b) Verificare che la curva $y - x^3 = 0$ è nonsingolare mentre la sua chiusura proiettiva ha un punto singolare ("acquista una singolarità all' infinito").*
- *c) Per quali valori di $\lambda \in K$ la curva*

$$x^3 + y^3 + z^3 + 3\lambda xyz = 0$$

è nonsingolare in $\mathbb{P}^2(K)$? Soluzione: per $\lambda \neq -1, -\rho, -\rho^2$ (ρ radice cubica dell'unità) e nei casi esclusi si spezza in 3 rette.

- *d) Per quali valori di $\lambda \in K$ la curva*

$$x^3 + y^3 + z^3 + \lambda(x+y+z)^3 = 0$$

è nonsingolare in $\mathbb{P}^2(K)$? Soluzione: per $\lambda \neq -1/9, -1$.

Per curve proiettive la nonsingolarità può essere espressa anche attraverso alcuni "invarianti". Ad esempio è ben noto che la conica definita da $\sum_{i,j=0}^{2} a_{ij}x_ix_j$ con a_{ij} matrice simmetrica 3×3 è nonsingolare se e solo se $det\ a_{ij} \neq 0$. In generale una curva di grado d è nonsingolare se un certo polinomio omogeneo di grado $3(d-1)^2$ nei coefficienti della curva (detto discriminante) è $\neq 0$.

Gli esercizi seguenti estendono alcuni dei concetti precedenti al caso proiettivo.

Esercizio 12.18. *1. Provare che la retta tangente alla curva $F(x,y,z) \subset \mathbb{P}^2$ nel punto nonsingolare P è data da*

$$xF_x(P) + yF_y(P) + zF_z(P) = 0$$

2. *Sia K algebricamente chiuso. Provare che 2 curve piane in $\mathbb{P}^2(K)$ si incontrano almeno in un punto.*

3. *Provare che un punto nonsingolare $P \in V(F) \subset \mathbb{P}^2$ è un flesso se e solo se*

$$H(F) \quad = \quad det \begin{vmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{xy} & F_{yy} & F_{yz} \\ F_{xz} & F_{yz} & F_{zz} \end{vmatrix} \quad = \quad 0$$

Dedurre che se deg $F = d$ allora $V(F)$ ha al più $3d(d-2)$ flessi.

Una cubica ha al più 9 flessi (sono esattamente 9 se contati in modo opportuno). Un celebre teorema di Newton afferma che la retta congiungente due flessi di una cubica incontra la cubica ancora in un punto di flesso.

13 Metodi effettivi per la diagonalizzazione

Ricordiamo, dal corso di Geometria I, le definizioni e i concetti principali sulle matrici diagonalizzabili. Una matrice A $n \times n$ è diagonalizzabile quando ammette una base di autovettori. Con una formulazione equivalente, A è diagonalizzabile se e solo se esiste G invertibile tale che $G^{-1}AG$ è diagonale. Infatti le colonne di G formano la base richiesta di autovettori.

Una matrice a coefficienti reali è diagonalizzabile (sui reali) se e solo se tutti gli autovalori sono reali (cioè tutte le radici del polinomio caratteristico sono reali) e la molteplicità algebrica di ogni autovalore è uguale alla molteplicità geometrica.

Una matrice a coefficienti complessi è diagonalizzabile se e solo se la molteplicità algebrica di ogni autovalore è uguale alla molteplicità geometrica.

Siccome gli autovalori sono calcolabili solo in modo approssimato, questo algoritmo non è effettivo in aritmetica esatta. Ci proponiamo in questa sezione di trovare un algoritmo effettivo, per verificare la diagonalizzabilità di una matrice, basato sulla nozione di polinomio minimo.

Una matrice con autovalori distinti è diagonalizzabile sui complessi, le matrici (complesse) con autovalori distinti formano un aperto (denso) nell'insieme di tutte le matrici complesse.

Per ogni polinomio $p(x) \in K[x]$ (siamo interessati ai casi $K = \mathbb{R}$ oppure $K = \mathbb{C}$) e ogni matrice A a coefficienti in K ha senso considerare la matrice $p(A)$. Se $p(x) = \sum_{i=0}^{d} a_i x^i$ allora $p(A) = \sum_{i=0}^{d} a_i A^i$, con la convenzione che A^0 è la matrice identità.

Lemma 13.1. *Per ogni matrice invertibile G vale $p(G^{-1}AG) = G^{-1}p(A)G$.*

Dimostrazione.

$$p(G^{-1}AG) = \sum_{i=0}^{d} a_i (G^{-1}AG)^i = \sum_{i=0}^{d} a_i G^{-1}A^iG = G^{-1}(\sum_{i=0}^{d} a_i A^i)G = G^{-1}p(A)G$$

□

Notiamo che per ogni polinomio p, q vale $p(A)q(A) = q(A)p(A) = pq(A)$.

Teorema 13.2 (Hamilton-Cayley)**.** *Ogni matrice a coefficienti reali (o complessi) è radice del polinomio caratteristico $p_A(t) = \det(A - tI)$, cioè $p_A(A) = 0$.*

Dimostrazione. Conviene dimostrare il risultato direttamente per le matrici complesse. Il risultato è immediato per le matrici diagonali, che hanno sulla diagonale gli autovalori, proprio perché gli autovalori sono le radici del polinomio caratteristico. Sia A diagonalizzabile. Allora esiste G tale che $G^{-1}AG = D$ è diagonale. Sappiamo che $p_D(t) = p_A(t)$. Allora dal lemma 13.1 $p_A(A) = p_A(GDG^{-1}) = Gp_A(D)G^{-1} = Gp_D(D)G^{-1} = 0$ e quindi il risultato è vero per le matrici diagonalizzabili.

Adesso la funzione $p_A(A)$ ha coefficienti che sono polinomi nei coefficienti di A, che valgono zero su un aperto denso , pertanto tali polinomi sono identicamente nulli e quindi $p_A(A) = 0 \ \forall A$. □

Fissata A, consideriamo il morfismo

$$K[x] \to M_n$$

$$p(x) \mapsto p(A)$$

Il nucleo di questo morfismo è un ideale di $K[x]$, il quale, come tutti gli ideali di $K[x]$, è principale. Il generatore di questo ideale (normalizzato in modo che sia monico) si dice il *polinomio minimo* di A.

Il teorema di Hamilton-Cayley afferma che il polinomio caratteristico appartiene al nucleo appena definito, quindi il polinomio minimo divide il polinomio caratteristico.

Notiamo anche che A e $G^{-1}AG$ hanno lo stesso polinomio minimo.

Lemma 13.3. *Se v è autovettore di A con autovalore λ e $p(x) \in K[x]$ è un polinomio allora*

$$p(A)v = p(\lambda)v$$

cioè v è autovettore di $p(A)$ con autovalore $p(\lambda)$.

Dimostrazione. Da $Av = \lambda v$ segue $A^k v = \lambda^k v$ per ogni $k \geq 0$. Quindi, se $p(x) = \sum_{i=0}^{d} a_i x^i$ abbiamo $p(A)v = \sum_{i=0}^{d} a_i A^i v = \sum_{i=0}^{d} a_i \lambda^i v = p(\lambda)v$ □

Proposizione 13.4. *Ogni autovalore di A è radice del polinomio minimo di A.*

Dimostrazione. Sia p il polinomio minimo di A, pertanto $p(A) = 0$. Sia λ un autovalore di A con autovettore v. Dal lemma precedente $0 = p(A)v = p(\lambda)v$ da cui $p(\lambda) = 0$. □

Esercizio 13.5. *Fissata A, per ogni $v \in K^n$ definiamo p_v come il generatore (monico) dell'ideale nucleo del morfismo*

$$K[x] \to K^n$$

$$p(x) \mapsto p(A)v$$

(i) Provare che p_v divide il polinomio minimo di A.

(ii) Provare che per ogni base $v_1, \ldots, v_n$, il minimo comune multiplo di p_{v_i} è il polinomio minimo di A.

Il polinomio minimo è ben definito per ogni applicazione lineare $A\colon K^n \to K^n$, nella trattazione seguente adotteremo questo punto di vista. Spesso denotiamo $V = K^n$.

Esercizio 13.6. *(i) Fissata A, sia $W \subset V$ un sottospazio A-invariante. Provare che il polinomio minimo di A ristretta a W divide il polinomio minimo di A.*

(ii) Sia $W_1 \oplus W_2 = V$ una decomposizione in due sottospazi A-invarianti. Per $i = 1, 2$, sia p_i il polinomio minimo di A ristretta a W_i. Provare che il polinomio minimo di A coincide con il minimo comune multiplo tra p_1 e p_2. Generalizzare l'enunciato a una somma di un numero finito di sottospazi.

Se $A\colon V \to V$ è un endomorfismo, e λ è un autovalore di A, abbiamo la catena di inclusioni

$$\ldots \subseteq \ker(A - \lambda I)^n \subseteq \ker(A - \lambda I)^{n+1} \subseteq \ldots$$

Per motivi dimensionali, la catena precedente diventa stazionaria per n sufficientemente grande. Inoltre si può verificare (esercizio) che se $\ker(A-\lambda I)^n = \ker(A-\lambda I)^{n+1}$ allora $\ker(A - \lambda I)^n = \ker(A - \lambda I)^m$ per ogni $m \geq n$.

Definizione 13.7. *Denotiamo $V_\lambda^\infty = \cup_n \ker(A - \lambda I)^n$, che è uguale a $\ker(A - \lambda I)^m$, per qualche m sufficientemente grande. Vedremo che si può sempre sempre scegliere $m \leq$ molteplicità algebrica di λ.*

Prima della Prop. 13.9 premettiamo il seguente (provvisorio)

Lemma 13.8. $\dim V_\lambda^\infty \leq$ *molteplicità algebrica di λ.*

Dimostrazione. V_λ^∞ è un sottospazio A-invariante. Il polinomio minimo di A ristretto a V_λ^∞ ha la forma $(x-\lambda)^m$ per qualche m, pertanto per la Proposizione 13.4 il polinomio caratteristico di A ristretto a V_λ^∞ è uguale a $(x - \lambda)^d$ dove $d = \dim V_\lambda^\infty$. Segue che il polinomio caratteristico di A è divisibile per $(x - \lambda)^d$ da cui la tesi. □

Proposizione 13.9. *Sia $A\colon V \to V$ è un endomorfismo. Se $K = \mathbb{C}$ abbiamo $V = \oplus_\lambda V_\lambda^\infty$ dove la somma è estesa a tutti gli autovalori. Inoltre* $\dim V_\lambda^\infty$ *è uguale alla molteplicità algebrica di λ. V_λ^∞ si dice autospazio generalizzato.*

Dimostrazione. Sia $f(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$ il polinomio minimo di A, fattorizzato nella chiusura algebrica. Poniamo $p_j(x) = \prod_{i \neq j} (x - \lambda_i)^{n_i}$, che sono primi tra loro. Pertanto esistono $a_j(x)$ tali che $\sum_{j=1}^k a_j(x)p_j(x) = 1$. Applicando i polinomi alla matrice A si ottiene $\sum_{j=1}^k a_j(A)p_j(A) = I$. Notiamo che $\operatorname{Im} p_j(A) \subset V_{\lambda_j}^\infty$ perché $(A-\lambda_j I)^{n_j} p_j(A) = f(A) = 0$. Per ogni vettore $v \in V$ abbiamo $v = \sum_{j=1}^k p_j(A)a_j(A)v$, dove $p_j(A)a_j(A)v \in V_{\lambda_j}^\infty$ per quanto appena visto. Segue che V si decompone nella somma dell'enunciato. Per provare che la somma è diretta, per il Lemma 13.8, la dimensione della somma dei $V_{\lambda_j}^\infty$ è minore o uguale della somma delle molteplicità algebriche, che è pari a $\dim V$, la dimensione dello spazio ambiente. Pertanto nel Lemma 13.8 deve valere l'uguaglianza. Inoltre la dimensione della somma dei $V_{\lambda_j}^\infty$ è uguale alla somma delle dimensioni dei $V_{\lambda_j}^\infty$, e questo accade precisamente quando la somma è diretta. □

Corollario 13.10. *Sia n_i il più piccolo intero tale che* $\ker(A-\lambda_i I)^{n_i} = \ker(A-\lambda_i I)^{n_i+1}$*, allora il polinomio minimo di A è* $\prod_{i=1}^{k}(x-\lambda_i)^{n_i}$

Dimostrazione. E' sufficiente provare che il polinomio minimo di A ristretto a ogni V_λ^∞ è uguale a $(x-\lambda_i)^{n_i}$, che è evidente dalla definizione. □

Teorema 13.11. *Una matrice è diagonalizzabile su K se e solo se il suo polinomio minimo ha tutte le radici in K di molteplicità uno (si spezza come prodotto di fattori lineari distinti).*

Dimostrazione. Se A è diagonalizzabile ha lo stesso polinomio minimo di una matrice diagonale D. Se D ha sulla diagonale gli elementi distinti $d_1, \dots d_k$ (nel senso che alcuni di questi valori possono apparire più volte sulla diagonale) allora $\prod_{i=1}^{k}(x-d_i)$ è il polinomio minimo di D. Viceversa, se il polinomio minimo di A ha tutte le radici di molteplicità uno allora dal Corollario 13.10 segue che $\ker(A-\lambda_i I) = V_{\lambda_i}^\infty$ e dalla Prop. 13.9 segue che A ha una base di autovettori. □

Proposizione 13.12. *Se λ_i per $i = 1, \dots, k$ sono tutti gli autovalori (complessi) di una matrice A, ripetuti con la loro molteplicità n_i, allora, per ogni polinomio $h(x)$, $h(\lambda_i)$ per $i = 1, \dots, k$ sono tutti gli autovalori (complessi) della matrice $h(A)$, ripetuti n_i volte.*

Dimostrazione. La decomposizione $V = \oplus_{\lambda_i} V_{\lambda_i}^\infty$ della Prop. 13.9 è $h(A)$-invariante. Sia n_i la dimensione di $V_{\lambda_i}^\infty$.

Se $v_i \in V_{\lambda_i}^\infty$ abbiamo $(A-\lambda_i)^{n_i} v_i = 0$. Siccome $(x-\lambda_i)$ divide $h(x)-h(\lambda_i)$, segue $(h(A)-h(\lambda_i)I)^{n_i} v_i = 0$, da cui il polinomio minimo di $h(A)$ ristretto a $V_{\lambda_i}^\infty$ ha la forma $(x-h(\lambda_i))^m$ per qualche m, e per la prop. 13.4 $h(\lambda_i)$ è l'unico autovalore di $h(A)$ ristretto a $V_{\lambda_i}^\infty$. Segue che il polinomio caratteristico di $h(A)$ ristretto a $V_{\lambda_i}^\infty$ è $(x-h(\lambda_i))^{n_i}$. □

Osservazione 13.13. *Una dimostrazione alternativa della Prop. 13.12 può essere ottenuta prendendo una forma triangolare per A.*

Algoritmo per il calcolo del polinomio minimo Si scelga il primo valore d tale che $A^0, A^1, A^2, \dots, A^d$ sono dipendenti. Allora esiste un unico vettore (a meno di costanti) $(a_0, \dots, a_d)$ tale che $\sum_{i=0}^{d} a_i A^i = 0$ (se ci fossero due tali vettori, entrambi dovrebbero avere $a_d \neq 0$, ed una loro combinazione lineare darebbe una relazione di dipendenza tra $A^0, A^1, \dots, A^{d-1}$).

Il polinomio $p(t) = \sum_{i=0}^{d} \frac{a_i}{a_d} t^i$ è il polinomio minimo di A.

Algoritmo per la verifica della diagonalizzabilità di una matrice, sui complessi Sia data una matrice A. Si costruisce p polinomio minimo di A con l'algoritmo precedente.

Si calcola MCD (p, p') con l'algoritmo euclideo, A è diagonalizzabile se e solo se MCD $(p, p') = 1$.

Spiegazione: p non ha fattori ripetuti se e solo se MCD $(p, p') = 1$ e questo equivale alla diagonalizzabilità per il teorema precedente.

Esercizio 13.14. *Sia A matrice $n \times n$ il cui polinomio minimo ha grado n e che ha tutti gli autovalori in K (una tale A si dice regolare). Provare che esiste $v \in K^n$ tale che $v, Av, A^2v, \dots, A^{n-1}v$ sono indipendenti.*

Suggerimento: provare prima il caso in cui il polinomio minimo ha la forma $(x-\lambda)^n$. In generale, se $p = \prod p_i^{n_i}$ con p_i primi tra loro, si può provare con una tecnica simile a quella usata nella dimostrazione del Teorema 13.11 che $K^n = \oplus_i \ker p_i(A)$ e ci si può ricondurre al caso precedente.

Esercizio 13.15. *Sia A una matrice regolare (definita nell'esercizio precedente). Provare che il centralizzante di A ha dimensione n ed è generato da $\{A^0, A^1, \dots, A^{n-1}\}$.*

Suggerimento: Il polinomio minimo di A è dato da $p_1^{n_1} \dots p_k^{n_k}$ con p_i polinomi di grado uno distinti. Posto $V_i = \ker p_i(A)$ si può provare (vedi dimostrazione del Teorema 13.11) che $K^n = \oplus_i V_i$. Se B commuta con A allora ogni V_i è B-invariante. Quindi per dimostrare l'asserto ci si può ricondurre al caso in cui il polinomio minimo di A ha la forma p^n con p di grado uno. In questo caso l'esercizio precedente mostra che abbiamo un unico blocco di Jordan. La conoscenza della forma di Jordan è utile ma non indispensabile per la comprensione di queste note. Il lettore che conosce la forma di Jordan può osservare che le matrici compagne, che introdurremo nella sezione 2, sono regolari e quindi hanno un unico blocco di Jordan per ogni autovalore. Applicando l'esercizio precedente ad $A - I$, otteniamo che esiste $v \in K^n$ tale che $v, (A-I)v, (A-I)^2v, \dots, (A-I)^{n-1}v$ sono indipendenti. Un calcolo esplicito, scrivendo la matrice di A rispetto a questa base, mostra che la condizione $AB = BA$ impone $n^2 - n$ condizioni indipendenti su B. Pertanto lo spazio vettoriale generato da $\{A^0, A^1, \dots, A^{n-1}\}$, che è sempre contenuto nel centralizzante di A, deve coincidere col centralizzante di A perché ha la stessa dimensione n.

14 Polinomi in una variabile e matrici compagne

In questa sezione ricordiamo come il calcolo delle radici di un polinomio può essere ricondotto al calcolo degli autovalori di una matrice, detta matrice compagna. Questo è interessante perché esistono molti algoritmi per calcolare numericamente gli autovalori di una matrice.

La struttura algebrica che nasce da questa problema ha delle generalizzazioni nel caso di più variabili, che studieremo successivamente a partire dalla sezione 17.

Sia $f(x) = \sum_{i=0}^{d} a_i x^i$ un polinomio di grado d, che possiamo supporre monico, cioè $a_d = 1$. Chiamiamo R l'anello quoziente $K[x]/(f(x))$ che ha dimensione d ed è generato dalle classi $[1], [x], \dots, [x^{d-1}]$. Infatti x^d può essere scritto come combinazione delle potenze precedenti modulo f, $[x^d] = -\sum_{i=0}^{d-1} a_i [x^i]$ e così anche le potenze successive, ad esempio

$[x^{d+1}] = -\sum_{i=0}^{d-1} a_i [x^{i+1}] = -\sum_{i=0}^{d-2} a_i [x^{i+1}] + a_{d-1} \sum_{i=0}^{d-1} a_i [x^i]$

Definizione 14.1. *La moltiplicazione per x induce un'applicazione lineare*

$$R \xrightarrow{M_x} R$$

$$[g] \mapsto [gx]$$

la matrice di M_x rispetto alla base $[1], [x], \ldots, [x^{d-1}]$ si dice matrice compagna di $f(x)$.

Analogamente la moltiplicazione per $h(x)$ induce un'applicazione lineare

$$R \xrightarrow{M_{h(x)}} R$$

$$[g] \mapsto [gh]$$

Proposizione 14.2. *$\forall\ h(x), k(x) \in K[x]$ vale che*
(i) $M_{h(x)} + M_{k(x)} = M_{h(x)+k(x)}$
(ii) $M_{ah(x)} = aM_{h(x)} \quad \forall a \in K$
(iii) $M_{h(x)} \cdot M_{k(x)} = M_{k(x)} \cdot M_{h(x)} = M_{h(x)k(x)}$
(iv) $M_{h(k(x))} = h(M_{k(x)})$, in particolare $M_{h(x)} = h(M_x)$.

Dimostrazione. (i), (ii) e (iii) sono immediate dalla definizione. Notiamo che la commutatività in (iii) segue dalla commutatività dei polinomi. Per provare (iv) supponiamo dapprima $h = x^i$. Allora $M_{(k(x))^i} = \left(M_{k(x)}\right)^i$ come diretta applicazione di (iii).

In generale, se $h(x) = \sum a_i x^i$ allora utilizzando anche (i)-(iii)

$$M_{h(k(x))} = M_{\sum a_i (k(x))^i} = \sum M_{a_i (k(x))^i} = \sum a_i M_{(k(x))^i} = \sum a_i \left(M_{(k(x))}\right)^i = h(M_{k(x)})$$

□

Proposizione 14.3. *La matrice compagna di $f(x)$ è*

$$\begin{bmatrix} 0 & 0 & \ldots & -a_0 \\ 1 & 0 & \ldots & -a_1 \\ 0 & 1 & \ldots & -a_2 \\ \vdots & & & \vdots \\ 0 & 0 & 1 & -a_{d-1} \end{bmatrix}$$

Dimostrazione. È un calcolo immediato dalle espressioni precedenti. □

Teorema 14.4. *(i) A meno di scalari moltiplicativi, $f(x)$ è il polinomio minimo di M_x.*
(ii) A meno di scalari moltiplicativi, $f(x)$ è il polinomio caratteristico di M_x.
(iii) M_x è diagonalizzabile su $\mathbb{C}$ se e solo se $f(x)$ ha radici distinte.
(iv) $f(x_0) = 0$ se e solo se x_0 è un autovalore di M_x.
(v) $tr(M_{h(x)}) = \sum_{i=1}^{n} h(\lambda_i)$ dove $\lambda_1, \ldots, \lambda_n$ sono le radici di $f(x)$
(vi) $\det(M_{h(x)}) = \prod_{i=1}^{n} h(\lambda_i)$, in particolare $\det(M_{h(x)})$ si annulla se e solo se f e h hanno una radice in comune, e costituisce un metodo alternativo (in generale più economico) di calcolo del risultante $Res(f, h)$.

Dimostrazione. Per ogni polinomio $p(x)$ abbiamo che $p(M_x) = M_{p(x)}$, che è nulla esattamente quando $p(x)$ è un multiplo di $f(x)$ (basta applicare $M_{p(x)}$ alla classe di 1). Pertanto $f(x)$ è il polinomio minimo di M_x, e siccome divide il polinomio caratteristico, che ha lo stesso grado, segue che polinomio minimo e polinomio caratteristico

coincidono, a meno del segno. (iii) è conseguenza di (i) e del Teorema 13.11. (iv) è immediata da (ii). (v) e (vi) seguono dalla Prop. 13.12.

□

Osservazione L'importanza del Teorema 14.4 risiede nel fatto che otteniamo le radici del polinomio f dal calcolo degli autovalori della matrice compagna, che può essere effettuato ad esempio col metodo QR.

15 La forma di Killing e il numero delle radici reali di un polinomio

15.1 Il teorema cinese dei resti e l'interpolazione polinomiale

Consideriamo la fattorizzazione di un intero $n = \prod_{i=1}^{k} p_i^{m_i}$ con p_i primi distinti. La forma classica del teorema cinese dei resti, su $\mathbb{Z}$, descrive l'isomorfismo

$$q\colon \mathbb{Z}/n \to \oplus_{i=1}^{k} \mathbb{Z}/(p_i^{m_i})$$

L'estensione naturale ai polinomi è data dal seguente

Teorema 15.1. *[Teorema cinese dei resti per polinomi] Sia $f(x) = \prod_{i=1}^{k} f_i^{n_i}$ con f_i irriducibili distinti, dunque $f_i^{n_i}$ primi tra loro a due a due. Sia $g_i = \frac{f}{f_i^{n_i}}$, siano b_i tali che $\sum_{i=1}^{k} b_i g_i = 1$ (ricavabili con l'algoritmo euclideo, vedi l'osservazione 15.2). L'applicazione naturale*

$$q\colon K[x]/(f(x)) \to \oplus_{i=1}^{k} K[x]/(f_i^{n_i})$$

definita da $q(h) = (h, \dots, h)$ è un isomorfismo con inversa data da

$$\tilde{q}\colon \oplus_{i=1}^{k} K[x]/(f_i^{n_i}) \to K[x]/(f(x))$$

definita da $\tilde{q}(a_1, \dots, a_k) = \sum_{i=1}^{k} b_i g_i a_i$.

Dimostrazione. $b_i g_i$ modulo $f_j^{n_j}$ è uguale a 1 se $i = j$, è uguale a 0 se $i \neq j$. Pertanto $\sum_{i=1}^{k} b_i g_i a_i$ modulo $f_j^{n_j}$ è uguale ad a_j $\forall j$. Da queste condizioni si ricava che, componendo $q\tilde{q}$ e $\tilde{q}q$, si ottiene in entrambi i casi l'identità. □

Osservazione 15.2. *Con Macaulay2, b_i puo' essere trovato tramite il comando*

```
quotientRemainder(matrix{{1}},matrix{{g_1,...g_k}})
```

Notiamo che per calcolare questi quozienti è necessario trovare la base di Groebner dell'ideale $(g_1, \dots, g_k)$, che è uguale a 1, e che essenzialmente questo calcolo corrisponde all'algoritmo euclideo del calcolo del MCD tra i g_i.

La seguente applicazione all'interpolazione polinomiale è interessante.

Corollario 15.3 (Interpolazione di Hermite, facoltativa)**.** *Siano $c_1, \dots, c_k \in \mathbb{R}$ punti distinti. Assegniamo, rispetto a ogni punto c_i, uno sviluppo di Taylor $a_i(x)$ di grado $< m_i$, questo equivale a dare $a_i(x) = \sum_{j=0}^{m_i-1} \frac{\alpha_{j,i}}{j!}(x-c_i)^j$ dove $\alpha_{j,i} = a_i^{(j)}(c_i)$ è la derivata j-esima di a_i valutata in c_i.*

Esiste un polinomio $H(x)$ di grado $< \sum_{i=1}^k m_i$, tale che $H^{(j)}(c_i) = \alpha_{j,i}$ per $0 \le j < m_i$, definito da $H = \sum_{i=1}^k b_i g_i a_i$ modulo $f(x) = \prod_{i=1}^k (x-c_i)^{m_i}$ (cioè $H\%ideal(f)$) dove , posto $f_i = (x-c_i)^{m_i}$, $g_i = \prod_{j\neq i}(x-c_j)^{m_j}$, b_i viene ottenuto dall'algoritmo euclideo come nel Teorema 15.1, cioè dalla condizione $b_i g_i + r_i f_i = 1$.

Dimostrazione. La condizione $H(x) = a_i(x)$ modulo $(x-c_i)^{m_i}$, equivale a $H^{(j)}(c_i) = a_i^{(j)}(c_i)$. Il risultato segue allora dal Teor. 15.1. □

Osservazione 15.4. *Il polinomio $H(x) = \sum_{i=1}^k b_i(x) g_i(x) a_i(x)$ ha le derivate richieste nei punti c_i. Considerare la sua forma normale rispetto a f serve soltanto per abbassare il grado fino a renderlo $< \sum_{i=1}^k m_i = \deg f$.*

Per $m_i = 1$ abbiamo come caso particolare l' interpolazione di Lagrange, dove si ricava un polinomio di grado $k-1$ che assume k valori assegnati.

Corollario 15.5 (Interpolazione di Lagrange)**.** *Siano $c_1, \dots, c_k \in \mathbb{R}$ punti distinti. Assegniamo, rispetto a ogni punto c_i, un valore $a_i \in K$.*

Esiste un polinomio $H(x)$ di grado $< k$, tale che $H(c_i) = a_i$, definito da $H = \sum_{i=1}^k b_i g_i a_i$ dove $g_i = \prod_{j\neq i}(x-c_j)$, $b_i = \frac{1}{\prod_{j\neq i}(c_i-c_j)}$.

Dimostrazione. Rispetto all'interpolazione di Hermite, si può scegliere b_i come lo scalare nell'enunciato, infatti $b_i g_i = \frac{\prod_{j\neq i}(x-c_j)}{\prod_{j\neq i}(c_i-c_j)}$ vale 1 su c_i e vale 0 su c_j per $j \neq i$. Il grado di $\sum_{i=1}^k b_i g_i a_i$ risulta immediatamente $< k$. □

L'interpolazione di Hermite è utile nello sviluppo di una funzione razionale in fratti semplici, necessaria per la sua integrazione. Infatti, con le notazioni precedenti, dopo aver trovato un'espressione $1 = \sum_{i=1}^k b_i g_i$, dividendo per $f(x)$ si ottiene $\frac{1}{f(x)} = \sum_{i=1}^k \frac{b_i(x)}{(x-c_i)^{m_i}}$ da cui per un polinomio g

$$\frac{g(x)}{f(x)} = \sum_{i=1}^k \frac{g(x) b_i(x)}{(x-c_i)^{m_i}}$$

Calcolando lo sviluppo di Taylor dei numeratori $g(x)b_i(x)$ rispetto al punto c_i, si ottiene l'espressione che puo' essere facilmente integrata.

Esercizio 15.6. *Si trovi, con l'ausilio di Macaulay2, un polinomio di grado 8 tale che $H(2) = \alpha$, $H'(2) = \beta$, $H(3) = 5$, $H'(3) = 7$, $H''(3) = 11$, $H(4) = 13$, $H'(4) = 17$, $H''(4) = 19$, $H'''(4) = 21$, al variare di α, β.*

15.2 La forma traccia di Killing

Sia $f(x) \in K[x]$ un polinomio. Per ogni $a, b \in K[x]/(f(x))$ è definita la forma (bilineare) B di Killing

$$\begin{array}{ccccc} K[x]/(f(x)) & \times & K[x]/(f(x)) & \longrightarrow & K \\ a & & b & \mapsto & B(a,b) := tr(M_{ab}) \end{array}$$

dove M_{ab} è l'applicazione lineare $K[x]/(f(x)) \xrightarrow{M_{ab}} K[x]/(f(x))$ data dalla moltiplicazione per ab. Notiamo che $M_a M_b = M_b M_a = M_{ab}$.

È associata la forma quadratica $a \mapsto B(a,a) = tr(M_{a^2})$.

Proposizione 15.7. *Sia $n = \deg f$. La matrice della forma di Killing nella base $\{1, x, \dots x^{n-1}\}$ ha la forma*

$$\begin{bmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & \dots & & s_{n+1} \\ \vdots & & & & \vdots \\ s_{n-1} & s_n & \dots & & s_{2n-2} \end{bmatrix}$$

dove s_i è la i-esima somma di potenze nelle radici $x_1, \dots x_n$ di f, cioè $s_i = \sum_{j=1}^{n} x_j^i$ e prende il nome di Bezoutiante. Numerando le righe e le colonne da 0 a $n-1$, il coefficiente di posto (i,j) è s_{i+j}.

Dimostrazione. Numeriamo le righe e le colonne da 0 a $n-1$. Allora il coefficiente di posto (i,j) è $tr(M_{x^{i+j}})$. Per il Teorema 14.4 (iv) gli autovalori di M_x sono le radici $x_1, \dots, x_n$ di $f(x)$ e quindi gli autovalori di $M_{x^{i+j}} = (M_x)^{i+j}$ sono $x_1^{i+j}, \dots, x_n^{i+j}$ (per il Teor. 14.4 (v)) e la loro somma coincide con s_{i+j}. □

Il Teorema di Sylvester 15.9 è il risultato fondamentale di questa area. Lega il numero di radici reali di $f(x)$ con la segnatura della forma di Killing associata a $f(x)$. Prima di enunciarlo, abbiamo

Teorema 15.8. *I sottospazi di $K[x]/(f(x))$ isomorfi a $K[x]/(p_i(x)^{n_i})$, identificati tramite l'isomorfismo dato dal Teorema cinese 15.1, sono ortogonali a due a due rispetto alla forma di Killing B.*

Dimostrazione. Due sottospazi sono generati rispettivamente da $b_i(x)g_i(x)$ e da $b_j(x)g_j(x)$ Abbiamo che $f(x)$ divide $g_i(x)g_j(x)$ e quindi l'applicazione lineare $M_{b_i(x)g_i(x)}M_{b_j(x)g_j(x)}$ è nulla in $K[x]/(f(x))$, pertanto la sua traccia è nulla. Questo prova che $\forall a \in K[x]/(p_i(x)^{n_i}), b \in K[x]/(p_j(x)^{n_j})$, vale $B(a,b) = 0$, come volevamo. □

Osservazione Distinguiamo due casi.

1. Gli autospazi generalizzati della matrice compagna M_x corrispondono agli addendi $\mathbb{R}[x]/(p_i(x)^{n_i})$ quando $p_i(x) = x - c_i$, cioè corrispondono alle radici c_i di f. Se la molteplicità della radice c_i è 1 allora l'autospazio generalizzato coincide con l'autospazio generalizzato relativo a c_i. Infatti $(M_x - c_i I)^{n_i} = M_{(x-c_i)^{n_i}} = 0$.

2. Quando $p_i(x)$ è un polinomio di secondo grado con una coppia di radici complesse coniugate $\{\alpha_i, \overline{\alpha_i}\}$, $\mathbb{R}[x]/(p_i(x)^{n_i})$ è M_x-invariante, ed è somma dei due autospazi generalizzati corrispondenti alla coppia di radici su $\mathbb{C}$.

In entrambi i casi $\mathbb{R}[x]/(p_i(x)^{n_i})$ è un *anello locale*, con unico ideale massimale generato dalla classe di $p_i(x)$. Il quoziente rispetto all'ideale massimale è isomorfo a $\mathbb{R}$ nel primo caso ed a $\mathbb{C}$ nel secondo caso. Gli elementi in $\mathbb{R}[x]/(p_i(x)^{n_i})$ possono essere riguardati come sviluppi di Taylor in $x = c_i$ nel primo caso e come una coppia di sviluppi di Taylor coniugati rispetto a $x = \alpha_i$ e $x = \overline{\alpha_i}$ nel secondo caso. In particolare gli elementi invertibili di $\mathbb{R}[x]/(p_i(x)^{n_i})$ corrispondono nel primo caso alle classi dei polinomi $q(x) \in \mathbb{R}[x]$ tali che $q(c_i) \neq 0$, mentre nel secondo caso alle classi dei polinomi $q(x) \in \mathbb{R}[x]$ tali che $q(\alpha_i) \neq 0$ (e dunque anche $q(\overline{\alpha_i}) \neq 0$.

$\mathbb{R}[x]/(p_i(x)^{n_i})$ è uno spazio vettoriale su $\mathbb{R}$ di dimensione n_i nel primo caso e di dimensione $2n_i$ nel secondo caso.

15.3 Il numero di radici reali

Teorema 15.9. *[**Sylvester**] Sia B la matrice Bezoutiante di f, cioè la matrice della forma di Killing associata.*

(i) f ha n radici reali e distinte se e solo se B è definita positiva.

(ii) f ha tutte le radici reali se e solo se B è semidefinita positiva.

(iii) Il rango di B è il numero di radici (reali o complesse) distinte di f.

(iv) il numero di radici reali (distinte) di f è uguale al numero di autovalori positivi di B meno il numero di autovalori negativi di B.

Dimostrazione. Il Teorema 15.8 permette di ricondurre il calcolo della segnatura della forma di Killing su $K[x]/(f(x))$ a quello della segnatura su ogni addendo $K[x]/(p_i(x)^{n_i})$. Tutte le applicazioni lineari della forma

$$M_{p(x)} \colon K[x]/(p(x)^n) \to K[x]/(p(x)^n)$$

sono nilpotenti, quindi hanno tutti gli autovalori nulli e la traccia nulla.

Nel caso in cui $K = \mathbb{R}$ abbiamo due casi da distinguere, dove $p(x) = x - c$ (radice reale) oppure dove $p(x) = x^2 + ax + b$ con $a^2 - 4b < 0$ (due radici complesse coniugate).

Nel primo caso, l'anello $K[x]/((x-c)^n)$ ha la base $1, x-c, (x-c)^2, \ldots, (x-c)^{n-1}$ e per quanto visto la matrice della forma di Killing ha tutti gli elementi nulli escluso quello in alto a sinistra, dove vale n, che è la traccia dell'identità. Pertanto il suo rango è 1 e la sua segnatura è 1, cioè ha un solo autovalore positivo e nessun autovalore negativo.

Nel secondo caso, l'anello $K[x]/((x^2+ax+b)^n)$ ha la base $1, x, (x^2+ax+b), x(x^2+ax+b), (x^2+ax+b)^2 \ldots, x(x^2+ax+b)^{n-1}$ e per quanto visto la matrice della forma di Killing ha tutti gli elementi nulli escluso quelli del blocco 2×2 in alto a sinistra, che corrisponde a

$$\begin{bmatrix} tr(1) & tr(M_x) \\ tr(M_x) & tr(M_{x^2}) \end{bmatrix}$$

Adesso $tr(1) = 2n$. Gli autovalori di M_x sono le radici di $(x^2 + ax + b)^n$. Chiamo $\alpha, \overline{\alpha}$ le due radici di $x^2 + ax + b = 0$, da cui $tr(M_x) = n(\alpha + \overline{\alpha})$, $tr(M_{x^2}) = n(\alpha^2 + \overline{\alpha}^2)$.

Quindi la segnatura della forma di Killing su $K[x]/((x^2+ax+b)^n)$ equivale alla segnatura della matrice

$$n\begin{bmatrix} 2 & \alpha+\overline{\alpha} \\ \alpha+\overline{\alpha} & \alpha^2+\overline{\alpha}^2 \end{bmatrix} = n\begin{bmatrix} 1 & 1 \\ \alpha & \overline{\alpha} \end{bmatrix}\begin{bmatrix} 1 & \alpha \\ 1 & \overline{\alpha} \end{bmatrix}$$

Dividendo la matrice per n (questo non modifica la segnatura), il determinante è $(\overline{\alpha}-\alpha)^2 = (-2i\cdot Im(\alpha))^2 = -4(Im\alpha)^2 < 0$, da cui abbiamo un autovalore positivo e un autovalore negativo. Inoltre il rango è 2.

Sommando i contributi di tutti gli addendi, si ottiene la segnatura della forma di Killing come nel teorema di Sylvester. □

La segnatura di una forma quadratica può essere determinata facilmente, ispezionando il polinomio caratteristico. Infatti è noto che tutti gli autovalori di una matrice simmetrica reale sono reali, e si può applicare la

Teorema 15.10 (Regola di Cartesio sui segni)**.** *Sia $p(x)$ un polinomio a coefficienti reali con tutte le radici reali. Il numero delle radici positive è uguale al numero delle variazioni di segno tra due suoi coefficienti non nulli consecutivi. Radici multiple sono contate quanto la loro molteplicità.*

Per una dimostrazione si veda [Aba]. Nel teorema 15.12 vedremo un modo di contare le radici positive di un polinomio anche senza l'ipotesi che tutte le radici siano reali.

Riguardo la regola di Cartesio, siccome la molteplicità della radice nulla è uguale all'esponente del più piccolo monomio che appare, possono essere calcolate esattamente il numero delle radici positive, nulle e negative (queste ultime per differenza).

Notiamo che M_x è simmetrico rispetto alla forma traccia di Killing, nel senso che $B(M_x a, b) = tr(M_{xab}) = B(a, M_x b)$. Attenzione, perché non è lecito applicare il teorema spettrale a meno che la forma non sia definita positiva, e infatti se la forma è definita positiva sappiamo, grazie al Teor. 14.4 e al Teor. 15.9 che f ha n radici reali distinte e che M_x è diagonalizzabile, in accordo col teorema spettrale.

Osservazione 15.11. *Supponiamo che f abbia radici $x_1, \dots x_n$ tutte reali. Definiamo la matrice di Vandermonde*

$$V = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ \vdots & & & \vdots \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix}$$

Un facile calcolo mostra che $V^t \cdot V$ coincide con la matrice Bezoutiante. Questa osservazione fornisce un argomento alternativo al fatto che la Bezoutiante è semidefinita positiva se f ha tutte le radici reali ed è definita positiva se le radici reali sono anche distinte. Questa è stata la strada seguita nel XIX secolo da Sylvester.

Per ogni $a, b \in K[x]/(f(x))$, $h \in K[x]$ è definita la forma (bilineare) B_h (generalizzazione della forma di Killing)

$$\begin{array}{ccccc} K[x]/(f(x)) & \times & K[x]/(f(x)) & \longrightarrow & K \\ a & & b & \mapsto & tr(M_{hab}) \end{array}$$

È associata la forma quadratica $a \mapsto tr(M_{ha^2})$. La funzione $h(x)$ va pensata come una sorta di "funzione test", scegliendo $h(x)$ opportune di gradi 1 o 2 si possono avere informazioni rilevanti sulla localizzazione delle radici di f secondo la seguente proposizione, che generalizza il teorema di Sylvester. La forma di Killing B definita in precedenza corrisponde a B_1 (cioè con $h = 1$).

Teorema 15.12. *Per un polinomio reale $h(x)$ di grado n, sia B_h la matrice della forma definita da $B_h(a, b) = tr(M_{hab})$ per ogni $a, b \in K[x]/(f(x))$. Notiamo che, rispetto alle notazioni precedenti, $B_1 = B$.*

(i) f ha n radici reali distinte p tali che $h(p) > 0$ se e solo se B_h è definita positiva.

(ii) Il rango di B_h è il numero di radici (reali o complesse) distinte p di f tali che $h(p) \neq 0$.

(iii) il numero di radici reali (distinte) di f tali che $h(p) > 0$ meno il numero di radici reali (distinte) di f tali che $h(p) < 0$ è uguale alla segnatura di B_h.

Inoltre supponiamo che $h(p) \neq 0$ $\forall p \in V(I)$, ipotesi soddisfatta se h è primo con f.

(iv) il numero di radici reali (distinte) di f tali che $h(p) > 0$ è uguale al numero di autovalori positivi di B_h meno il numero di autovalori negativi di B.

(v) il numero di radici reali (distinte) di f tali che $h(p) < 0$ è uguale al numero di autovalori negativi di B_h meno il numero di autovalori negativi di B.

Dimostrazione. Ancora riconduciamo il calcolo della segnatura di B_h su $K[x]/(f(x))$ a quello della segnatura su ogni addendo $K[x]/(p_i(x)^{n_i})$, che sono ancora ortogonali, per lo stesso ragionamento fatto per la forma di Killing B.

Nel caso in cui $K = \mathbb{R}$ abbiamo due casi da distinguere, dove $p(x) = x - c$ (radice reale) oppure dove $p(x) = x^2 + ax + b$ con $a^2 - 4b < 0$ (due radici complesse coniugate).

Nel primo caso, l'anello $K[x]/((x-c)^n)$ ha la base $1, x-c, (x-c)^2, \ldots, (x-c)^{n-1}$ e per quanto visto la matrice della forma B_h, ristretta al sottospazio $K[x]/((x-c)^n)$, ha tutti gli elementi nulli escluso quello in alto a sinistra, dove vale $\mathrm{tr}(M_h) = \mathrm{tr}\,(h(M_x))$, per la Prop. 14.2 (iv). Siccome M_x ha il solo autovalore c, con molteplicità algebrica n, la traccia di $h(M_x)$ vale $nh(c)$. Pertanto, il rango di B_h, ristretta al sottospazio $K[x]/((x-c)^n)$, vale 1 se $h(c) \neq 0$ e vale 0 se $h(c) = 0$, la sua segnatura è 1 se $h(c) > 0$ mentre è -1 se $h(c) < 0$.

Nel secondo caso, l'anello $K[x]/((x^2+ax+b)^n)$ ha la base $1, x, (x^2+ax+b), x(x^2+ax+b), (x^2+ax+b)^3 \ldots, x(x^2+ax+b)^{n-1}$ e per quanto visto la matrice della forma B_h ha tutti gli elementi nulli escluso quelli del blocco 2×2 in alto a sinistra, che corrisponde a

$$\begin{bmatrix} tr(M_h) & tr(M_{xh}) \\ tr(M_{xh}) & tr(M_{x^2h}) \end{bmatrix}$$

Infatti la matrice 2×2 può essere scritta come

$$\begin{bmatrix} h(\alpha)+h(\overline{\alpha}) & h(\alpha)\alpha+h(\overline{\alpha})\overline{\alpha} \\ h(\alpha)\alpha+h(\overline{\alpha})\overline{\alpha} & h(\alpha)\alpha^2+h(\overline{\alpha})\overline{\alpha}^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \alpha & \overline{\alpha} \end{bmatrix} \cdot \begin{bmatrix} h(\alpha) & 0 \\ 0 & h(\overline{\alpha}) \end{bmatrix} \cdot \begin{bmatrix} 1 & \alpha \\ 1 & \overline{\alpha} \end{bmatrix} = U^t D U \tag{15.1}$$

e ha quindi rango due se $h(\alpha) \neq 0$ e α ha parte immaginaria non nulla. Quando questa condizione è soddisfatta, la sua segnatura è $(1,1)$ perché il suo determinante è negativo in quanto vale $(\det U)^2 \det D$ ed abbiamo $(\det U)^2 = (-2i \cdot Im\alpha)^2 < 0$, $\det D = h(\alpha)h(\overline{\alpha}) = h(\alpha)\overline{h(\alpha)} = |h(\alpha)|^2 > 0$.

□

Nel caso $h(x) = x$, il Teorema 15.12 permette di calcolare le radici positive. Nel caso $h(x) = x - a$, il Teorema 15.12 permette di calcolare le radici $> a$.

Osservazione 15.13. *E' possibile calcolare il numero di radici contenute in un qualunque intervallo. Infatti il numero di radici nell'intervallo $(a, b]$ si ottiene sottraendo dal numero di radici $> a$ quelle che sono $> b$.*

La dimostrazione della proposizione seguente è formalmente identica a quella della Prop. 15.7 e viene omessa.

Proposizione 15.14. *Sia $n = \deg f$. La matrice della forma B_x nella base $\{1, x, \ldots x^{n-1}\}$ è*

$$B_x = \begin{bmatrix} s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ s_3 & s_4 & \cdots & & s_{n+2} \\ \vdots & & & & \vdots \\ s_n & s_{n+1} & \cdots & & s_{2n-1} \end{bmatrix}$$

Numerando le righe e le colonne da 0 a $n-1$, il coefficiente di posto (i,j) di B' è s_{i+j+1}.

Osservazione 15.15. *La matrice di B_h nella base standard $\{1, x, \ldots, x^{n-1}\}$ si può calcolare nel modo seguente. Se $h(x) = \sum_k a_k x^k$ allora al posto (i,j) (con la solita numerazione da 0 a $n-1$) abbiamo $\sum_k a_k s_{i+j+k}$ (dove s_i è la i-esima somma di potenze nelle radici di f), calcolabile facilmente con Macaulay 2 come $tr(h(M_x)M_x^{i+j}) = tr(M_{h(x)x^{i+j}})$, dove M_x è la matrice compagna.*

15.4 Criteri effettivi

Criterio per calcolare il numero di radici reali di un polinomio Sia dato un polinomio $f(x) \in \mathbb{R}[x]$. Dalla traccia della matrice compagna e delle sue potenze si calcola la matrice Bezoutiante. Dalla regola di Cartesio si può calcolare, mediante il polinomio caratteristico della matrice Bezoutiante B, il numero di autovalori positivi e il numero di autovalori negativi di B.

Allora, per il teorema di Sylvester, il numero di radici reali (distinte) è dato dal numero di autovalori positivi di B meno il numero di autovalori negativi di B.

Calcolo effettivo delle molteplicità delle radici. Per conoscere effettivamente le molteplicità di ciascuna radice di $f(x)$, si può calcolare $f_{rid} = \frac{f}{MCD(f,f')}$ e poi continuare induttivamente a valutare le radici di $f_2 = f/f_{rid} = MCD(f, f')$, $f_3 = f_2/f_{2rid}$, e cosìvia. Se chiamiamo d_i il numero di radici distinte di f_i , calcolabile mediante il Teor. 15.9 (ii), allora il numero di radici di molteplicità i è uguale a $d_i - d_{i+1}$. Analogamente, se chiamiamo r_i il numero di radici reali distinte di f_i , calcolabile mediante il Teor. 15.9 (iii), allora il numero di radici reali di molteplicità i è uguale a $r_i - r_{i+1}$. Analogamente si possono trovare le molteplicità delle radici in un qualunque intervallo.

Criterio per stabilire se una matrice reale è diagonalizzabile su $\mathbb{R}$ Sia data A matrice reale. Si calcola il polinomio minimo di A, si veda l'algoritmo prima dell'eserc. 13.14. Si calcola la matrice Bezoutiante B del polinomio , mediante il metodo esposto nell'Osservazione 15.15 con $h = 1$. A è diagonalizzabile se e solo se B è definita positiva. Questo segue da (i) del Teor. 15.9 e dal Teor. 13.11. Ricordiamo che B è definita positiva se e solo se tutti i suoi minori principali sono positivi.

15.5 Esercizi sulle matrici compagne, il bezoutiante e le radici di polinomi reali in una variabile

• Si consideri il polinomio $f(x) = \prod_{i=1}^{10}(x - i)$.

1. Si calcoli la matrice compagna M_x di f. E' diagonalizzabile ? Si confronti la risposta col calcolo numerico degli autovalori di f.

2. Si calcoli la matrice bezoutiante B ed il suo rango.

3. Si calcoli il polinomio caratteristico e la segnatura della matrice bezoutiante B e si provi che è definita positiva. Esiste V reale tale che $VV^t = B$?

4. Si trovino dieci elementi e_i per $i = 1, \ldots, 10$ le cui classi corrispondono agli elementi unità dei dieci addendi della decomposizione

$$\mathbb{R}[x]/(f(x)) = \bigoplus_{i=1}^{10} \mathbb{R}[x]/(x - i).$$

   e_i sono autovettori di M_x ? (Un modo rapido per calcolare $M_x(e_i)$ è $x * e_i \% I$.)

5. Si calcoli la matrice simmetrica $e_i e_j$ per $1 \leq i, j \leq 10$.

6. Si calcoli $\sum_{i=1}^{10} e_i$ e $\sum_{i=1}^{10} e_i^2$.

7. $*$ Si calcoli il numero di radici reali di $f(x) - x^9/10$. Si calcoli il numero di radici reali di $f(x) + x^9 t$ al variare di $t \in \mathbb{R}$.

8. Si calcoli la matrice di valutazione $e_i(j)$ per $1 \leq i, j \leq 10$. Si trovino tutti i polinomi $h(x)$ tale che $h(i) = 1$, per $i = 1, \ldots 10$.

- Si consideri il polinomio $f(x) = \prod_{i=1}^{5}(x-i)^2$.

1. Si calcoli la matrice compagna M_x di f. E' diagonalizzabile ? Si confronti la risposta col calcolo numerico degli autovalori di f. Il quadrato M_x^2 è diagonalizzabile ?

2. Si calcoli la matrice bezoutiante B ed il suo rango.

3. Si calcoli il polinomio caratteristico e la segnatura della matrice bezoutiante B e si provi che è semidefinita positiva. Esiste V reale tale che $VV^t = B$?

4. Si trovino cinque elementi e_i per $i = 1, \ldots, 5$ le cui classi corrispondono agli elementi unità dei cinque addendi della decomposizione f

$$\mathbb{R}[x]/(f(x)) = \bigoplus_{i=1}^{5} \mathbb{R}[x]/(x-i)^2.$$

   e_i sono autovettori di M_x ? (Un modo rapido per calcolare $M_x(e_i)$ è $x * e_i \% I$.) Trovare gli autovettori di M_x.

5. Si calcoli la matrice simmetrica $e_i e_j$ per $1 \leq i, j \leq 5$.

6. Si calcoli $\sum_{i=1}^{5} e_i$ e $\sum_{i=1}^{5} e_i^2$.

7. Si calcoli la matrice di valutazione $e_i(j)$ per $1 \leq i, j \leq 5$. *(La domanda seguente richiede l'interpolazione di Hermite.)* Si trovino tutti i polinomi $h(x)$ tale che $h(i) = 0$, $h'(i) = 1$ per $i = 1, \ldots 5$. Qual è il grado minimo di tali polinomi ?

- Si ripeta i primi 6 punti dell'esercizio precedente con il polinomio

$$f(x) = \prod_{i=1}^{5}(x^2 - i).$$

16 Ideali zero-dimensionali

Il seguente risultato è una conseguenza dell'additività della forma normale rispetto a un ideale (si veda la definizione 2.15).

Teorema 16.1. *Sia $I \subset K[x_1, \ldots, x_n]$ un ideale, fissiamo un ordine monomiale e sia $S =< x^\alpha | x^\alpha \notin LT(I) >$ sottospazio (su K) di $K[x_1, \ldots, x_n]$. Allora $K[x_1, \ldots, x_n]/I \simeq S$ come spazi vettoriali su K.*

Dimostrazione. Sia $G = \{g_1, \ldots, g_s\}$ una base di Gröbner per I. Indichiamo con $[f] \in K[x_1, \ldots, x_n]/I$ gli elementi del quoziente ($f \in K[x_1, \ldots, x_n]$) e definiamo $\phi : K[x_1, \ldots, x_n]/I \to S$ data da $\phi([f]) = \overline{f}^G$ (resto della divisione per G, che abbiamo

indicato anche come $f \,\%\, I$). Se $f = \sum h_i g_i + r$ è la divisione poniamo quindi $\phi([f]) = r$. ϕ è ben definito perché se $f - f' \in I$ abbiamo $f = \sum h_i g_i + r$, $f' = \sum h'_i g_i + r'$ (divisioni per G) e quindi $r - r' \in I$. Pertanto se $r \neq r'$ abbiamo $LT(r - r')$ divisibile per qualche $LT(g_j)$ (per definizione di base di Gröbner) in contraddizione con l'algoritmo di divisione. ϕ è suriettiva perchè se $x^\alpha \in S$ abbiamo $\phi([x^\alpha]) = x^\alpha$. Inoltre se $\overline{f}^G = \overline{f'}^G$ è ovvio che $f - f' \in I$, quindi ϕ è iniettiva. E' facile verificare che $\phi(k[f]) = k\phi([f])$. Rimane da far vedere che ϕ conserva la somma. Infatti se $f = g + r$, $f' = g' + r'$ dove nessun termine di r, r' appartiene a $LT(I)$ segue che $f + f' = (g + g') + (r + r')$ dove nessun termine di $(r + r')$ appartiene a $LT(I)$. Sia r''il resto della divisione di $f + f'$ per G, allora $f + f' = g'' + r''$ dove nessun termine di r'' appartiene a $LT(I)$. Quindi se $(r + r') - r'' \neq 0$ abbiamo $LT[(r + r') - r''] \in LT(I)$ che è una contraddizione. Pertanto $r + r' = r''$, cioè $\phi([f]) + \phi([f']) = \phi([f + f'])$. □

Teorema 16.2. *Sia $K = \mathbb{R}$ oppure $K = \mathbb{C}$ e $I \subset K[x_1, \ldots, x_n]$ un ideale. Sono equivalenti:*

- *i) $V_{\mathbb{C}}(I)$ è finito*
- *ii) $\forall i = 1, \ldots, n \ \exists m_i \geq 0$ tale che $x_i{}^{m_i} \in LT(I)$*
- *iii) lo spazio vettoriale $K[x_1, \ldots, x_n]/I (\simeq S$ vedi teor. 16.1) ha dimensione finita*

Dimostrazione.

- *i)* $\Rightarrow$ *ii)* Se $V_{\mathbb{C}}(I) = \emptyset$ abbiamo $1 \in LT(I)$ dal teorema degli zeri e basta porre $m_i = 0$. Se $V_{\mathbb{C}}(I) \neq \emptyset$ siano $\{a_j\}_{j \in J(i)}$ tutte le i-esime coordinate dei punti di $V_{\mathbb{C}}(I)$. Posto $f_i = \prod_{j \in J(i)} (x_i - a_j)$ abbiamo $f_i \in I(V_{\mathbb{C}}(I)) = \sqrt{I}$ dal teorema degli zeri.

  Quindi $\exists \quad k_i : f_i{}^{k_i} \in I$ e $LT(f_i{}^{k_i}) =: x_i{}^{m_i} \in LT(I)$ come volevamo.

- *ii)* $\Rightarrow$ *iii)* Se $x^\alpha \in S =< x^\alpha | x^\alpha \notin LT(I) >$ (l'uguaglianza per il teor. 16.1) deve essere $\alpha_i \leq m_i - 1$. Quindi $dim S \leq \prod_{i=1}^n m_i$.

- *iii)* $\Rightarrow$ *i)* Per ipotesi gli elementi $[1], [x_1], [x_1{}^2], \ldots, [x_1{}^n], \ldots$sono linearmente dipendenti, quindi $\exists \quad c_j \in J$ tali che $\sum c_j x_1{}^j \in I$. In particolare i punti di $V_{\mathbb{C}}(I)$ possono avere solo un numero finito di coordinate differenti al primo indice (le radici del polinomio precedente). Il ragionamento si ripete anche per gli altri indici.

□

La condizione ii) del teorema precedente dà un algoritmo per stabilire se l'insieme delle soluzioni di un sistema di polinomi è finito, una volta calcolata una base di Gröbner. In caso affermativo la dimostrazione mostra anche che il numero delle soluzioni è limitato da $\prod_{i=1}^n m_i$ (se ci fossero $\prod_{i=1}^n m_i + 1$ soluzioni potrei costruire $\prod_{i=1}^n m_i + 1$ polinomi ciascuno dei quali si annulla in tutti i punti escluso uno, e questi polinomi sarebbero indipendenti in $K[x_1, \ldots, x_n]/I$).

Definizione 16.3. *Un ideale che soddisfa una delle condizioni equivalenti del Teorema 16.2 si dice zero-dimensionale.*

Esercizio 16.4. *Nei casi seguenti determinare quando $V(I)$ è finito e limitare il numero dei punti di $V(I)$.*

1. *$I \subset \mathbb{C}[x,y,z]$ con base di Gröbner data da (x^2y, y^3z, z^5, xz)*
2. *$I \subset \mathbb{C}[x,y,z]$con base di Gröbner data da $(x^2, xy, y^3, yz^{10}, z^7)$*
3. *$I \subset \mathbb{C}[x_1,x_2,x_3,x_4]$ con base di Gröbner data da $(x_1^2, x_2^2, x_3^2, x_1x_4, x_2x_4, x_3x_4)$*

Risposta: $V(I)$ è finito solo nel caso ii) ed il numero di punti è≤ 42.

17 Decomposizione primaria di ideali zero-dimensionali e molteplicità

17.1 Diagonalizzazione simultanea di più matrici

Teorema 17.1. *Diagonalizzazione simultanea*

(i) Siano $A, B\colon V \to V$ due endomorfismi diagonalizzabili. Allora esiste una base di autovettori comuni a A e B se e solo se $AB = BA$.

(ii) Siano $A_1, \dots A_n\colon V \to V$ endomorfismi diagonalizzabili. Allora esiste una base di autovettori comuni a A_i per $i = 1, \dots, n$ se e solo se $A_iA_j = A_jA_i$ per ogni i, j.

Dimostrazione. (i) Sia $\{v_1, \dots v_n\}$ una base comune di autovettori. Allora $Av_i = \lambda_i v_i$ e $Bv_i = \mu_i v_i$. Quindi $ABv_i = \lambda_i\mu_i v_i = BAv_i$. Quindi AB e BA assumono gli stessi valori su una base di V e pertanto sono uguali. Viceversa siano $V_{\lambda_i} = \ker(A - \lambda_i I)$ gli autospazi di A. Se $v \in V_{\lambda_i}$ allora $A(Bv) = B(Av) = B(\lambda_i v) = \lambda_i(Bv)$, da cui $Bv \in V_{\lambda_i}$. Quindi $B(V_{\lambda_i}) \subseteq V_{\lambda_i}$, cioè gli autospazi di A sono B-invarianti. Pertanto il polinomio caratteristico di B si fattorizza come prodotto dei polinomi caratteristici di $B_{|V_{\lambda_i}}$ e quindi gli autovalori di $B_{|V_{\lambda_i}}$ sono tutti in K. Inoltre il polinomio minimo di $B_{|V_{\lambda_i}}$ divide il polinomio minimo di B, siccome quest'ultimo per ipotesi non ha fattori ripetuti, neanche il polinomio minimo di $B_{|V_{\lambda_i}}$ ha fattori ripetuti e pertanto $B_{|V_{\lambda_i}}$ è diagonalizzabile. Mettendo insieme le basi di autovettori per $B_{|V_{\lambda_i}}$, si ottiene una base di autovettori comuni a A e B.

(ii) Se abbiamo una base comune di autovettori l'argomento è lo stesso del punto (i). Viceversa, possiamo ragionare per induzione su n. Se V_i sono gli autospazi di A_n, abbiamo come nel punto (i) che $A_j(V_i) \subseteq V_i$ per ogni i, j. Il ragionamento del punto (i) mostra che, per ogni i, gli $n-1$ endomorfismi $A_{j\,|V_i}$ per $j = 1, \dots, n-1$ commutano a due a due e sono diagonalizzabili e quindi hanno una base di autovettori comuni su V_i. Mettendo insieme queste basi di autovettori comuni, si ottiene una base di autovettori comuni a tutti gli A_i.

□

Proposizione 17.2. *Siano A_i per $i = 1, \dots, n$ endomorfismi tali che $A_iA_j = A_jA_i$ per ogni i, j. Sia A_1 diagonalizzabile con autovalori distinti. Allora ogni A_i è diagonalizzabile e tutti gli A_i hanno una base comune di autovettori.*

Dimostrazione. Per ipotesi, gli autospazi di A_1 $V_{\lambda_i} = \ker(A_1 - \lambda_i I)$ hanno dimensione uno. Come nella dimostrazione precedente abbiamo $A_j(V_{\lambda_i}) \subset V_{\lambda_i}$, e questo vuol dire che i generatori di V_{λ_i} sono autovettori anche per A_j. Pertanto A_j è diagonalizzabile e gli autovettori trovati sono comuni a tutti gli A_j. □

Proposizione 17.3. *Siano A_i per $i = 1, \ldots, n$ endomorfismi tali che $A_iA_j = A_jA_i$ per ogni i, j. Esiste un autovettore (complesso) comune a A_i.*

Dimostrazione. Sia λ_1 un autovalore di A_1. Allora $V_1 = \ker(A_1 - \lambda_1 I)$ è A_2-invariante, pertanto A_2 ha un autovettore su V_1 con autovalore λ_2 e $V_2 = \bigcap_{i=1}^{2} \ker(A_i - \lambda_i I) \neq 0$. Analogamente, V_2 è A_3-invariante, per cui esiste λ_3 tale che $V_3 = \bigcap_{i=1}^{3} \ker(A_i - \lambda_i I) \neq 0$. Proseguendo in questo modo, si trova $V_n = \bigcap_{i=1}^{n} \ker(A_i - \lambda_i I) \neq 0$, e ogni vettore non nullo in V_n è un autovettore comune a A_i. □

Dalla Proposizione precedente segue che un sottospazio di endomorfismi che commutano può essere ridotto (simultaneamente) a forma triangolare.

17.2 Matrici compagne in più variabili

Sia $I = (f_1, \ldots, f_k)$ un ideale zero dimensionale di $K[x_1, \ldots, x_n]$, dove $K = \mathbb{R}$ oppure $\mathbb{C}$. Questo significa che il sistema $f_1 = \ldots = f_k = 0$ ha un numero finito di soluzioni pari a d (contate con la relativa molteplicità), il quoziente $K[x_1, \ldots, x_n]/I$ ha dimensione d ed è generato dalle classi $[x^\alpha]$ di monomi non contenuti in $LT(I)$. Per ogni $f \in K[x_1, \ldots, x_n]/I$, calcolando con Macaulay2 $f\%I$, si ottiene l'espressione di f in questa base.

La moltiplicazione per x_i induce un'applicazione lineare

$$R \xrightarrow{M_{x_i}} R$$

$$[g] \mapsto [gx_i]$$

le matrici di M_{x_i} rispetto alla base $\{[x^\alpha] | x^\alpha \notin LT(I)\}$ si dicono *matrici compagne* di I. Nel caso di una variabile, la diagonalizzazione di M_x giocava un ruolo fondamentale. Nel caso di più variabili, è la diagonalizzazione simultanea delle M_{x_i} che entra in gioco.

Proposizione 17.4. $\forall h(x), k(x) \in K[x_1, \ldots, x_n]$ *vale che*
(i) $M_{h(x)} + M_{k(x)} = M_{h(x)+k(x)}$
(ii) $M_{ah(x)} = aM_{h(x)} \quad \forall a \in K$
(iii) $M_{h(x)} \cdot M_{k(x)} = M_{k(x)} \cdot M_{h(x)} = M_{h(x)k(x)}$
(iv) $M_{h(x_1,\ldots,x_n)} = h(M_{x_1}, \ldots, M_{x_n})$.

Dimostrazione. (i), (ii) e (iii) sono immediate dalla definizione. Per provare (iv) supponiamo dapprima $h = x_j^i$. Allora $M_{x_j^i} = \left(M_{x_j}\right)^i$ come diretta applicazione di (iii).

In generale, se $h(x) = \sum a_\alpha x^\alpha$ allora, utilizzando anche (i)-(iii),

$$M_{h(x_1,\ldots,x_n)} = M_{\sum a_\alpha x^\alpha} = \sum a_\alpha M_{x^\alpha} = \sum a_\alpha \left(M_{x_1}\right)^{\alpha_1} \ldots \left(M_{x_n}\right)^{\alpha_n} = h(M_{x_1}, \ldots, M_{x_n})$$

□

Lemma 17.5. *Se $M_{x_i}v = \lambda_i v$ e $p \in K[x_1, \ldots, x_n]$ allora*

$$p(M_{x_1}, \ldots, M_{x_n})v = p(\lambda_1, \ldots, \lambda_n)v$$

Dimostrazione. È l'analogo multidimensionale del Lemma 13.3.

17.3 Decomposizione primaria e definizione di molteplicità

Sia $I \subset K[x_1, \ldots, x_n]$,dove $K = \mathbb{R}$ oppure $\mathbb{C}$, e sia $V_{\mathbb{C}}(I) = \{p_1, \ldots, p_k\} \subset \mathbb{C}^n$, cioè supponiamo che le soluzioni complesse del sistema definito da I siano un numero finito. Un tale ideale I è detto zero-dimensionale, equivalentemente il polinomio di Hilbert di I è costante, e questo può essere verificato in modo effettivo tramite Macaulay2.

In questa sezione assegneremo una molteplicità a ogni punto $p_i \in V(I)$, analogamente a quanto avviene per le radici di un polinomio in una variabile, in modo che la somma delle molteplicità di tutte le soluzioni sia uguale alla dimensione di $K[x_1, \ldots, x_n]/I$ (si veda il Corollario 17.16).

Sia M_i l'ideale massimale dei polinomi che si annullano in $p_i = ((p_i)_1, \ldots, (p_i)_n)$. L'ideale M_i è generato dai polinomi $x_j - (p_i)_j$. Con un piccolo abuso di notazione, indicheremo con M_i anche la sua immagine nel quoziente $\mathbb{C}[x_1, \ldots, x_n]/I$, che è ancora un ideale massimale.

Lemma 17.6. *(i)* $V(M_1 \cap \ldots \cap M_k) = p_1 \cup \ldots \cup p_k$.

(ii) $\sqrt{I} = M_1 \cap \ldots \cap M_k$.

Questo significa che $g \in \sqrt{I}$ se e solo se $g(p_i) = 0$ per $i = 1, \ldots, k$. In particolare, per ogni elemento del quoziente $g \in \mathbb{C}[x_1, \ldots, x_n]/I$, la valutazione $g(p_i) \in \mathbb{C}$ è ben definita e non dipende dal rappresentante.

Dimostrazione. (i) è elementare e segue dalla Proposizione 6.9. Per provare (ii), se $f \in \sqrt{I}$ allora esiste $m > 0$ tale che $f^m(p_i) = 0$, da cui $f(p_i) = 0$ e quindi $f \in M_1 \cap \ldots \cap M_k$. Viceversa, per il teorema degli zeri, $\sqrt{I} = I(V(I)) = I(p_1 \cup \ldots \cup p_k) = I(V(M_1 \cap \ldots \cap M_k)) = \sqrt{M_1 \cap \ldots \cap M_k} \supset M_1 \cap \ldots \cap M_k$.

□

Lemma 17.7. *Dato $V(I) = \{p_1, \ldots p_k\} \subset \mathbb{C}^n$, esiste un polinomio $h(x)$ tale che $h(p_i)$ siano distinti (si veda la Figura 17.6). Se I è generato da polinomi a coefficienti reali, allora $h(x)$ puo' essere scelto a coefficienti reali. In tale caso, per ogni coppia di punti complessi coniugati $\{p, \overline{p}\}$, abbiamo $h(\overline{p}) = \overline{h(p)}$.*

Dimostrazione. Il prodotto scalare euclideo si può estendere (algebricamente) a $\mathbb{C}^n$ ponendo $(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = \sum_{i=1}^n z_i w_i$, $\forall (z_1, \ldots, z_n), (w_1, \ldots, w_n) \in \mathbb{C}^n$. E' sufficiente scegliere un vettore $H = (h_1, \ldots, h_n)$ tale che il prodotto scalare euclideo $H \cdot (p_i - p_j) \neq 0$ $\forall i \neq j$. Questo è possibile perché $(p_i - p_j)$ sono un numero finito di vettori. Allora $h(x) = \sum_{i=1}^n h_i x_i$ soddisfa la condizione richiesta. □

Lemma 17.8. *Un elemento del quoziente $g \in \mathbb{C}[x_1, \ldots, x_n]/I$ è invertibile se e solo se $g(p_i) \neq 0$ $\forall i$.*

Dimostrazione. Se g è invertibile, segue immediatamente che $g(p_i) \neq 0$. Viceversa, supponiamo $g(p_i) \neq 0$ $\forall i$. Per il lemma 17.7, esiste $h(x)$ tale che $h(p_i)$ sono distinti. Definiamo $g'(x) = \sum_{i=1}^{k} \frac{1}{g(p_i)} \prod_{j \neq i} \frac{h(x)-h(p_j)}{h(p_i)-h(p_j)}$, che soddisfa le uguaglianze $g(p_i)g'(p_i) = 1$ $\forall i$. Per il Lemma 17.6 (ii) abbiamo $1 - gg' \in \sqrt{I}$, da cui esiste $m > 0$ tale che $(1-gg')^m \in I$. Espandendo la potenza m-esima e raccogliendo i termini che contengono g, si trova $\tilde{g}$ tale che $1 - \tilde{g}g \in I$, da cui g è invertibile nel quoziente , come volevamo.□

Lemma 17.9. *Sia $V(I) = \{p_1, \ldots p_k\} \subset \mathbb{C}^n$, e sia $h(x) \in \mathbb{C}[x_1, \ldots, x_n]$. Gli autovalori di $M_{h(x)} \colon K[x_1, \ldots, x_n]/I \to K[x_1, \ldots, x_n]/I$ coincidono con i valori $h(p_i) \in \mathbb{C}$.*

Dimostrazione. Sia λ l' autovalore di $M_{h(x)}$ corrispondente all'autovettore $v(x)$. Allora $(h(x) - \lambda)\, v(x) \in I$. Affermiamo che $h(p_i) = \lambda$ per qualche i. Se per assurdo $h(p_i) - \lambda \neq 0$ $\forall i$, allora $h(x) - \lambda$ è invertibile per il Lemma 17.8. Quindi $v(x) \in I$, che è una contraddizione perché gli autovettori sono non nulli.

Viceversa, proviamo che $h(p_i)$ è un autovalore di $M_{h(x)}$. Sia $q(t)$ il polinomio minimo di $M_{h(x)}$. Allora $0 = q(M_{h(x)}) = M_{q(h(x))}$. Quindi $q(h(x)) \in I$, da cui, valutando in p_i, $q(h(p_i)) = 0$, pertanto $h(p_i)$ è un autovalore per la Prop. 13.4. □

Osservazione 17.10. *Applicando il Lemma 17.9 alle matrici M_{x_i} si ottiene che le coordinate i-esime dei punti di $V(I)$ coincidono con gli autovalori di M_{x_i}. Questa è già un'informazione importante per calcolare i punti di $V(I)$, ma richiede lavoro supplementare per stabilire quali coordinate corrispondono allo stesso punto. Un metodo più efficiente è descritto dalla proposizione 18.4.*

Ricordiamo che un ideale J si dice *primario* se $fg \in J$ implica $f \in J$ oppure $g^m \in J$ per qualche $m > 0$. Per gli ideali valgono le implicazioni

$$\text{massimale} \Longrightarrow \text{primo} \Longrightarrow \text{primario}.$$

e ciascuna delle implicazioni precedenti è stretta. Se J è primario allora $\sqrt{J}$ è primo, ma il viceversa è falso. Però se $\sqrt{J}$ è massimale allora J è primario.

Segue immediatamente dalla definizione che il radicale di un ideale primario è primo.

Teorema 17.11. *[Decomposizione primaria] Sia $V(I) = \{p_1, \ldots, p_k\}$. Sia $h(x) \in \mathbb{C}[x_1, \ldots, x_n]$ tale che $h(p_i)$ siano distinti (si veda il Lemma 17.7).*

Considero per $i = 1, \ldots, k$ le applicazioni lineari

$$M_{h(x)-h(p_i)} \colon \mathbb{C}[x_1, \ldots, x_n]/I \to \mathbb{C}[x_1, \ldots, x_n]/I$$

(i) Posto $A_i := \ker \left[M_{h(x)-h(p_i)}\right]^{\infty}$ (sono gli autospazi generalizzati di $M_{h(x)}$ per il Lemma 17.9, sono sottoalgebre e anche ideali di $\mathbb{C}[x_1, \ldots, x_n]$), abbiamo la decomposizione diretta di sottoalgebre

$$\mathbb{C}[x_1, \ldots, x_n]/I = \oplus_{i=1}^{k} A_i. \tag{17.1}$$

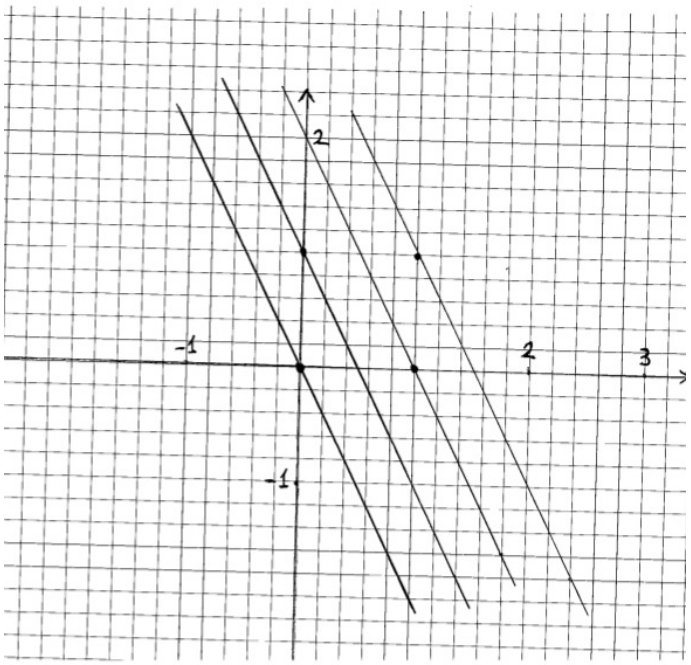


Figura 17.6: In figura i quattro punti $p_1 = (0,0)$, $p_2 = (1,0)$, $p_3 = (0,1)$, $p_4 = (1,1)$ corrispondono a $V(I)$ dove $I = (x(x-1), y(y-1))$. Posto $h(x,y) = 2x+y$, il fascio di rette parallele $h(x,y) = \lambda$ incontra $V(I)$ per i 4 valori $\lambda = h(p_i)$. I quattro autovalori di $M_{h(x)}$ sono $h(p_i)$. Ciascun punto ha molteplicità 1. In questo caso, M_x e M_y non hanno autovalori distinti (quali sono?).

Se $v(x) \in A_i$ allora $v(p_j) = 0$ $\forall j \neq i$. Ogni sottoalgebra A_i ha un elemento unità e_i, che soddisfa le proprietà $e_i^2 = e_i$, $e_i e_j = 0$ per $i \neq j$. Inoltre, valutando in p_j, $e_i(p_j) = \delta_{ij}$. Gli elementi $g \in A_i$ sono invertibili in A_i se e solo se $g(p_i) \neq 0$.

(ii)Posto $J_i = \oplus_{j \neq i} A_j$, ideale di $\mathbb{C}[x_1, \ldots, x_n]/I$, la sua retroimmagine $\tilde{J}_i \subset \mathbb{C}[x_1, \ldots, x_n]$ è un ideale primario, tale che $\sqrt{\tilde{J}_i} = M_i$, $A_i = \mathbb{C}[x_1, \ldots, x_n]/(\tilde{J}_i)$,

$$\cap_{i=1}^{n} \tilde{J}_i = I. \tag{17.2}$$

L' intersezione (17.2) si dice la decomposizione primaria di I. Notiamo che e_i corrisponde alla classe di 1 modulo $\tilde{J}_i$.

Dimostrazione. (i) La somma diretta segue dalla Prop. 13.9 e dal Lemma 17.9, come somma di spazi vettoriali. E' facile verificare dalla definizione che A_i è un ideale. Se $v(x) \in A_i$ allora esiste n_i tale che $(h(x) - h(p_i))^{n_i} v(x) \in I$, da cui valutando per $x = p_j$ segue $v(p_j) = 0$. L'unità e_i di ogni sottoalgebra A_i proviene dalla decomposizione in somma diretta $1 = \sum_{i=1}^{k} e_i$, risolubile dividendo 1 per i generatori di ciascuna A_i (aggiungendo eventualmente i generatori di I), si può applicare il comando *"quotientRemainder"* di M2. L'elemento e_i funge da unità in A_i perché, preso

$a_i \in A_i \subset \mathbb{C}[x_1, \ldots, x_n]/I$, moltiplicando per 1 abbiamo

$$a_i = 1 \cdot a_i = \sum_{j=1}^{k} e_j a_i = e_i a_i.$$

Se $i \neq j$, abbiamo $e_i e_j \in A_i \cap A_j = 0$, da cui $1 = 1^2 = \sum_{i=1}^{k} e_i^2$ e per l'unicità della decomposizione $e_i^2 = e_i$. L'affermazione sull'invertibilità segue applicando il Lemma 17.8 al caso in cui $V(I)$ contiene un solo punto ($k = 1$).

(ii) Per come è definito l'ideale $\tilde{J}_i$, abbiamo $\mathbb{C}[x_1, \ldots, x_n]/\tilde{J}_i \simeq A_i$. Notiamo che $\cap_{i=1}^{n} J_i = 0$, da cui prendendo le retroimmagini $\cap_{i=1}^{n} \tilde{J}_i = I$. Valutando gli elementi unità e_i abbiamo $p_i = V(\tilde{J}_i)$, dal Nullstellensatz segue che $\sqrt{\tilde{J}_i} = M_i$, ideale massimale, segue che $\tilde{J}_i$ è primario. □

Osservazione 17.12. *Gli anelli A_i hanno come unico ideale massimale M_i e sono quindi anelli locali. La decomposizione (17.1) spiega l'origine del termine locale. Ogni A_i corrisponde a localizzare in un punto p_i , cioè la classe di un polinomio in A_i è influenzata soltanto dal comportamento vicino a p_i e può essere ricostruita da un opportuno sviluppo di Taylor nel punto p_i (rispetto ai monomi $\notin \tilde{J}_i$). Per approfondimenti sugli anelli locali si può vedere il cap. 4 di [CLO2]. In particolare la localizzazione di A in M_i coincide con A_i. Infatti l'elemento $e_i \notin M_i$ soddisfa $e_i A_j = 0$ per $j \neq i$ e quindi localizzando rispetto a M_i gli addendi A_j per $j \neq i$ sono identificati a zero.*

Definizione 17.13. $\dim A_i$ *si dice <u>molteplicità</u> di p_i in I, la indicheremo con m_{p_i}, non dipende dal polinomio $h(x)$ scelto nel Teorema 17.11.*

Il fatto che la molteplicità sia ben definita e non dipenda da $h(x)$ segue dal fatto che A_i ha una definizione intrinseca come localizzato di A rispetto a M_i. Per provare che la molteplicità non dipende da $h(x)$, con un ragionamento diretto ed elementare, prendiamo un altro polinomio $h'(x)$ che assume valori distinti sui punti p_i. Si osserva che A_i è $h'(x)$-invariante. Siccome $A_i = \mathbb{C}[x_1, \ldots, x_n]/(\tilde{J}_i)$ e $V(\tilde{J}_i) = V(\sqrt{\tilde{J}_i}) = \{p_i\}$ dal Teor. 17.11 (ii), segue che l'unico autovalore di $M_{h'(x)}$ su A_i è $h'(p_i)$ per il Lemma 17.9, applicato al caso $I = \tilde{J}_i$. Pertanto l' autospazio generalizzato A_i di $M_{h(x)}$ relativo all'autovalore $h(p_i)$ è contenuto nell' autospazio generalizzato A'_i di $M_{h'(x)}$ relativo all'autovalore $h'(p_i)$. Sia la somma dei A_i che quella dei A'_i sono entrambe dirette, quindi vale l'uguaglianza $A_i = A'_i$.

Come conseguenza di questo ragionamento enunciamo esplicitamente la

Proposizione 17.14. *(i) Sia $h(x)$ un polinomio che assume valori distinti sui punti p_i. Gli ideali A_i sono gli autospazi generalizzati di $M_{h(x)}$ e l'unico autovalore di $M_{h(x)}$ su A_i è $h(p_i)$.*

(ii) Per ogni polinomio $k(x)$, l'unico autovalore di $M_{k(x)}$ su A_i è $k(p_i)$ (segue dal Lemma 17.9 applicato al caso $I = \tilde{J}_i$).

(iii) La sottoalgebra A_i del teorema 17.11 (i) dipende solo da I.

(iv) La decomposizione primaria $I = \cap_{i=1}^{n} \tilde{J}_i$ del Teorema 17.11 (ii) è unica.

Esercizio 17.15. *Modificando la figura 17.6, consideriamo* $I = \big(x^2(x-1), y(y-1)\big)$. *Provare che, con le notazioni della figura 17.6,* $V(I) = \{p_1, p_2, p_3, p_4\}$, *la molteplicità di* p_1, p_3 *è* 2, *la molteplicità di* p_2, p_4 *è* 1.

Corollario 17.16. *La somma delle molteplicità di ciascun* p_i *in* I *è uguale alla dimensione di* $\mathbb{C}[x_1, \ldots, x_n]/I$. *Questo valore è uguale al polinomio di Hilbert di* I, *che ha grado zero ed è costante.*

Corollario 17.17. *Vale* $I = \sqrt{I}$ *se e solo se tutti i punti hanno molteplicità* 1 *in* I.

Dimostrazione. Basta confrontare

$$\mathbb{C}[x_1, \ldots, x_n]/I = \oplus_{i=1}^{k} \mathbb{C}[x_1, \ldots, x_n]/\tilde{J}_i$$

con

$$\mathbb{C}[x_1, \ldots, x_n]/\sqrt{I} = \oplus_{i=1}^{k} \mathbb{C}[x_1, \ldots, x_n]/M_i,$$

dove nella seconda somma gli addendi hanno dimensione 1, si veda il Lemma 17.6 (ii). □

La decomposizione primaria del Teorema 17.11 si generalizza a ideali qualunque di $K[x_1, \ldots, x_n]$, dove $V(I)$ può avere dimensione positiva, secondo un risultato classico di Lasker-Noether. In questi casi, trovare una decomposizione esplicita è più difficile, e non è necessariamente unica, come nel caso zero-dimensionale. Per dettagli si può consultare il cap. 4 §7 di [CLO1].

17.4 Calcolo effettivo della molteplicità di ogni soluzione

Il numero di radici con una data molteplicità può essere calcolato con un metodo analogo a quello descritto in 15.4, calcolando il radicale di un ideale (algoritmo di Krick-Logar) ed iterando il comando quotient(I, radical I). Questo metodo è complesso, sia in teoria che in pratica. In questa sezione accenniamo a una tecnica alternativa che ha successo con probabilità uno ed è più semplice.

Per ogni polinomio h, la molteplicità algebrica dell'autovalore λ per

$$M_h \colon K[x_1, \ldots x_n]/I \to K[x_1, \ldots x_n]/I$$

è uguale a $\sum_{\{p|h(p)=\lambda\}} m_p$, infatti $K[x_1, \ldots x_n]/I = \oplus A_i$, ogni A_i è M_h- invariante e l'unico autovalore di M_h su A_i è proprio $h(p_i)$ (si veda il Lemma 17.9).

Pertanto, scegliendo una forma lineare h generale che prende valori distinti su ogni punto di $V(I)$, le molteplicità m_{p_i} possono essere calcolate come le molteplicità algebriche degli autovalori di M_h. A_i sono gli autospazi generalizzati per M_h, come definiti nella Prop. 13.9. Con probabilità uno, una forma lineare h scelta casualmente (random) soddisfa questa condizione e permette di calcolare effettivamente le molteplicità. Purtroppo, la scelta random non garantisce a priori di trovare il polinomio h richiesto.

Un criterio sufficiente per verificare se h prende valori distinti su $V(I)$, senza ancora conoscere $V(I)$, è verificare se M_h è regolare, cioè se il suo polinomio minimo e caratteristico coincidono (a meno del segno). Purtroppo, il criterio è solo sufficiente. Ad esempio consideriamo $K[x, y]/(x^2, y^2)$ che corrisponde al punto $(0, 0)$ con molteplicità

4. La matrice M_x ha un autospazio di dimensione 2 generato da x, xy, e ci sono due blocchi di Jordan di ordine 2. La matrice M_y ha un autospazio di dimensione 2 generato da y, xy, e ci sono due blocchi di Jordan di ordine 2. La matrice $M_{x+y} = M_x + M_y$ ha un autospazio di dimensione 2 generato da $xy, x - y$, e ci sono un blocco di Jordan di ordine 1 e un blocco di Jordan di ordine 3, e questo è il comportamento per $M_{\alpha x+\beta y}$ generale.

Una tecnica che garantisce il calcolo effettivo delle molteplicità m_{p_i} è di calcolare le molteplicità degli autovalori di M_{x_1}, isolando ciascun autovalore in un intervallo, seguendo l'Osservazione 15.13 (si veda il Teorema 19.1). Per ciascuno di questi intervalli, si calcolano le molteplicità degli autovalori di M_{x_2}, isolandoli in intervalli rispetto a x_2, e cosìvia. Questa procedura è laboriosa ma ha sempre successo.

18 Sistemi zero dimensionali in più variabili

Sia $K = \mathbb{R}$ oppure $K = \mathbb{C}$. Consideriamo $f_i \in K[x_1, \ldots, x_n]$ per $i = 1, \ldots, k$. Sia $I = (f_1, \ldots, f_k)$ l'ideale generato da questi polinomi.

Teorema 18.1 (Stickelberger)**.** *Sia I un ideale zero dimensionale e siano $M_{x_i} \colon \mathbb{C}[x_1, \ldots, x_n]/I \to \mathbb{C}[x_1, \ldots, x_n]/I$ le applicazioni lineari (compagne) indotte dalla moltiplicazione per x_i. Esiste un autovettore v comune a M_{x_i} con autovalori λ_i, cioè $M_{x_i} v = \lambda_i v$, se e solo se $(\lambda_1, \ldots, \lambda_n) \in V(I)$.*

Dimostrazione. Sia v un autovettore tale che $M_{x_i} v = \lambda_i v \ \forall i$. Se $f \in I$, ricordiamo che $M_{f(x_1,\ldots,x_n)} = 0$, quindi $0 = M_{f(x_1,\ldots,x_n)} v = f(M_{x_1}, \ldots, M_{x_n}) v = f(\lambda_1, \ldots, \lambda_n) v$, l'ultima uguaglianza per il Lemma 17.5, da cui $f(\lambda_1, \ldots, \lambda_n) = 0$.

Viceversa, dobbiamo provare che le coordinate di ogni $p_i \in V(I)$ sono autovalori di un autovettore comune delle matrici M_{x_j}. Decomponiamo $\mathbb{C}[x_1, \ldots, x_n]/I = \oplus_{i=1}^{k} A_i$ secondo il Teorema 17.11. A_i è M_{x_j}-invariante per $j = 1, \ldots, n$ e $M_{x_1}, \ldots M_{x_n}$ commutano. Pertanto per la Prop. 17.3 (triangolarizzazione simultanea) esiste un autovettore comune per gli endomorfismi $(M_{x_j})_{|A_i}$, per $j = 1, \ldots, n$, il cui autovalore relativo a $(M_{x_j})_{|A_i}$ è la coordinata j-esima di p_i, per il Lemma 17.9 e la Proposizione 17.14 (i), come volevamo. □ □

Corollario 18.2. *Nelle ipotesi del teorema di Stickelberger, il polinomio monico generatore dell'ideale di eliminazione $I \cap \mathbb{C}[x_i]$ coincide con il polinomio minimo di M_{x_i} (sostituendo $x = x_i$).*

Dimostrazione. Sia $p(x)$ il polinomio generatore di $I \cap \mathbb{C}[x_i]$ e sia $h(x)$ il polinomio minimo di M_{x_i}. Siccome $p(M_{x_i}) = M_{p(x_i)}$ è l'applicazione nulla da $K[x_1, \ldots, x_n]/I$ in sè stesso, perché $p \in I$, segue che h divide p. Viceversa $h(M_{x_i}) = M_{h(x_i)}$ è l'applicazione nulla, quindi applicata ad 1 mostra che $h(x_i) \in I$, da cui p divide h. □

Teorema 18.3. *Sia $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ un ideale zero-dimensionale. Le seguenti condizioni sono equivalenti*

(i) M_{x_i} sono diagonalizzabili

(ii) M_{x_i} sono diagonalizzabili simultaneamente (cioè con una base comune di autovettori).

(iii) $V(I)$ ha punti distinti, cioè I è radicale. (eserc. 12 pag. 61 di [CLO2])

Dimostrazione. (i) e (ii) sono equivalenti per il Teor. 17.1.

(iii)$\Longrightarrow$(i) Se $V(I)$ ha punti distinti, scegliamo per il Lemma 17.7 una combinazione lineare $h = \sum_i a_i x_i$ che assume valori distinti su $V(I)$. Allora dal Lemma 17.9 M_h ha d autovalori distinti e quindi è diagonalizzabile. Per la Prop. 17.2 otteniamo che M_{x_i} (che commutano con M_h) sono tutte diagonalizzabili.

(ii)$\Longrightarrow$(iii) Se M_{x_i} sono diagonalizzabili, allora ogni elemento di A_j è un autovettore per M_{x_i} con autovalore $(p_j)_i$. In particolare $e_j(x_i - (p_j)_i) \in \tilde{J}_j$ per ogni i, e dall'invertibilità di e_j (modulo $\tilde{J}_j$) segue $(x_i - (p_j)_i) \in \tilde{J}_j$ per ogni i. Abbiamo che i generatori di M_j appartengono a $\tilde{J}_j$ e quindi $M_j = \tilde{J}_j$, da cui la molteplicità di p_j è 1.

□

Proposizione 18.4. *Siano $x^{\alpha(1)}, \dots x^{\alpha(m)}$ i monomi non in $LT(I)$ che generano $K[x_1, \dots, x_n]/I$. Per ogni punto p di $V(I)$ e ogni polinomio h, il vettore $p^{\alpha(1)}, \dots p^{\alpha(m)}$ (ottenuto calcolando i monomi in p) è un autovettore di M_h^t con autovalore $h(p)$.*

Dimostrazione. Siano m_{ij} i coefficienti di M_h. Abbiamo $[x^{\alpha(j)}h] = M_h([x^{\alpha(j)}]) = \sum_{i=1}^m m_{ij}[x^{\alpha(i)}]$ (nell'ultima uguaglianza si pensa $[x^{\alpha(j)}]$ come vettore colonna con coefficienti zero tranne $x^{\alpha(j)}$ al j-esimo posto.

Valutando in p otteniamo $p^{\alpha(j)}h(p) = \sum_{i=1}^m m_{ij}p^{\alpha(i)}$, che equivale alla tesi. □

Seconda dimostrazione. Diamo una dimostrazione più elegante, senza usare le coordinate. La valutazione in p, $ev(p) \in (K[x_1, \dots, x_n]/I)^\vee$ è definita da $ev(p)(h) = h(p)$. Ricordiamo (vedi [Aba, 8C.2]) che se $A\colon V \to W$ è una applicazione lineare, la sua trasposta $A^t\colon W^\vee \to V^\vee$ è definita da $A^t(w^*)(v) = w^*(A(v))\ \forall w^* \in W^\vee, v \in V$. La notazione è giustificata dal fatto che la matrice di A^t, calcolata nelle basi duali, coincide con la trasposta della matrice di A. Facciamo vedere che $ev(p)$ è autovettore della trasposta M_h^t, con autovalore $h(p)$. Infatti, $\forall b \in K[x_1, \dots, x_n]/I$

$$M_h^t(ev(p))(b) = ev(p)(M_h(b)) = ev(p)(hb) = h(p)b(p) = h(p)ev(p)(b)$$

da cui

$$M_h^t(ev(p)) = h(p)ev(p)$$

che equivale alla tesi. □

La Proposizione 18.4 è utile per il calcolo delle soluzioni di un sistema polinomiale, soprattutto quando tra i monomi non in $LT(I)$ appaiono i generatori x_i. Ad esempio l'ideale $((x+y)^4 + 2 * x * y^2, x^2 + y^2 - x) \subset \mathbb{Q}[x, y]$ ha una base di $\mathbb{Q}[x, y]/LT(I)$ data dagli otto monomi $\{1, x, xy, xy^2, xy^3, y, y^2, y^3\}$. Notiamo che nelle posizioni $0, 1, 5$ appaiono rispettivamente $\{1, x, y\}$. Allora per ogni autovettore $v = \{v_0, \dots, v_7\}$ di M_h^t, le espressioni $x = v_1/v_0$, $y = v_5/v_0$ forniscono le coordinate di un punto $p \in V(I)$, tale che l'autovalore corrispondente a v è proprio $h(p)$.

19 La forma traccia in più variabili e il numero di soluzioni reali

In questa sezione, I è un ideale zero-dimensionale di $\mathbb{R}[x_1, \ldots, x_n]$. Le sue soluzioni in $\mathbb{C}^n$ si dividono in punti reali e in coppie di punti complessi coniugati. I può essere visto come ideale in $\mathbb{C}[x_1, \ldots, x_n]$ (generato da polinomi reali), e il suo quoziente R si spezza, come nel Teorema 17.11, nella somma diretta di A_i, dove alcuni A_i corrispondono ai punti reali, mentre altre coppie A_j, $\overline{A_j}$ corrispondono a coppie di punti complessi coniugati. Il coniugio agisce su $\mathbb{C}[x_1, \ldots, x_n]$, coniugando i coefficienti di ogni polinomio, ed è un morfismo di anelli, che lascia invarianti le sottoalgebre A_i corrispondenti ai punti reali e scambia tra loro le sottoalgebre coniugate A_j e $\overline{A_j}$. Ne segue che in corrispondenza dei punti reali le unità $e_i \in A_i$ sono reali, mentre coniugando l'unità $e_j \in A_j$ relativa a una coppia coniugata si trova l'unità $\overline{e_j} \in \overline{A_j}$. Notiamo che la somma $A_j \oplus \overline{A_j}$ è una sottoalgebra con unità $e_j + \overline{e_j}$, che è sempre un anello locale. Questo permette di decomporre sui reali $\dim \mathbb{R}[x_1, \ldots, x_n]/I$, che diventa somma delle sottoalgebre A_i generate dalle unità reali e_i corrispondenti ai punti reali e dalle sottoalgebre generate da $e_j + \overline{e_j}$ nel caso di coppie di punti coniugati. Il campo residuo di ciascuna sottoalgebra (vista come anello locale) è $\mathbb{R}$ nel primo caso e $\mathbb{C}$ nel secondo caso. In alternativa, scelto $h \in \mathbb{R}[x_1, \ldots, x_n]$ che assume valori distinti su $V_{\mathbb{C}}(I)$, la sottoalgebra relativa a p nel primo caso è $\ker M^{\infty}_{h(x)-h(p)}$, mentre la sottoalgebra relativa alla coppia $\{p, \overline{p}\}$ nel secondo caso è $\ker M^{\infty}_{(h(x)-h(p))(h(x)-h(\overline{p}))}$.

La forma traccia è definita analogamente al caso unidimensionale, cioè

$$B_h(a, b) = Tr(M_{hab})$$

per ogni $a, b \in R$. La decomposizione $\oplus_i A_i$ è ortogonale rispetto a B_h (basta calcolarla sulle unità di ogni sottoalgebra).

Il seguente teorema generalizza il criterio di Sylvester al caso multidimensionale e permette di calcolare, in aritmetica esatta, il numero di punti reali e di coppie di punti complessi coniugati in $V(I)$, oltre ad altre informazioni che, per particolari h, permettono di localizzare le radici (ad esempio studiando a quale ottante appartengono).

Teorema 19.1 (Hermite)**.** *(i) Sia* $\dim \mathbb{R}[x_1, \ldots, x_n]/I = m$*, sia* $h \in \mathbb{R}[x_1, \ldots, x_n]$*. La varietà* $V_{\mathbb{R}}(I)$ *consiste di* m *punti distinti* p *tali che* $h(p) > 0$ *se e solo se* B_h *è definita positiva. In particolare* $V_{\mathbb{R}}(I)$ *consiste di* m *punti distinti se e solo se* B_1 *è definita positiva.*

(ii) Il rango di B_h *è il numero di punti distinti* $p \in V_{\mathbb{C}}(I)$ *tali che* $h(p) \neq 0$*. In particolare il rango di* B_1 *è il numero di punti distinti in* $V(I)$*.*

(iii) il numero di punti reali (distinti) $p \in V_{\mathbb{R}}(I)$ *tali che* $h(p) > 0$ *meno il numero di punti reali (distinti)* $p \in V_{\mathbb{R}}(I)$ *tali che* $h(p) < 0$ *è uguale alla segnatura di* B_h*. In particolare il numero dei punti reali (distinti)* $p \in V_{\mathbb{R}}(I)$ *è uguale alla segnatura di* B_1*.*

Inoltre supponiamo che $h(p) \neq 0 \ \forall p \in V(I)$ *non reale, ipotesi soddisfatta se* h *assume valori distinti su* $V_{\mathbb{C}}(I)$*.*

(iv) il numero di punti reali (distinti) $p \in V_{\mathbb{R}}(I)$ tali che $h(p) > 0$ è uguale al numero di autovalori positivi di B_h meno il numero di autovalori negativi di B_1.

(v) il numero di punti reali (distinti) $p \in V_{\mathbb{R}}(I)$ tali che $h(p) < 0$ è uguale al numero di autovalori negativi di B_h meno il numero di autovalori negativi di B_1.

Dimostrazione. Sviluppiamo i punti salienti della dimostrazione, che è analoga a quella del Teorema 15.12. Per ogni polinomio h, la matrice M_h si decompone su ciascuna sottoalgebra A_i, inoltre la traccia di M_h su A_i è uguale a $m_{p_i}h(p_i)$. Infatti l'unico autovalore di M_h su A_i è dato da $h(p_i)$ per il Lemma 17.9.

Per calcolare la forma traccia, osserviamo che A_i ha come ideale massimale l'ideale M_i dei polinomi che si annullano in p_i. Pertanto si può scegliere come base di A_i i polinomi e_if_j (vedi il Teor. 17.11 (i)) per $j = 0, \dots m_{p_i} - 1$ dove $f_0 = 1$ e $f_j(p_i) = 0$ per $j \geq 1$. Siccome $(x - p_i)^{\alpha_k}$ formano una base dell'anello dei polinomi, al variare di $\alpha_k \in \mathbb{Z}^n_{\geq 0}$, (e questo per ogni p_i), si può anche scegliere $f_j = (x - p_i)^{\alpha_j}$ per convenienti α_j.

Pertanto calcolando la matrice di B_h rispetto a questa base, per una radice reale rimane solo il contributo di $tr(M_{he_ie_i})$ che vale $h(p_i)$ e tutti gli altri elementi hanno la forma $tr(M_{he_if_{j_1}f_{j_2}})$ dove uno tra j_1 e j_2 è positivo, e quindi $he_if_{j_1}f_{j_2}$ vale zero in p_i a per quanto visto $tr(M_{he_if_{j_1}f_{j_2}}) = 0$. Il rango di B_h ristretta a A_i vale 1.

Nel caso di una coppia di radici complesse coniugate $\{p_i, \overline{p_i}\}$ allora possiamo considerare la base $e_if_j + \overline{e_if_j}$, $\frac{1}{\sqrt{-1}}\left(e_if_j - \overline{e_if_j}\right)$, dove e_i e f_j sono gli stessi polinomi visti nel caso complesso relativi a p_i ed i loro coniugati $\overline{e_i}$, $\overline{f_j}$ vanno intesi come i polinomi con le stessi monomi ed i coefficienti coniugati. In particolare $\overline{e_i}$ è l'unità della sottoalgebra relativa a $\overline{p_i}$, (si veda l'esempio 19.2).

La matrice di B_h rispetto a questa base è nulla tranne il blocco 2×2 in alto a sinistra, corrispondente agli elementi della base $\{e_i + \overline{e_i}, \frac{1}{\sqrt{-1}}(e_i - \overline{e_i})\}$, che è

$$m_{p_i}\begin{bmatrix} h(p_i) + h(\overline{p_i}) & \frac{1}{\sqrt{-1}}(h(p_i) - h(\overline{p_i})) \\ \frac{1}{\sqrt{-1}}(h(p_i) - h(\overline{p_i})) & -(h(p_i) + h(\overline{p_i})) \end{bmatrix} = m_{p_i}U^tDU$$

dove $U = \begin{bmatrix} 1 & -\sqrt{-1} \\ 1 & \sqrt{-1} \end{bmatrix}$, $D = \begin{bmatrix} h(p_i) & 0 \\ 0 & h(\overline{p_i}) \end{bmatrix}$

Siccome $\det U = 2\sqrt{-1}$, $detD = |h(p_i)|^2$ segue $\det(U^tDU) = (\det U)^2 \det D = -4|h(p_i)|^2$, quindi B_h ha rango 2 se $h(p_i) \neq 0$ e rango zero altrimenti, mentre se il rango è 2 allora $\det U^tDU < 0$ e la segnatura di B_h vale zero. Notiamo che la segnatura è zero in ogni caso. □

Esempio 19.2. *Nel caso $\mathbb{C}[x]/((x-a)(x-\overline{a})) = \mathbb{C}[x]/(x-a) \oplus \mathbb{C}[x]/(x-\overline{a})$ l'unità della sottoalgebra $\mathbb{C}[x]/(x-a)$ è $e = \frac{x-\overline{a}}{a-\overline{a}}$ mentre l'unità dell'altra sottoalgebra $\mathbb{C}[x]/(x-\overline{a})$ è $\overline{e} = \frac{x-a}{\overline{a}-a} = 1 - e$. In questo caso $\frac{1}{\sqrt{-1}}(e - \overline{e}) = -\frac{x-Re(a)}{Im(a)}$ ed abbiamo $\frac{1}{\sqrt{-1}}(e-\overline{e})(a) + \frac{1}{\sqrt{-1}}(e-\overline{e})(\overline{a}) = 0$ e più in generale $\frac{1}{\sqrt{-1}}(e-\overline{e})(a)h(a) + \frac{1}{\sqrt{-1}}(e-\overline{e})(\overline{a})h(\overline{a}) = \frac{1}{\sqrt{-1}}(h(a) - h(\overline{a}))$. Inoltre $\left(\frac{1}{\sqrt{-1}}(e-\overline{e})\right)^2 = -1$.*

Riferimenti bibliografici

[Aba] M. Abate, *Geometria*, McGraw-Hill, 1996

[CLO1] D. Cox, J. Little, D. O'Shea, *Ideal, Varieties and Algorithms*, Springer, 1992

[CLO2] D. Cox, J. Little, D. O'Shea, *Using Algebraic Geometry*, Springer, 1998

[EM] M. Elkadi, B. Mourrain, *Introduction à la rèsolution de systèmes polynomiaux*, Mathèmatiques et Applications 59, Springer, Berlin, 2007

[Stu] B. Sturmfels, *Solving systems of polynomial equations*, CBMS Regional Conference Series in Mathematics, 97, AMS, 2002