LA FORMA CANONICA DI JORDAN

R. BENEDETTI

Inquadriamo la questione e richiamiamo alcuni fatti che saranno considerati noti al lettore.

Sia V uno spazio vettoriale su $\mathbb{K}$, $\dim V = n$. Sullo spazio End (V) degli endomorfismi di V, consideriamo la relazione di equivalenza definita da:

$f, g \in$ End (V) *si dicono* coniugati *e scriveremo* $f \sim g$, *se è solo se esiste* $h \in GL(V)$ *tale che* $g = h^{-1} \circ f \circ h$.

Nel caso in cui $V = \mathbb{K}^n$, End $(\mathbb{K}^n)$ è canonicamente identificato con lo spazio delle matrici $n \times n$ $M(n, \mathbb{K})$, $GL(V)$ con $GL(n, \mathbb{K})$ e si usa dire che due matrici coniugate sono *simili*. Vogliamo studiare il quoziente End $(V)/\sim$ ed in particolare $M(n, \mathbb{K})/\sim$. Per ogni endomorfismo f di V ed ogni base $\mathcal{B}$ di V, indicheremo con $M_{\mathcal{B}}(f)$ la matrice che esprime f in termini del sistema di coordinate associato a $\mathcal{B}$.

- Sono fatti tra loro equivalenti:
  (1) $f \sim g$
  (2) Per qualsiasi base $\mathcal{B}$, le matrici $M_{\mathcal{B}}(f)$ e $M_{\mathcal{B}}(g)$ sono simili.
  (3) Esistono basi $\mathcal{B}$ e $\mathcal{B}'$ di V tali che $M_{\mathcal{B}}(f) = M_{\mathcal{B}'}(g)$.
- **Alcuni invarianti.** Se $g \sim f$ allora:
  (1) f e g hanno lo stesso polinomio caratteristico, $p_f(t) = p_g(t) \in \mathbb{K}[t]$.
  (2) f e g hanno lo stesso spettro di autovalori, Spettro(f) = Spettro(g)
  (3) Per ogni autovalore λ che hanno necessariamente in comune, le molteplicità algebriche in quanto radici del polinomio caratteristico coincidono, $m_f(\lambda) = m_g(\lambda)$, e i rispettivi autospazi hanno la stessa dimensione, $\dim V_\lambda(f) = \dim V_\lambda(g)$.
- **Endomorfismi triangolabili.** Un endomorfismo f è triangolabile se esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}(f)$ è triangolare superiore. Sono fatti equivalenti
  (1) f è triangolabile.
  (2) Il polinomio caratteristico di f è completamente fattorizzabile in $\mathbb{K}[t]$:

$$p_f(t) = \pm \prod_{j=1}^{k} (t - \lambda_j)^{m_j}$$

$$\lambda_j \neq \lambda_i, \text{ se } i \neq j; \ \sum_j m_j = n \ .$$

  (3) Esiste una base $\mathcal{B} = \{v_1, \dots, v_n\}$, con bandiera Span $\{v_1\} \subset$ Span $\{v_1, v_2\} \subset$ Span $\{v_1, v_2, v_3\} \subset \dots \subset V$ f–invariante.

  Essere o no triangolabile è una proprietà invariante rispetto alla coniugazione, per cui possiamo restringere la relazione di equivalenza all'insieme $\mathcal{T}(V)$ degli endomorfismi triangolabili. Poiché $\mathbb{C}$ è un campo algebricamente chiuso, allora $\mathcal{T}(V) =$ End (V) per ogni $\mathbb{C}$-spazio vettoriale V.

La teoria della "forma canonica di Jordan" fornirà un insieme completo di invarianti ed una descrizione esaustiva del quoziente $\mathcal{T}(V)/\sim$.

1. Ideale e polinomio caratteristico di un endomorfismo

Dati $f \in \text{End }(V)$ e un polinomio $p(t) = a_0 + a_1 t + \cdots + a_k t^k \in \mathbb{K}[t]$, poniamo

$$p(f) := a_0 f^0 + a_1 f^1 + \ldots a_k f^k \in \text{End }(V)$$

dove $f^0 = Id$, $f^1 = f$, $f^n = f^{n-1} \circ f$.

Il seguente semplice Lemma è cruciale in molti degli argomenti che seguono.

Lemma 1.1. *Se $p(t) = a(t)b(t)$ allora $p(f) = a(f) \circ b(f) = b(f) \circ a(f)$.*

Dimostrazione. Ci si riduce al caso evidente di un monomio per cui cui $t^m = t^r t^s$, $f^m = f^r \circ f^s = f^s \circ f^r$.

□

Definiamo

$$I(f) = \{p(t) \in \mathbb{K}[t];\ p(f) = 0\}$$

Lemma 1.2.
(1) *$I(f)$ contiene polinomi di grado ≥ 1.*
(2) *Per ogni $p(t), q(t) \in I(f)$, allora $p(t) + q(t) \in I(f)$; per ogni $p(t) \in I(f)$ e ogni $q(t) \in \mathbb{K}[t]$, allora $q(t)p(t) \in I(f)$. Un sottoinsieme non vuoto di $\mathbb{K}[t]$ che verifichi queste proprietà è detto un ideale, pertanto $I(f)$ è detto* l'ideale dell'endomorfismo.
(3) *Se $f \sim g$, allora $I(f) = I(g)$.*

Dimostrazione. Se $f = 0$, allora $t^n \in I(f)$ per ogni $n \geq 1$. Sia $f \neq 0$ Se $m > n^2$, allora gli endomorfismi $f^0, f^1, \ldots, f^m$ sono linearmente dipendenti per cui esiste $p(t)$ di grado k $m \geq k \geq 1$ tale che $p(f) = 0$. Le proprietà che fanno di $I(f)$ un ideale di $\mathbb{K}[t]$ sono di verifica immediata, utilizzando anche il Lemma 1.1. Se $g = h \circ f \circ h^{-1}$, allora per ogni polinomio $p(t)$, $p(g) = h \circ p(f) \circ h^{-1}$. Poiché h è invertibile, $p(g) = 0$ se e solo se $p(f) = 0$.

□

Un polinomio minimo per f è un polinomio $q(t) \in I(f)$ di grado minimo tra i polinomi di grado ≥ 1 che stanno in $I(f)$.

Lemma 1.3. *$I(f)$ è generato da ogni suo polinomio minimo $q(t)$, cioè $I(f) = (q(t)) := \{p(t)q(t);\ p(t) \in \mathbb{K}[t]\}$.*

Dimostrazione. Sia $a(t) \in I(f)$. Eseguendo la divisione Euclidea ("con il resto") sapiamo che esistono unici polinomi $Q(t)$ e $R(t)$ tali che

$$a(t) = Q(t)q(t) + R(t), \quad \text{grado } R(t) < \text{grado } q(t)\ .$$

Poiché $I(f)$ è un ideale, $R(t) = a(t) - Q(t)q(t) \in I(f)$. Poiché $q(t)$ è un polinomio minimo, allora $R(t) = R$ è uno scalare. Infine $RId = a(f) - Q(f) \circ q(f) = 0 - Q(f) \circ 0 = 0$, implica necessariamente che $R = 0$. Dunque $a(t) = Q(t)q(t)$ come voluto.

□

Una conseguenza immediata del precedente Lemma è che esiste un unico polinomio minimo *monico* $q_f(t)$ (cioè con coefficiente del monomio di grado massimo uguale a 1). Questo $q_f(t)$ è detto "il"*polinomio minimo di f*. Segue immediatamente da quanto visto che

Corollario 1.4. *Se $f \sim g$, allora $q_f(t) = q_g(t)$ (cioè disponiamo di un altro invariante polinomiale oltre il già noto polinomio caratteristico $p_f(t)$).*

In effetti i due polinomi invarianti sono strettamente imparentati.

Proposizione 1.5. *$p_f(t) \in I(f)$ e quindi $q_f(t)$ divide $p_f(t)$.*

Dimostrazione. Lo dimostreremo nel caso in cui $f \in \mathcal{T}(V)$. Daremo poi alla fine una indicazione di come dimostrarlo in generale. Sia allora $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V con bandiera f–invariante, per cui $f(v_1) = \mu_1 v_1$, $f(v_2) = *v_1 + \mu_2 v_2$, $f(v_3) = *v_1 + *v_2 + \mu_3 v_3$, ecc., dove gli asterischi stanno a indicare qualche scalare che non ha importanza specificare, i μ_j sono gli autovalori di f e sono possibili ripetizioni. Sappiamo che

$$p_f(f) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_1 Id)$$

e che i fattori possono essere permutati a piacimento senza alterare il risultato della composizione. Basta dimostrare che $p_f(f)(v_j) = 0$ per ogni vettore di $\mathcal{B}$. Adesso $\pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_2 Id) \circ (f - \mu_1 Id)(v_1) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_2 Id)(0) = 0$; $\pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_3 Id) \circ (f - \mu_1 Id) \circ (f - \mu_2 Id)(v_2) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_3 Id) \circ (f - \mu_1 Id)(*v_1) = \pm(f - \mu_n Id) \circ \cdots \circ (f - \mu_3 Id)(0) = 0$. Iterando in modo induttivo la procedura si conclude appunto che $p_f(f)(v_j) = 0$ per ogni j, come voluto.

Vediamo come si può operare quando f non è necessariamente triangolabile. Consideriamo intanto il caso reale $\mathbb{K} = \mathbb{R}$. In questo caso possiamo usare la procedura della "complessificazione" descritta per esempio nella nota [3] sul TSH reperibile in nella sezione "Didattica" di

http://www.dm.unipi.it/ benedett

Allora $p_f(t)$ coincide con il polinomio caratteristico $p_{f_\mathbb{C}}(t)$ dell'endomorfismo complessificato $f_\mathbb{C} \in \text{End}\ (V_\mathbb{C})$. Poiché f è la restrizione di $f_\mathbb{C}$ alla parte reale V di $V_\mathbb{C}$, e $p_{f_\mathbb{C}}(f_\mathbb{C}) = 0$ (perchè su $\mathbb{C}$ tutti gli endomorfismi sono triangolabili), possiamo concludere che anche $p_f(f) = 0$. In generale, dato f, supponiamo esista un campo $\mathbb{F}$ che estende $\mathbb{K}$ (per cui $\mathbb{K}[t] \subset \mathbb{F}[t]$) tale che $p_f(t) \in \mathbb{K}[t]$ sia completamente fattorizzabile in $\mathbb{F}[t]$. Fissiamo una base $\mathcal{B}$ di V e sia A la matrice che rappresenta f in quel sistema di coordinate. Allora $p_f(f) = 0$ se e solo se $p_A(A) = 0$. Adesso $X \to AX$ definisce anche un endomorfismo di $\mathbb{F}^n$, che sappiamo essere triangolabile. Allora $p_A(A) = 0$ che comunque è l' identità tra matrici su $\mathbb{K} \subset \mathbb{F}$ che volevamo. L'esistenza di estensioni $\mathbb{K} \subset \mathbb{F}$ con questa proprietà è garantita da una teoria elementare che almeno gli studenti di Matematica dovrebbero incontrare nel corso di Algebra del primo anno della laurea triennale.

□

Lemma 1.6. *Se $p(t) \in I(f)$ e $\lambda \in \mathbb{K}$ è un autovalore di f allora $p(\lambda) = 0$.*

Dimostrazione. Sia $v \neq 0$, un autovettore tale che $f(v) = \lambda v$. Allora $0 = p(f)(v) = p(\lambda)v$. Poichè $v \neq 0$, necessariamente $p(\lambda) = 0$.

□

Tiriamo allora alcune conseguenze di quanto visto quando $f \in \mathcal{T}(V)$.

Corollario 1.7. *Se f è un endomorfismo triangolabile allora*

$$p_f(t) = \pm \prod_{j=1}^{k} (t - \lambda_j)^{m_j}$$

$$q_f(t) = \pm \prod_{j=1}^{k} (t - \lambda_j)^{r_j}$$

$$\lambda_j \neq \lambda_i, \text{ se } i \neq j;\ \sum_j m_j = n;\ 1 \leq r_j \leq m_j,\ j = 1, \dots k\ .$$

Dimostrazione. Il polinomio caratteristico è completamente fattorizzabile in $\mathbb{K}[t]$ perché f è triangolabile. Il polinomio minimo $q_f(t)$ è di quella forma per qualche $0 \leq r_j \leq m_j$, perché il polinomio minimo divide quello caratteristico. Infine ogni $r_j \geq 1$ perché ogni autovalore è radice di $q_f(t) \in I(f)$.

□

2. Decomposizione primaria

Si ricordi per esempio che f è diagonalizzabile se e solo se $V = \oplus V_{\lambda_j}(f)$ è la somma diretta degli autospazi di f, cioè è la somma diretta di sottospazi f–invarianti sui quali la restrizione di f è particolarmente semplice, trattandosi di un multiplo dell'identità. Analogamente, il programma che ci porterà infine alla "forma normale di Jordan" per un arbitrario endomorfismo f che sia solo triangolabile, si propone di decomporre in modo "canonico" V in somma diretta di sottospazi f–invarianti sui quali la restrizione di f sia la "più semplice possibile". Il "Teorema di decomposizione primaria" che andiamo a discutere è il passo fondamentale in quella direzione. Richiamiamo prima alcune nozioni riguardanti i polinomi (almeno gli studenti di Matematica potranno ancora una volta riferirsi al corso di Algebra). Dati due polinomi non nulli $a(t), b(t) \in \mathbb{K}[t]$, poniamo

$$I(a(t), b(t)) = \{p(t) = m(t)a(t) + n(t)b(t);\ m(t), n(t) \in \mathbb{K}[t]\}\ .$$

Non è difficile mostrare usando come prima la divisione Euclidea tra polinomi che

(1) $I = I(a(t), b(t))$ è un ideale in $\mathbb{K}[t]$ che contiene polinomi non nulli.
(2) Esiste un unico polinomio monico $d(t)$ di grado minimo tra i polinomi non nulli contenuti nell'ideale, e questo $d(t)$ genera $I = (d(t))$.
(3) Il polinomio $d(t)$ verifica le seguenti proprietà:
   - E' un divisore comune di $a(t)$ e $b(t)$, cioè $d(t)|a(t)$ e $d(t)|b(t)$;
   - Se $p(t)$ è un altro divisore comune allora $p(t)|d(t)$. Per questo motivo $d(t)$ è detto il "massimo comun divisore" $\text{MCD}(a(t), b(t))$

Segue in particolare dalla precedente discussione che il MCD (come tutti gli elementi di $I(a(t), b(t))$) può essere scritto nella forma

(Identità di Bezout) $d(t) = m(t)a(t) + n(t)b(t)$ per qualche $m(t), n(t) \in \mathbb{K}[t]$.

Possiamo infine enunciare

Proposizione 2.1. (Decomposizione primaria.) *Sia* $f \in \text{End}\ (V)$ *e sia* $p(t) \in I(f)$ *non nullo. Supponiamo che* $p(t) = a(t)b(t)$ *e che* $MCD(a(t), b(t)) = 1$ *(cioè* $p(t)$ *è prodotto di due fattori "coprimi"). Allora:*

(1) Si ha la decomposizione in somma diretta

$$V = \text{Ker}(a(f)) \oplus \text{Ker}(b(f))$$

e i due addendi diretti sono f*–invarianti.*

(2) Se $g = h \circ f \circ h^{-1}$*, così che* $V = \text{Ker}(a(f)) \oplus \text{Ker}(b(f)) = \text{Ker}(a(g)) \oplus \text{Ker}(b(g))$*, allora* $h(\text{Ker}(a(f))) = \text{Ker}(a(g))$ *e* $h(\text{Ker}(b(f))) = \text{Ker}(b(g))$*.*

Dimostrazione. Usando l'identità di Bezout abbiamo che

$$id = m(f) \circ a(f) + n(f) \circ b(f)$$

da cui per ogni $v \in V$

$$v = m(f) \circ a(f)(v) + n(f) \circ b(f)(v)\ .$$

Osserviamo che $b(f) \circ m(f) \circ a(f)(v) = m(f) \circ p(f)(v) = 0$ perché $p(t) \in I(f)$. Dunque $m(f) \circ a(f)(v) \in \text{Ker}(b(f))$. Analogamente $n(f) \circ b(f)(v) \in \text{Ker}(a(f))$. Abbiamo quindi dimostrato che $V = \text{Ker}(a(f)) + \text{Ker}(b(f))$. D'altra parte se $v \in \text{Ker}(a(f)) \cap \text{Ker}(b(f))$ $v = m(f) \circ a(f)(v) + n(f) \circ b(f)(v) = 0$, per cui è una somma diretta. Il fatto che gli addendi diretti siano f-invarianti è di semplice verifica usando ancora una volta il Lemma 1.1.

Venendo al punto (2), si ha per esempio $a(g) = h \circ a(f) \circ h^{-1}$ da cui per ogni $v \in V$, $a(g) \circ h(v) = h \circ a(f)(v)$ da cui ricaviamo l'inclusione $h(\mathrm{Ker}(a(f))) \subset \mathrm{Ker}(a(g))$. Per simmetria ricaviamo anche l'inclusione opposta. Si ragiona analogamente con $b(t)$. La Proposizione è dimostrata. □

Tiriamo le conseguenze quando $f \in \mathcal{T}(V)$. Poiché $p_f(t) \in I(f)$ è completamente fattorizzabile, e i fattori corrispondenti ad autovalori distinti sono chiaramente coprimi, applicando ripetutamente la Proposizione precedente abbiamo

Corollario 2.2. *Se f è triangolabile allora*

$$V = \oplus_{j=1}^{k} \mathrm{Ker}(f - \lambda_j Id)^{m_j}$$

ed ogni addendo diretto è f-invariante.

Vogliamo adesso cominciare ad analizzare le proprietà della restrizione di f ad ogni addendo diretto. Premettiamo alcune osservazioni che valgono per ogni endomorfismo.

Lemma 2.3. *Sia $h \in \mathrm{End}\,(V)$. Allora:*

(1) Per ogni $n \geq 1$, $\mathrm{Ker}(h^n) \subset \mathrm{Ker}(h^{n+1})$.

(2) Se $\mathrm{Ker}(h^n) = \mathrm{Ker}(h^{n+1})$ allora $\mathrm{Ker}(h^n) = \mathrm{Ker}(h^m)$ per ogni $m \geq n$.

(3) Se $p = k \circ h \circ k^{-1}$ è coniugato ad h, allora per ogni $n \geq 1$, $k(\mathrm{Ker}(h^n)) = \mathrm{Ker}(p^n)$. In particolare ogni dimensione $d_n = \dim \mathrm{Ker}(h^n)$ è invariante per coniugazione.

Dimostrazione. La prima affermazione è evidente. Per mostrare la seconda basta far vedere che $\mathrm{Ker}(h^{n+1}) = \mathrm{Ker}(h^{n+2})$, e concludere poi per induzione. In effetti basta dimostrare che vale l'inclusione $\mathrm{Ker}(h^{n+2}) \subset \mathrm{Ker}(h^{n+1})$. Ma $h^{n+2}(v) = h^{n+1}(h(v)) = 0$ se e solo se $h(v) \in \mathrm{Ker}(h^{n+1}) = \mathrm{Ker}(h^n)$. Dunque $h^n(h(v)) = h^{n+1}(v) = 0$ come voluto. Per dimostrare il punto (3) si ragiona come nella dimostrazione di (2) in Proposizione 2.1.

□

Proposizione 2.4. *Sia f un endomorfismo triangolabile con polinomio caratteristico e polinimio minimo come stabilito nel Corollario 1.7. Sia λ un autovalore di f di molteplicità m rispetto al polinomio caratteristico, di molteplicità r rispetto al polinomio minimo. Indichiamo con $W_\lambda := \mathrm{Ker}(f - \lambda Id)^m$ l'addendo diretto corrispondente a λ nella decomposizione primaria associata al polinomio caratteristico $p_f(t)$. Indichiamo con $g_\lambda = f|W_\lambda$. Allora:*

(1) $\dim W_\lambda = m$, $p_{g_\lambda}(t) = \pm(t-\lambda)^m$, $q_{g_\lambda}(t) = \pm(t-\lambda)^r$.

(2) Ogni dimensione $d_n = \dim \mathrm{Ker}(f - \lambda Id)^n$ è invariante per coniugazione; $\dim V_\lambda(f) = d_1 < d_2 \cdots < d_r = \dim W_\lambda = m$.

(3) La decomposizione primaria associata al polinomio caratteristico di f coincide con la decomposizione primaria associata al polinomio minimo.

Dimostrazione. Applicando il Lemma precedente a $h := (f - \lambda Id)$ vediamo che la stringa dei successivi nuclei è contenuta in W_λ; inizia con l'autospazio $V_\lambda(f)$ e termina con tutto W_λ. Ne segue che ogni g_λ ammette solo l'autovalore λ. D'altra parte, usando un sistema di coordinate adattato alla decomposizione primaria, si vede immediatamente che $p_f(t) = \prod_\lambda p_{g_\lambda}(t)$ da cui necessariamente $p_{g_\lambda}(t) = \pm(t-\lambda)^m$, e quindi $\dim W_\lambda = m$. Se uno dei g_λ avesse polinomio minimo di grado strettamente minore di r, allora $\prod_\lambda q_{g_\lambda}(t)$ sarebbe un polinomio appartenente a $I(f)$ di grado ≥ 1 e strettamente minore del grado di $q_f(t)$, che è impossibile. Il punto (1) è così dimostrato. La prima affermazione del punto (2) segue dal punto (3) del Lemma 2.3 osservando che $f \sim f'$ se e solo se $(f - \alpha Id) \sim (f' - \alpha Id)$ per ogni scalare α. La seconda segue dal precedente punto (1) e dal punto (2) del Lemma 2.3. Il punto (3) è ora evidente. □

Riassumenndo quanto fatto, lo studio degli endomorfismi triangolabili a meno di coniugazione è ridotto allo studio degli endomorfismi "speciali" $g : W \to W$ tali che $p_g(t) = \pm(t-\lambda)^m$ (per cui $\dim W = m$), e $p_g(t) = \pm(t-\lambda)^r$, per qualche $1 \leq r \leq m$. Inoltre posto $h := g - \lambda Id$, la successione di dimensioni strettamente crescenti $\dim V_\lambda(g) = d_1 < d_2 \cdots < d_r = \dim W = m$ è invariante per coniugazione. Tutte queste informazioni sono incorporate nella stringa di invarianti numerici

$$[\lambda, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)] \; .$$

Mostreremo che questo è l'invariante completo che cercavamo.

Come già osservato, se g e g' sono endomorfismi speciali su W con lo stesso autovalore λ, allora $g \sim g'$ se e solo se $h \sim h'$. La dimostrazione del seguente Lemma è immediata.

Lemma 2.5. *Se g è speciale su W con con stringa di invarianti $[\lambda, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$, allora $h = g - \lambda Id$ è speciale su W con stringa di invarianti $[0, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$.*

Un endomorfismo speciale h con autovalore $\lambda = 0$ è detto *nilpotente*. Se g è speciale di autovalore λ, allora $h = g - \lambda Id$ è detto la parte nilpotente di g. Possiamo allora riassumere quanto osservato qui sopra dicendo che non è restrittivo assumere che $\lambda = 0$, cioè che tutto lo studio è ridotto al caso degli endomorfismi nilpotenti. Osserviamo anche che se W viene decomposto in somma diretta di sottospazi h-invarianti, allora la restizione di h ad ogni addendo diretto è nilpotente.

Cominciamo con l' analizzare i due casi estremi: $r = 1$, $r = m$.

$r = 1$: questo equivale a dire che la stringa invariante è $[0, 1, (m)]$, cioè che $W = V_0(h)$ e siamo nel caso diagonalizzabile.

$r = m$:

Definizione 2.6. Una base di W è detta *ciclica* per h nilpotente se è della forma $\mathcal{B} = \{h^{m-1}(v), h^{m-2}(v), \dots, h(v), v\}$ per qualche $v \in W$.

La dimostrazione del seguente lemma è immediata.

Lemma 2.7. *Indicate con E^j, $j = 1, \dots, m$, le colonne della matrice identità $m \times m$. Allora $\mathcal{B}$ è ciclica per h se e solo se le colonne della matrice $M_\mathcal{B}(h)$ sono $C^1 = 0$, $C^2 = E^1$, ..., $C^m = E^{m-1}$.*

La matrice $m \times m$ della forma descritta nel Lemma è indicata con $J(0, m)$ ed è detta *blocco di Jordan di taglia m e autovalore* 0. Analogamente, per ogni $\lambda \in \mathbb{K}$, poniamo $J(\lambda, m) = \lambda I + J(0, m)$ il corrispondente blocco di Jordan di autovalore λ.

Possiamo infine enunciare

Lemma 2.8. *Sia h nilpotente. Allora sono fatti tra loro equivalenti:*

(1) $p_h(t) = q_h(t) = t^m$;

(2) h ha stringa invariante $[0, m, (1, 2, 3, \dots, m)]$.

(3) Esiste una base ciclica per h.

Dimostrazione. (1) è equivalente a (2) perché le dimensioni sono forzate a crescere ad ogni passo di 1. E' evidente che (3) implica (1). Dimostriamo l'altra implicazione. Poiché $r = m$, esiste sicuramente v tale che $h^{m-1}(v) \neq 0$. Basta allora dimostrare che gli m vettori di $\mathcal{B} = \{h^{m-1}(v), h^{m-2}(v), \dots, h(v), v\}$ sono linearmente indipendenti. Sia allora $\sum_{j=0}^{m-1} a_j h^j(v) = 0$ e mostriamo che necessariamente tutti gli $a_j = 0$. Applicando h, ricaviamo la relazione $\sum_{j=1}^{m-1} a_{j-1} h^j(v) = 0$. Iterando l'applicazione di h otteniamo un sistema di relazioni che termina con $a_0 h^{m-1}(v) = 0$. Poiché $h^{m-1}(v) \neq 0$, necessariamente $a_0 = 0$, e risalendo in modo induttivo le relazioni trovate si conclude che tutti i coefficienti sono nulli.

□

Abbiamo così trovato i primi due casi di forma normale di Jordan per gli endomorfismi speciali e nilpotenti:

Il blocco $J(0, m)$ corrispondente alla stringa $[0, m, (1, 2, 3, \dots, m)]$;

La matrice diagonale ("a blocchi") formata da m blocchi di Jordan $J(0, 1)$.

Possiamo infine enunciare il teorema della forma normale di Jordan per gli endomorfismi nilpotenti in generale.

Theorem 2.9. *Sia h un endomorfismo nilpotente su W con stringa invariante* $[0, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$. *Allora:*

(1) Esiste una decomposizione di $W = \oplus Z_j$ in somma diretta di sottospazi h-invarianti tale che la restrizione di h ad ogni Z_j ammette una base ciclica. Indicata con $\mathcal{B}$ la base di W ottenuta unendo tali basi cicliche, allora $M_{\mathcal{B}}(h)$ è una matrice diagonale a blocchi dove ogni blocco è di Jordan del tipo $J(0, s_j)$, $\sum_j s_j = m$. Questa $M_{\mathcal{B}}(h)$ è detta una forma di Jordan per h, mentre $\mathcal{B}$ è detta una base di Jordan per h.

(2) La forma di Jordan di h è unica a meno di permutazione dei blocchi ed è completamente ed esplicitamente determinabile in funzione soltanto della stringa invariante $[0, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$, *che risulta quindi essere un invariante completo.*

Dimostrazione. Consideriamo un frammento a tre termini del tipo $\mathrm{Ker}(h^{j-2}) \subset \mathrm{Ker}(h^{j-1}) \subset \mathrm{Ker}(h^j)$, $1 \leq j-2$, $j \leq r$, della successione strettamente crescente di nuclei. Illustriamo qui la costruzione fondamentale che porterà alla dimostrazione del Teorema. Fissiamo una decomposizione in somma diretta

$$\mathrm{Ker}(h^j) = \mathrm{Ker}(h^{j-1}) \oplus U$$

dove $t = \dim U = d_j - d_{j-1}$, e fissiamo una base $\{u_1, \dots u_t\}$ di U. Applichiamo h ottenendo $\{h(u_1), \dots h(u_t)\}$. Affermiamo che valgono i seguenti fatti:

(1) $\{h(u_1), \dots h(u_t)\}$ è contenuto in $\mathrm{Ker}(h^{j-1})$.
(2) I vettori in $\{h(u_1), \dots h(u_t)\}$ sono lineramente indipendenti.
(3) $\mathrm{Ker}(h^{j-2}) \cap \mathrm{Span}\{h(u_1), \dots h(u_t)\} = \{0\}$.

Dimostriamo queste affermazioni. La prima è evidente. Sia $\sum_i a_i h(u_i) = 0$. Allora $\sum_i a_i u_i \in \mathrm{Ker}(h) \subset \mathrm{Ker}(h^{j-1})$ ma allora tutti i coefficienti sono nulli perché gli u_i sono linearmente indipendenti e $\mathrm{Ker}(h^{j-1}) \cap U = \{0\}$. Questo prova il secondo punto. La prova del terzo è del tutto simile.

Procediamo adesso alla costruzione passo per passo della base $\mathcal{B}$ di W con le proprietà volute. Il lettore deve immaginare che i vettori di questa base saranno arrangiati in una tabella a gradini, che scendono verso destra di altezza e lunghezza variabili. Daremo una dopo l'altra le righe di questa tabella di vettori. Queste righe saranno allineate sul bordo sinistro della tabella. La lunghezza e l'altezza di ciascun gradino sarà esplicitamente calcolabile solamente in funzione della stringa invariante di h.

Partiamo dal frammento a tre termini all'estremità destra: $\mathrm{Ker}(h^{r-2}) \subset \mathrm{Ker}(h^{r-1}) \subset \mathrm{Ker}(h^r)$. Applichiamo la costruzione fondamentale. La prima riga della tabella di vettori sarà data da $(u_1, \dots, u_t)$ dove $t = m - d_{r-1}$. Applichiamo h e otteniamo un primo segmento $(h(u_1), \dots, h(u_t))$ della seconda riga. Consideriamo adesso il frammento a tre termini ottenuto scivolando di una posizione verso sinistra: $\mathrm{Ker}(h^{r-3}) \subset \mathrm{Ker}(h^{r-2}) \subset \mathrm{Ker}(h^{r-1})$ e applichiamo ancora una volta la costruzione fondamentale *raffinandola però come segue*: imponiamo che la base di U sia della forma $h(u_1), h(u_2), \dots, h(u_t), u_{t+1}, \dots, u_{t'}$, e questi vettori formeranno la seconda riga della tabella, mentre $h^2(u_1), h^2(u_2), \dots, h^2(u_t), h(u_{t+1}), \dots, h(u_{t'})$ sarà il segmento iniziale della terza riga. Questo è possibile perché grazie al fatto (3) osservato prima, $\mathrm{Span}\{h(u_1), h(u_2), \dots, h(u_t)\}$ può far parte di un complementare U di $\mathrm{Ker}(h^{r-2})$ in $\mathrm{Ker}(h^{r-1})$. Notiamo che $t' = t + t" = d_{r-1} - d_{r-2}$, $t = d_r - d_{r-1}$, per cui $0 \leq t" = t' - t = 2d_{r-1} - (d_{r-2} + d_r)$. Se $t" = 0$, allora le prime due righe fanno parte dello stesso gradino di altezza maggiore o uguale a 2 e lunghezza uguale a t. Se $t" > O$, allora il primo gradino ha lunghezza uguale a t e altezza uguale a 1, mentre il secondo gradino ha lunghezza maggiore o uguale a $t"$. Consideriamo adesso il frammento a tre termini ottenuto scivolando di un' altra posizione verso sinistra. E applichiamo ancora una volta la versione raffinata della costruzione fondamentale. Iteriamo la procedura fino a quando è possibile. Ne risulta una tabella a gradini di vettori che verifica le proprietà dette in precedenza. Per costruzione l'unione di questi vettori forma una base $\mathcal{B}$ di W. D'altra parte, leggendo dal basso verso l'alto ciascuna colonna della tabella vediamo vettori indipendenti del tipo $\{h^j(u), h^{j-1}(u), \dots, u\}$ tali che, ancora per costruzione, $h^{j+1}(v) = 0$. Dunque lo spazio generato da ogni colonna è h-invariante e munito per costruzione di una base ciclica per la restrizione di h. Dunque l'esistenza di una forma di Jordan per h è dimostrata. Riguardo all'unicità. A meno di permutazione dei blocchi (cioè a meno di permutazioni dei vettori di una "base di Jordan" per h) possiamo assumere che i blocchi siano disposti lungo la diagonale in modo che la taglia decresca. Possiamo allora organizzare i vettori della base in una tabella a gradini con le proprietà formali della costruzione precedente, dove l'altezza e la lunghezza di ogni gradino è esplicitamente esprimibile in funzione delle taglie dei blocchi. A questo punto è facile convincersi che questa tabella è il risultato di una implementazione della costruzione che ha portato alla dimostrazione dell'esistenza, per cui la lunghezza e l'altezza di ogni gradino (quindi la taglia di ogni blocco) è esplicitamente esprimibile in funzione solamente della stringa invariante di h. Il teorema è così dimostrato. □

NO PARTE QUI SOTTO VEDI DOPO

Osservazioni 2.10. A parte quelle relative al primo gradino, volutamente non abbiamo esplicitato l' espressione delle lunghezze e delle altezze dei gradini in funzione di $[0, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$. Può essere un utile esercizio per il lettore farlo, seguendo la costruzione induttiva di una base di Jordan. Può essere anche istruttivo ricavare da $[0, r, (d_1 < d_2 < \cdots < d_{r-1} < d_r = m)]$ informazioni parziali relative alla forma normale di Jordan. Per esempio il lettore può dimostrare per esercizio che il numero totale dei blocchi è uguale a d_1, mentre la taglia massima al variare dei blocchi è uguale a r.

Per completezza enunciamo la versione generale del Teorema sulla forma canonica di Jordan per endomorfismi triangolabili arbitrari (non necessariamente nilpotenti). La dimostrazione è una conseguenza immediata della discussione precedente.

Teorema 2.1. *Sia f un endomorfismo triangolabile su V. Possiamo associare a f l'invariante per coniugazione $\mathcal{J}(h)$ dato dall'insieme delle stringhe invarianti $[\lambda, r_\lambda, (d_1^\lambda < \cdots < d_{r_\lambda}^\lambda = m_\lambda)]$, al variare di λ nello spettro di f. Allora:*

(1) Esiste una base $\mathcal{B}$ (detta di Jordan per f) tale che $M_{\mathcal{B}}(f)$ è una matrice diagonale a blocchi tale che ogni blocco è della forma $J(\lambda, s)$ dove λ varia nello spettro di f. Tale matrice è detta una forma di Jordan per f.

(2) La forma di Jordan è unica a meno di permutazione dei blocchi ed è completamente e esplicitamente determinata in funzione dell'invariante $\mathcal{J}$ che risulta quindi un invariante completo.

AGGIUNTA SU FORMA CANONICA DI JORDAN DI ENDOMORFISMI

Notiamo che il risultato precedente afferma che un **endomorfismo nilpotente**, che ha dunque spettro costituito solo da $\lambda = 0$, ammette sempre una base di Jordan, dove ciascun blocco e' della forma $J_k(0)$, per qualche $1 \leqslant k \leqslant n = \dim_{\mathbb{K}(V)}$, e dove la somma degli ordini dei blocchi e' n. Dimostriamo il risultato principale di queste note.

Teorema [Forma canonica di Jordan] *Sia V un $\mathbb{K}$-spazio vettoriale,* $\dim_{\mathbb{K}}(V) = n$, *e sia* $f \in \mathrm{End}_{\mathbb{K}}(V)$ *un endomorfismo tale che* $\mathrm{Sp}(f) = \{\lambda_1, \ldots, \lambda_k\} \subset \mathbb{K}$. *Allora esiste sempre una base di Jordan $\mathcal{J}$ per f.*

Piu' precisamente, se consideriamo:

• *il polinomio caratteristico di f,*

$$P_f(x) := \Pi_{j=1}^{k}(\lambda_j - x)^{m_j},$$

con $m_j := m_a(f, \lambda_j)$ la molteplicita' algebrica di λ_j, $1 \leqslant i \leqslant k$

• *il polinomio minimo di f*

$$q_f(x) := \Pi_{j=1}^{k}(x - \lambda_j)^{r_j},$$

dove r_j esponente tale che $1 \leqslant r_j \leqslant m_j = m_a(f, \lambda_j)$, $1 \leqslant j \leqslant k$, e

• *$m_g(f, \lambda_j)$ la molteplicita' geometrica dell'autovalore λ_j, $1 \leqslant j \leqslant k$.*

Allora, nella matrice $M_{\mathcal{J},\mathcal{J}}(f)$ che rappresenta f in forma canonica di Jordan*:*

(a) esiste almeno un blocco di Jordan di taglia r_j, i.e. della forma $J_{r_j}(\lambda_j)$ e tutti gli altri eventuali blocchi di Jordan presenti e relativi all'autovalore λ_j sono di ordine minore od uguale ad r_j, per ogni $1 \leqslant j \leqslant k$;

(b) il numero dei blocchi distinti di Jordan relativi al medesimo autovalore λ_j e' uguale alla molteplicita' geometrica $m_g(f, \lambda_j)$ dell'autovalore λ_j, per ogni $1 \leqslant j \leqslant k$;

(c) la somma degli ordini dei differenti $m_g(f, \lambda_j)$ blocchi di Jordan relativi al medesimo autovalore λ_j e' uguale alla molteplicita' algebrica $m_j = m_a(f, \lambda_j)$ dell'autovalore λ_j, per ogni $1 \leqslant j \leqslant k$.

Effettivamente vale un risultato ancora piu' forte come visto nel caso nilpotente, cioe' non solo un endomorfismo f come sopra puo' essere messo sempre in una opportuna forma canonica di Jordan, ma inoltre tale forma canonica di Jordan e' unica a meno di una permutazione dei blocchi (i.e. a meno di un diverso riordinamento dei vettori di una base di Jordan). Dimostrare l'unicita' richiederebbe ancor piu' tecnicismi e questo esula dall'obiettivo di queste note. In ogni caso le affermazioni (a), (b) e (c) nell'enunciato del Teorema [Forma canonica di Jordan] forniscono sufficienti limitazioni da dedurre spesso in modo univoco la forma canonica di Jordan di endomorfismi con spettro nel campo.

Dimostrazione del Teorema [Forma canonica di Jordan.] Dalle ipotesi, grazie al **Corollario 2.2** abbiamo che

$$V = G_{\lambda_1}(f) \oplus G_{\lambda_2}(f) \oplus \cdots \oplus G_{\lambda_m}(f)$$

dove ciascun

$$G_{\lambda_j}(f) := \mathrm{Ker}((\mathrm{f} - \lambda_\mathrm{j}\mathrm{Id})^{\mathrm{m_j}})$$

e' sottospazio vettoriale f-invariante, $1 \leqslant j \leqslant k$, i cui vettori vengono anche detti **autovettori generalizzati** dell'endomorfismo f rispetto all'autovalore λ_j, $1 \leqslant j \leqslant k$.

Dalla **Proposizione 2.4 (1)**, si ha che

$$\dim_{\mathbb{K}}(G_{\lambda_j}(f)) = m_a(f, \lambda_j) = m_j, \;\; 1 \leqslant j \leqslant k.$$

Pertanto se

$$\mathcal{B}_j := \{\underline{v}_{j,1}, \ldots, \underline{v}_{j,\, m_a(f,\lambda_j)}\}, \; 1 \leqslant j \leqslant k$$

e' una base del sottospazio $G_{\lambda_j}(f)$, allora

$$\mathcal{B} := \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_m$$

e' una base di V rispetto alla quale $M_{\mathcal{B},\,\mathcal{B}}(f)$ e' una matrice diagonale a blocchi

$$M_{\mathcal{B},\,\mathcal{B}}(f) = \begin{pmatrix} A_1 & O & O & \cdots & O \\ O & A_2 & O & \cdots & O \\ O & O & A_3 & \cdots & O \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ O & O & O & \cdots & A_m \end{pmatrix}$$

dove ciascuna matrice A_j e' quadrata di ordine $m_a(f, \lambda_j)$ e rappresenta in base $\mathcal{B}_j$ l'endomorfismo $f|_{G_{\lambda_j}(f)}$, i.e. $A_j = M_{\mathcal{B}_j,\,\mathcal{B}_j}(f|_{G_{\lambda_j}(f)})$, $1 \leqslant j \leqslant m$.

Per dimostrare la prima parte dell'enunciato e' allora sufficiente dimostrare che si possono scegliere opportunamente le basi $\mathcal{B}_j$ di modo tale che ogni matrice A_j sia rappresentata in forma canonica di Jordan relativamente all'autovalore λ_j, i.e. che A_j sia diagonale a blocchi dove ciascun blocco e' della forma

$$J_h(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}, \tag{0.1}$$

per qualche ordine (o taglia) $h \leqslant m_a(f, \lambda_j) = m_j$ relativamente all'autovalore λ_j, $1 \leqslant j \leqslant k$.

Per ogni $j \in \{1, \ldots, m\}$, denotiamo con

$$h_j := (f - \lambda_j Id_V)|_{G_{\lambda_j}(f)} \in \mathrm{End}_{\mathbb{K}}(G_{\lambda_j}(f))$$

e, contestualmente,

$$f_j := f|_{G_{\lambda_j}(f)} \in \mathrm{End}_{\mathbb{K}}(G_{\lambda_j}(f)).$$

Con queste notazioni, osserviamo che

$$f_j = h_j + \lambda_j Id_{G_{\lambda_j}(f)}, \; 1 \leqslant j \leqslant m.$$

Ricordando ora che $\dim_{\mathbb{K}}(G_{\lambda_j}(f)) = m_a(f, \lambda_j) \leqslant n$ e che $G_{\lambda_j}(f)$ e' f-invariante, si ha

$$G_{\lambda_j}(f) = \mathrm{Ker}((f_j - \lambda_j Id_{G_{\lambda_j}(f)})^{m_j}) = \mathrm{Ker}((g_j)^{m_j})$$

con h_j endomorfismo nilpotente su $G_{\lambda_j}(f)$ e $m_j = m_a(f, \lambda_j)$, $1 \leqslant j \leqslant k$.

La dimostrazione del **Teorema 2.9** ci assicura l'esistenza di una base ciclica $\mathcal{J}_j$ di $G_{\lambda_j}(f)$ rispetto alla quale l'endomorfismo nilpotente h_j si scrive in forma canonica di Jordan, con tutti blocchi della forma

$$A_j = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{0.2}$$

$1 \leqslant j \leqslant k$. Rispetto a questa stessa base, l'endomorfismo

$$f_j = h_j + \lambda_j Id_{G_{\lambda_j}(f)}$$

si scrive in forma canonica di Jordan, con tutti blocchi dello stesso ordine di quelli di h_j ma della forma (0.1), con $1 \leqslant j \leqslant k$. Pertanto, una base $\mathcal{J}$ come nella prima parte dell'enunciato sara' data da

$$\mathcal{J} := \mathcal{J}_1 \cup \mathcal{J}_2 \cup \cdots \cup \mathcal{J}_k.$$

Per quanto riguarda (c) nella seconda parte dell'enunciato, la somma degli ordini dei blocchi di Jordan relativi all'autovalore λ_j deve essere uguale all'ordine della matrice $M_{\mathfrak{J}_j, \mathfrak{J}_j}(f_j)$ che eguaglia la dimensione di $G_{\lambda_j}(f)$ che e' pari a $m_a(f, \lambda_j)0m_j$, $1 \leqslant j \leqslant k$.

Per quanto riguarda (b), la costruzione della base nella dimostrazione di Teorema 2.9 fa si che, partendo dai termini precedenti della catena di inclusioni proprie per costruire la base ed applicando successivamente le potenze di h, gli elementi $h_j^{\nu-1}(\underline{v}_{i,k})$, dove $\nu := \nu(h_j)$ denota l'indice di nilpotenza di h_j, insieme con i vettori $\underline{v}_{1,k}$ costituivano una base di $\mathrm{Ker}(h_j^{\nu-1}) = V_{\lambda_j}(f)$. Pertanto, visto che in ogni blocco di Jordan la prima colonna corrisponde ad un autovettore di f_j rispetto all'autovalore λ_j, abbiamo tanti blocchi distinti quanti sono gli autovettori nella base e dunque si hanno $m_g(f, \lambda_j) = \dim_{\mathbb{K}}(V_{\lambda_j}(f))$ blocchi distinti.

Per quanto riguarda (a), si tratta di far vedere che l'indice di stazionarieta' di g_j coincide con l'esponente r_j nel binomio corrispondente a λ_j nella fattorizzazione del polinomio minimo $q_f(x)$, dove ricordiamo che $r_j \leqslant m_a(f, \lambda_j) = m_j \leqslant n$, $1 \leqslant j \leqslant k$. Infatti se cosi' e', visto che

$$h_j = f_j - \lambda_j Id_{G_{\lambda_j}(f)},$$

avremo la catena di sottospazi

$$\{\underline{0}\} = \mathrm{Ker}(h_j^0) \subsetneq \mathrm{Ker}(h_j) = V_{\lambda_j}(f) \subsetneq \mathrm{Ker}(h^2) \subsetneq \cdots \subsetneq \mathrm{Ker}(h_j^{r_j-1}) \subsetneq \mathrm{Ker}(h_j^{r_j}) = \mathrm{Ker}(h_j^{r_j+1}) =$$

$$= \mathrm{Ker}(h_j^{r_j+2}) = \cdots = \mathrm{Ker}(h_j^{m_j}) = \cdots = \mathrm{Ker}(h_j^n) = G_{\lambda_j}(f)$$

e, dalla costruzione della base nella dimostrazione del **Teorema 2.9**, questo comportera' l'esistenza di almeno un blocco di Jordan di ordine r_j nella forma canonica di Jordan per f_j e che tutti gli altri blocchi di Jordan hanno ordine al piu' e_j.

Per dimostrare che l'indice di stazionarieta' di h_j e' proprio r_j, facciamo vedere che il polinomio minimo $q_{f_j}(x)$ dell'endomorfismo $f_j \in \mathrm{End}_{\mathbb{K}}(G_{\lambda_j}(f))$ e' proprio $q_{f_j}(x) = (x - \lambda_j)^{r_j}$; se cosi' e', per definizione di polinomio minimo, $q_{f_j}(f_j) = 0$ su $G_{\lambda_j}(f)$, i.e. $G_{\lambda_j}(f) = \mathrm{Ker}(q_{f_j}(f_j))$ e la potenza r_j e' la minima possibile per cui cio' accada.

Prima dimostriamo che $G_{\lambda_j}(f) = \mathrm{Ker}(h_j^{r_j})$, $1 \leqslant j \leqslant k$. Per semplicita' supponiamo $j = 1$ (la dimostrazione per un $j \in \{1, \ldots, k\}$ generale e' analoga). Consideriamo i polinomi

$$(x - \lambda_1)^{r_1} \;\text{ e }\; s(x) := \Pi_{j=2}^{m}(x - \lambda_j)^{r_j}.$$

Poiche' essi non hanno fattori comuni, i.e. sono coprimi in $\mathbb{K}[x]$, per l'algoritmo di divisione euclidea con resto, esistono polinomi $a(x), b(x) \in \mathbb{K}[x]$ tali che

$$1 = a(x)(x - \lambda_1)^{r_1} + b(x)s(x).$$

Per ogni vettore $\underline{v} \in V$ si ha:

$\underline{v} = \mathrm{Id}_V(\underline{v}) = (a(f) \circ (f - \lambda_1 Id_V)^{r_1} + b(f) \circ s(f))(\underline{v}) = a(f)((f - \lambda_1 Id_V)^{r_1}(\underline{v})) + b(f)(s(f)(\underline{v}))$.

Siccome

$s(f)(a(f)((f - \lambda_1 Id_V)^{r_1}(\underline{v}))) = a(f)(s(f)((f - \lambda_1 Id_V)^{r_1}(\underline{v}))) = a(f)(m_f(f)(\underline{v}) = a(f)(\underline{0}) = \underline{0}$,

visto che $m_f(f) = 0$ su V e visto che $a(f)$ endomorfismo, si ha allora che $a(f)((f - \lambda_1 Id_V)^{r_1}(\underline{v}) \in \mathrm{Ker}(s(f))$. In modo analogo si dimostra che $b(f)(s(f)(\underline{v}) \in \mathrm{Ker}((f - \lambda_1 Id_V)^{r_1})$. Pertanto

$$V = \mathrm{Ker}((f - \lambda_1 Id_V)^{r_1}) + \mathrm{Ker}(s(f)).$$

Se ora si considerano $\underline{u} \in \mathrm{Ker}((f - \lambda_1 Id_V)^{r_1})$ e $\underline{w} \in \mathrm{Ker}(s(f))$ tali che $\underline{u} + \underline{w} = \underline{0}$, allora

$$\underline{0} = (f - \lambda_1 Id_V)^{r_1}(\underline{0}) = (f - \lambda_1 Id_V)^{r_1}(\underline{u} + \underline{w}) = (f - \lambda_1 Id_V)^{r_1}(\underline{u}) + (f - \lambda_1 Id_V)^{r_1}(\underline{w}).$$

D'altra parte $\underline{u} \in \mathrm{Ker}((f - \lambda_1 Id_V)^{r_1})$ comporta che la precedente eguaglianza sia

$$\underline{0} = (f - \lambda_1 Id_V)^{e_1}(\underline{w}),$$

che implica

$$\underline{w} = \mathrm{Id}_V(\underline{w}) = (a(f) \circ (f - \lambda_1 Id_V)^{r_1} + b(f) \circ s(f))(\underline{w}) = (b(f)(s(f)(\underline{w})) = \underline{0}$$

visto che $\underline{w} \in \mathrm{Ker}(s(f))$. Ma allora si ha anche $\underline{u} = 0$ e dunque

$$V = \mathrm{Ker}((f - \lambda_1 Id_V)^{r_1}) \oplus \mathrm{Ker}(s(f)).$$

Ricordiamo pero' che si ha

$$V = G_{\lambda_1}(f) \oplus G_{\lambda_2}(f) \oplus \cdots \oplus G_{\lambda_m}(f),$$

con $G_{\lambda_1}(f) = \mathrm{Ker}((f - \lambda_1 Id_V)^{m_1})$ e con $\mathrm{Ker}((f - \lambda_1 Id_V)^{r_1}) \subseteq G_{\lambda_1}(f)$, essendo $r_1 \leqslant m_1 = m_a(f, \lambda_1) \leqslant n$. Pertanto, visto che $\mathrm{Ker}(s(f))$ non contiene autovettori generalizzati di f relativi a λ_1, essendo $s(x) = \Pi_{j=2}^{k}(x - \lambda_j)^{r_j}$, si ha piu' precisamente che

$$\mathrm{Ker}((f - \lambda_1 Id_V)^{r_1}) = \mathrm{Ker}((f - \lambda_1 Id_V)^{m_1}) = G_{\lambda_1}(f)$$

e

$$\mathrm{Ker}(s(f)) = G_{\lambda_2}(f) \oplus \cdots \oplus G_{\lambda_k}(f);$$

in particolare, l'indice di stazionarieta' di h_1 e' almeno r_1. La stessa conclusione si ottiene per ogni $G_{\lambda_j}(f)$ e per ogni h_j, $1 \leqslant j \leqslant r$.

Ora vogliamo dimostrare che il polinomio minimo di f_j e' $q_{f_j}(x) = (x - \lambda_j)^{r_j}$, $1 \leqslant j \leqslant r$. Visto che $G_{\lambda_j}(f) = \mathrm{Ker}((f - \lambda_j Id_V)^{r_j})$, ne segue che $(f_j - \lambda_j Id_V)^{r_j}$ e' identicamente nullo su $G_{\lambda_j}(f)$; si ha dunque che il polinomio minimo $q_{f_j}(x)$ divide $(x - \lambda_j)^{r_j}$ in $\mathbb{K}[x]$, $1 \leqslant j \leqslant k$. Si ha allora che $\Pi_{j=1}^{k} q_{f_j}(x)$ divide in $\mathbb{K}[x]$ il polinomio $\Pi_{j=1}^{k}(x - \lambda_j)^{r_j} = q_f(x)$, dove $q_f(x)$ polinomio minimo di f. D'altra parte, ogni $\underline{v} \in V$ si scrive in modo unico come

$$\underline{v} = \underline{v}_1 + \cdots + \underline{v}_k, \ \underline{v}_j \in G_{\lambda_j}(f), \ 1 \leqslant j \leqslant k.$$

Pertanto

$$\Pi_{i=1}^{k} q_{f_i}(f)(\sum_{j=1}^{k} \underline{v}_j) = \sum_{j=1}^{k} (\Pi_{i=1}^{k} q_{f_i}(f))(\underline{v}_j) = \sum_{j=1}^{k} (\Pi_{i \neq j}^{k} q_{f_i}(f))(q_{f_j}(f)(\underline{v}_j)) =$$

$$= \sum_{j=1}^{k} (\Pi_{i \neq j}^{k} q_{f_i}(f))(q_{f_j}(f_j)(\underline{v}_j)) = \sum_{j=1}^{k} (\Pi_{i \neq j}^{k} q_{f_i}(f))(\underline{0}) = \underline{0};$$

dunque l'endomorfismo $\Pi_{i=1}^{k} q_{f_i}(f)$ e' identicamente nullo su V e si ha che $q_f(x)$ deve dividere $\Pi_{i=1}^{k} q_{f_i}(f)$ in $\mathbb{K}[x]$. Da quanto visto precedentemente, segue che $q_f(x) = \Pi_{j=1}^{k} q_{f_j}(x)$ in $\mathbb{K}[x]$.

Visto che $q_f(x) = \Pi_{j=1}^{k}(x - \lambda_j)^{r_j}$, che $q_{f_j}(x)$ divide $(x - \lambda_j)^{r_j}$, $1 \leqslant j \leqslant k$, e vista l'eguaglianza $q_f(x) = \Pi_{j=1}^{k} m_{f_j}(x)$ appena dimostrata, segue che $q_{f_j}(x) = (x - \lambda_j)^{r_j}$, per ogni $1 \leqslant j \leqslant k$, come desiderato. $\square$