Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Triennale in Matematica

IL DIAMOND LEMMA E IL TEOREMA DI POINCARE'-BIRKHOFF-WITT SUGLI ANELLI

Tesi di Laurea in Algebra

Relatore:
Chiar.ma Prof.
Fioresi Rita

Presentata da:
Ronchetti Niccolò

Sessione I
Anno Accademico 2009/10

Introduzione

In questa tesi presentiamo una dimostrazione del Teorema di Poincaré-Birkhoff-Witt sugli anelli ottenuta sfruttando il Diamond Lemma, un importante risultato di Bergman.

Nel capitolo 1 introduciamo il concetto di algebra di Lie su di un anello, dando alcuni tra gli esempi più significativi, e il concetto di rappresentazione di un'algebra di Lie. La struttura algebrica più importante legata ad un'algebra di Lie è la sua algebra inviluppante universale, di cui diamo una costruzione dettagliata e una definizione alternativa tramite la proprietà universale.

Nel capitolo 2 discutiamo in dettaglio il Diamond Lemma, nella formulazione algebrica data da Bergman in [2], ispirata al Diamond Lemma di Newman in teoria dei grafi ([5]).
Il Diamond Lemma di Bergman si occupa del seguente problema. Sull'algebra libera $k\langle X\rangle$, con k anello commutativo unitario, definiamo un sistema di riduzione S, cioè un insieme di endomorfismi di un certo tipo. Ogni monomio in generale viene trasformato dalle riduzioni in S; un monomio m è irriducibile se, al contrario, resta invariante, cioè $s(m) = s\ \forall s \in S$. Se supponiamo che le riduzioni in S possano trasformare ciascun monomio in un espressione irriducibile, sorge spontanea la domanda se tale forma irriducibile sia unica.
Il Diamond Lemma fornisce una serie di condizioni equivalenti affinchè ciò accada ed inoltre, allo stesso tempo, ci dà un insieme di rappresentanti in $k\langle X\rangle$ per il sottomodulo degli elementi irriducibili. Il risultato di Bergman è molto generale ed ha applicazioni nei più diversi ambiti della matematica: l'autore stesso fornisce alcuni esempi in [2] tra cui una dimostrazione del teorema di Poincaré-Birkhoff-Witt di cui ci occupiamo più in dettaglio nel capitolo 3.

Il capitolo 3 introduce uno dei risultati più importanti dell'intera teoria

delle algebre di Lie: il Teorema di Poincaré-Birkhoff-Witt (PBW). Questo teorema fu originariamente enunciato e solo parzialmente dimostrato da Poincaré alla fine dell'Ottocento, ma solo nel 1937 Birkhoff e Witt indipendentemente ne diedero una dimostrazione completa.
Forniamo dapprima la trattazione del Teorema PBW per le algebre di Lie definite sugli anelli, per la dimostrazione del quale usiamo il Diamond Lemma del capitolo precedente, e quindi spostiamo l'attenzione sul Teorema PBW per algebre di Lie definite sui campi, un caso più comune, e presentiamo una dimostrazione diversa, che non fa uso del Diamond Lemma, tratta principalmente da [7]. Il lettore attento troverà comunque alcuni punti in comune fra le due prove.

E' importante ricordare una delle conseguenze più importanti del Teorema PBW: ogni algebra di Lie si immerge iniettivamente nella sua algebra inviluppante universale, che essendo associativa risulta generalmente più facile da studiare pur essendo, nella grande maggioranza dei casi, infinitodimensionale.
L'algebra inviluppante universale riveste un'importanza fondamentale nella teoria della rappresentazione delle algebre di Lie. Non riusciremo tuttavia a discutere alcun dettaglio.

Il capitolo 4 è dedicato alle applicazioni, in particolare a introdurre degli esempi in cui importanti algebre di Lie sono necessariamente definite sugli anelli: in questo modo giustifichiamo la necessità di ampliare la generalità in cui dimostriamo il Teorema PBW, cioè la necessità di studiare il Teorema per le algebre di Lie definite su anelli e non su campi, caso più comune. L'esempio fondamentale che introduciamo è quello dell' algebra di Lie associata ad un gruppo algebrico, definita funtorialmente.

Indice

Capitolo 1

Algebre di Lie

Sia k anello commutativo unitario con 2, 3 invertibili.
Tutti i moduli, i morfismi e le algebre, ove non diversamente specificato, sono da intendersi rispettivamente come k-moduli, k-morfismi e k-algebre.

1.1 Algebra di Lie

Definizione 1.1. Si dice *algebra di Lie* un modulo $\mathfrak{g}$ sul quale è definita una mappa bilineare

$$\begin{array}{ccc} \mathfrak{g} \times \mathfrak{g} & \longrightarrow & \mathfrak{g} \\ (x, y) & \mapsto & [x, y] \end{array}$$

tale che

1. $[x, y] = -[y, x] \quad \forall x, y \in \mathfrak{g}$;
2. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \quad \forall x, y, z \in \mathfrak{g}$.

La proprietà (1) è nota come *antisimmetria*, mentre la proprietà (2) si dice *identità di Jacobi*. $[\ ,\ ]$ è anche detta *bracket*.

Poichè abbiamo ipotizzato 2 invertibile, la condizione (1) è equivalente a $[x, x] = 0\ \forall x \in \mathfrak{g}$, infatti se vale (1) scegliendo $x = y$ si ha $[x, x] = -[x, x] \Rightarrow [x, x] = 0$; viceversa $0 = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y] = 0 + [x, y] + [y, x] + 0$ da cui (1).

Vediamo ora qualche esempio.

Esempio 1.2. Per ogni modulo $\mathfrak{g}$ se definiamo $[x, y] = 0, \forall x, y \in \mathfrak{g}$ otteniamo un'algebra di Lie con (1) e (2) della definizione verificate in modo banale. Tale $\mathfrak{g}$ si dice *commutativa*.

Esempio 1.3. Sia $\mathfrak{g}$ un'algebra associativa, cioè un modulo con una struttura di anello compatibile. Se definiamo $[x, y] = xy - yx$, $\forall x, y \in \mathfrak{g}$ otteniamo un'algebra di Lie.
Andiamo a verificare la definizione di algebra di Lie. La bracket è bilineare, infatti dati $x, y, z \in \mathfrak{g}$, $a, b \in k$ si ha

$$[ax + by, z] = (ax + by)z - z(ax + by) = axz + byz - azx - bzy =$$

$$= (axz - azx) + (byz - bzy) = a[x, z] + b[y, z]$$

e analogamente per il secondo argomento della bracket si trova:

$$[x, ay + bz] = a[x, y] + b[x, z].$$

L'antisimmetria è ovvia; per l'identità di Jacobi abbiamo:

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = [x, yz - zy] + [y, zx - xz] + [z, xy - yx] =$$

$$= [x, yz] - [x, zy] + [y, zx] - [y, xz] + [z, xy] - [z, yx] =$$

$$= xyz - yzx - xzy + zyx + yzx - zxy - yxz + xzy + zxy - xyz - zyx + yxz = 0.$$

Siamo ora in grado di dare esempi più concreti di algebre di Lie, a partire dall'algebra degli endomorfismi di un modulo.

Esempio 1.4. Dato un modulo M, l'insieme $\mathrm{End}(M)$ degli endomorfismi di M è in modo naturale un modulo. Se definiamo come prodotto la composizione di endomorfismi, $\mathrm{End}(M)$ risulta essere un'algebra associativa, e dunque un'algebra di Lie (vedi esempio 1.3).
Se M è un modulo libero di dimensione finita n, fissata una sua base l'insieme $\mathrm{End}(M)$ si identifica con $M_n(k)$, anello delle matrici quadrate di ordine n a coefficienti in k. $M_n(k)$ è algebra associativa e definendo

$$[X, Y] := XY - YX \qquad \forall X, Y \in M_n(k)$$

si ha un'algebra di Lie dall'esempio 1.3. Come algebra di Lie, questa è spesso denotata in letteratura $\mathfrak{gl}_n(k)$, in quanto algebra di Lie del gruppo generale lineare $GL_n(k)$.
Con la medesima definizione delle brackets, risultano essere algebre di Lie i seguenti sottomoduli di $M_n(k)$:

- $\mathfrak{sl}_n(k)$ matrici con traccia nulla (algebra di Lie speciale lineare);
- $\mathfrak{so}_n(k)$ matrici antisimmetriche (algebra di Lie ortogonale);
- matrici triangolari superiori;

- matrici triangolari inferiori.

In ognuno dei casi sopracitati la bracket ristretta al sottomodulo è ben definita, mentre le proprietà di bilinearità, antisimmetria e di Jacobi sono ereditate da $\mathfrak{gl}_n(k)$.

Definizione 1.5. Date $\mathfrak{g}, \mathfrak{g}'$ algebre di Lie, si dice *morfismo di Lie* una mappa lineare $\pi : \mathfrak{g} \longrightarrow \mathfrak{g}'$ che conserva le brackets, ossia

$$\pi([X,Y]) = [\pi(X), \pi(Y)] \quad \forall X, Y \in \mathfrak{g}.$$

Le rappresentazioni costituiscono l'esempio più importante di morfismi tra algebre di Lie.

Definizione 1.6. Sia $\mathfrak{g}$ una'algebra di Lie e M un modulo libero.
Si dice *rappresentazione di* $\mathfrak{g}$ *in* M un morfismo di algebre di Lie:

$$\rho : \mathfrak{g} \longrightarrow \mathrm{End}(M).$$

In altre parole una rappresentazione è una mappa $\rho : \mathfrak{g} \longrightarrow \mathrm{End}(M)$ tale che

1. ρ sia lineare;
2. $\rho([X,Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X) \quad \forall X, Y \in \mathfrak{g}$.

In tal caso si dice anche che $\mathfrak{g}$ *agisce su* M o che M *è un* $\mathfrak{g}$*-modulo.*
Se ρ è iniettiva, la rappresentazione si dice *fedele.*

Se M è modulo libero finito dimensionale e ρ è una rappresentazione fedele di $\mathfrak{g}$ in M, ρ ci permette di identificare $\mathfrak{g}$ con una sottoalgebra di $\mathrm{End}(M)$.

Vediamo degli esempi.

Esempio 1.7. Sia $\mathfrak{g} = \mathrm{span}_k\{H, X, Y\}$ con la bracket così definita sui generatori:

$$[X,Y] := H \quad [H,X] := 2X \quad [H,Y] := -2Y.$$

Tale bracket definisce una struttura di algebra di Lie su $\mathfrak{g}$, in quanto bilinearità e antisimmetria vengono dalla definizione, mentre l'identita di Jacobi risulta verificata per ogni possibile caso.

1. Consideriamo ora la mappa $\rho : \mathfrak{g} \longrightarrow \mathrm{End}(k^2)$ così definita sui generatori:

$$H \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad X \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad Y \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Verificando la proprietà (2) della definizione 1.6 per le brackets dei generatori, si ricava subito che questa è una rappresentazione di $\mathfrak{g}$ in $\mathfrak{sl}_2(k)$.
Poichè la rappresentazione ρ è iniettiva e suriettiva, possiamo identificare i generatori di $\mathfrak{sl}_2(k)$ con H, X, Y. Dunque $\mathfrak{g}$ è isomorfa a $\mathfrak{gl}_n(k)$ algebra di Lie speciale lineare.

2. Fissato $n \in \mathbb{N}$ siano $\mathfrak{g} = \mathfrak{sl}_2(k)$ e $V = \text{span}_k\{x^r y^{n-r}\}_{0 \leq r \leq n}$, lo spazio vettoriale dei polinomi omogenei di grado n sul campo k.
Definiamo la seguente rappresentazione di $\mathfrak{g}$ in V:
$\rho : \mathfrak{sl}_2(k) \longrightarrow \text{End}(V)$:

$$\begin{aligned} X &\mapsto \left(p \mapsto y\tfrac{\partial p}{\partial x}\right) \\ Y &\mapsto \left(p \mapsto x\tfrac{\partial p}{\partial y}\right) \\ H &\mapsto \left(p \mapsto y\tfrac{\partial}{\partial x}\left(x\tfrac{\partial p}{\partial y}\right) - x\tfrac{\partial}{\partial y}\left(y\tfrac{\partial p}{\partial x}\right)\right). \end{aligned}$$

Nuovamente, questa risulta rappresentazione una volta verificate le condizioni sulle brackets: vediamo come agiscono le immagini di X, Y, H sulla base di V.

$$\begin{aligned} \rho(X)(y^n) &= 0 \qquad \rho(X)(x^r y^{n-r}) = r x^{r-1} y^{n-r+1} \\ \rho(Y)(x^n) &= 0 \qquad \rho(Y)(x^r y^{n-r}) = (n-r) x^{r+1} y^{n-r-1} \\ &\qquad\quad \rho(H)(x^r y^{n-r}) = (n-2r) x^r y^{n-r}. \end{aligned}$$

A meno di uno scalare moltiplicativo, possiamo riassumere le infor-

mazioni come segue:

$$\begin{array}{cc} \text{autovalori di } H & \text{base di } V \\ \\ n & y^n \\ & \rho(X)\uparrow \quad \downarrow\rho(Y) \\ n-2 & xy^{n-1} \\ & \rho(X)\uparrow \quad \downarrow\rho(Y) \\ \vdots & \vdots \\ & \rho(X)\uparrow \quad \downarrow\rho(Y) \\ -n+2 & x^{n-1}y \\ & \rho(X)\uparrow \quad \downarrow\rho(Y) \\ -n & x^n. \end{array}$$

Si ha quella che viene detta *rappresentazione a scala* o *ladder*: n è detto *peso più alto* perchè è il massimo autovalore di H ed analogamente y^n si dice *vettore di peso più alto*.
E' possibile dimostrare che tutte le rappresentazioni finito dimensionali di $\mathfrak{sl}_2(k)$ sono isomorfe ad una somma diretta di rappresentazioni a scala (vedi [7], capitolo IV).

Esempio 1.8. Sia $\mathfrak{g}$ un'algebra di Lie. Definiamo la seguente mappa lineare

$$\begin{array}{rccc} \text{ad}: & \mathfrak{g} & \longrightarrow & \text{End}(\mathfrak{g}) \\ & X & \mapsto & \{Y \mapsto [X,Y]\} \end{array}$$

$Y \mapsto [X,Y]$ è un'endomorfismo di $\mathfrak{g}$, e la mappa ad : $X \mapsto \text{ad}X$ è detta *rappresentazione aggiunta* di $\mathfrak{g}$ in $\mathfrak{g}$.
Il nucleo della rappresentazione aggiunta, cioè l'insieme degli $X \in \mathfrak{g}$ tali che $[X,Y] = 0$, $\forall Y \in \mathfrak{g}$ è detto *centro* di $\mathfrak{g}$.

Definizione 1.9. Sia $\rho : \mathfrak{g} \longrightarrow \text{End}(M)$ una rappresentazione. Dato N sottomodulo di M, esso si dice *sottorappresentazione* o $\mathfrak{g}$-*sottomodulo* se

$$\rho(x)(N) \subseteq N \quad \forall x \in \mathfrak{g}.$$

Definizione 1.10. Una rappresentazione $\rho : \mathfrak{g} \longrightarrow \text{End}(M)$ si dice *irriducibile* quando le sole sottorappresentazioni che ammette sono quelle ovvie, cioè 0 e M stesso.
ρ si dice *completamente riducibile* se per ogni $N \subset M$ sottorappresentazione esiste $N' \subset M$ sottorappresentazione tale che $M = N \bigoplus N'$.

Se k è campo, e dunque M è spazio vettoriale, per N sottorappresentazione di M esiste sempre N' (non unico) sottospazio vettoriale tale che $M = N \bigoplus N'$. Tuttavia nella definizione 1.10 relativa alla completa riducibilità richiediamo che tale N' sia a sua volta una sottorappresentazione. Questa richiesta fa si chè, come vedremo nell'esempio seguente, non sempre tale N' esista.

Esempio 1.11. Consideriamo l'esempio 1.7 con $k = \mathbb{Z}_5$ ed $n = 5$ e identifichiamo gli endomorfismi $\rho(X)$, $\rho(Y)$, $\rho(H)$ con le rispettive matrici in $\mathrm{End}(V)$. Si ha

$$\begin{aligned} X : x^5 &\mapsto 5x^4y = 0 \\ Y : x^5 &\mapsto 0 \\ H : x^5 &\mapsto -5x^5 = 0 \end{aligned}$$

Dunque evidentemente $W = \mathrm{span}_k\{x^5\}$ è sottorappresentazione. Proviamo ora che non esiste una sottorappresentazione W' con $V = W \bigoplus W'$. Se W' esistesse, il suo generico elemento w sarebbe tale che

$$w = \sum_{i=1}^{5} a_i x^{5-i} y^i$$

e quindi si avrebbe

$$Y(w) = \sum_{i=1}^{5} a_i i x^{6-i} y^{i-1}.$$

In tale immagine è presente anche un termine x^5 e quindi $Y(w) \notin W'$, perciò la rappresentazione non è completamente riducibile.
In questo esempio il fatto che *char* $k \neq 0$ ha giocato un ruolo determinante e infatti è possibile dimostrare che nel caso di $\mathfrak{sl}_2(k)$ con *char* $k = 0$ e k algebricamente chiuso, tutte le rappresentazioni finito-dimensionali sono completamente riducibili (vedi [7], capitolo IV).

1.2 Algebra Inviluppante Universale

L'algebra inviluppante universale ha un ruolo fondamentale nella teoria delle algebre di Lie e delle loro rappresentazioni.

Prima di darne la definizione ricordiamo come è definito il prodotto tensoriale di due moduli.

Definizione 1.12. Siano M, N due moduli. Definiamo dapprima l'algebra libera

$$k^{M\times N} = span_k\{m \otimes n \text{ tali che } m \in M, n \in N\}$$

che è data dalle somme formali degli elementi $m \otimes n$ a coefficienti in k. In generale si tratta di un modulo libero infinito-dimensionale.
Consideriamo ora il sottomodulo generato dalle relazioni che seguono:

$$D = \begin{pmatrix} k(m\otimes n) & - & km\otimes n \\ k(m\otimes n) & - & m\otimes kn \\ (m_1+m_2)\otimes n & - & m_1\otimes n + m_2\otimes n \\ m\otimes(n_1+n_2) & - & m\otimes n_1 + m\otimes n_2 \end{pmatrix}.$$

Il *prodotto tensoriale di* M *e* N è il sottomodulo definito come

$$M \otimes N := k^{M\otimes N}/D.$$

Ed è inoltre definita la mappa bilineare

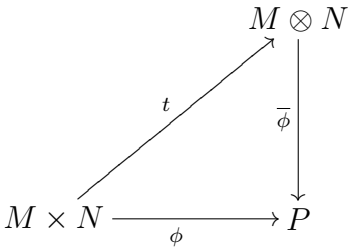
$$t : M \times N \longrightarrow M \otimes N \quad (m, n) \mapsto m \otimes n.$$

Proposizione 1.13. *Siano* V, W *due spazi vettoriali di basi rispettivamente* $\{v_i\}_{1\le i\le m}$ *e* $\{w_j\}_{1\le j\le n}$. *Allora* $V \otimes W$ *ha base* $\{v_i \otimes w_j\}$ *e risulta dunque spazio vettoriale di dimensione* mn.

Dimostrazione. Questa proposizione è un risultato immediato del corollario IV.5.12 di [4]. □

E' possibile dare una definizione alternativa del prodotto tensoriale tramite proprietà universale.

Teorema 1.14 (Proprietà universale del Prodotto Tensoriale)**.** *Siano* M, N, P *tre moduli. Per ogni mappa bilineare* $\phi : M \times N \longrightarrow P$ *esiste ed è unico un morfismo* $\overline{\phi} : M \otimes N \longrightarrow P$ *che rende commutativo il diagramma che segue*



Dimostrazione. La dimostrazione di questo teorema si può trovare in [1], alle pagine 24-25. □

Definizione 1.15. Sia M un modulo, poniamo dapprima per ogni $n \geq 0$

$$M^{\otimes n} = \underbrace{M \otimes \ldots \otimes M}_{n}.$$

Definiamo quindi il modulo dei tensori su M nel modo seguente:

$$T(M) = \bigoplus_{n=0}^{\infty} M^{\otimes n}.$$

Su di esso è possibile definire un prodotto come segue:

$$\Big(\sum a_{i_1} \ldots a_{i_r}\, m_{i_1} \otimes \ldots \otimes m_{i_r}, \sum b_{j_1} \ldots b_{j_s}\, m_{j_1} \otimes \ldots \otimes m_{j_s}\Big) \mapsto$$

$$\sum a_{i_1} \ldots a_{i_r} b_{j_1} \ldots b_{j_s}\, m_{i_1} \otimes \ldots \otimes m_{i_r} \otimes m_{j_1} \otimes \ldots \otimes m_{j_s}.$$

Con questa struttura $T(M)$ risulta dunque un'algebra associativa detta *algebra tensoriale su M*.

Osservazione 1.16. *Siano K campo e V un K-spazio vettoriale non banale. $T(V)$ risulta essere K-spazio vettoriale infinito dimensionale.*

L'algebra tensoriale ha due quozienti particolarmente interessanti: l'algebra polinomiale e l'algebra esterna.

Definizione 1.17. Sia M un modulo, e C l'ideale di $T(M)$ così generato:

$$C = \langle m \otimes n - n \otimes m | m, n \in M \rangle.$$

Definiamo *algebra simmetrica di M*:

$$\mathrm{Sym}(M) = T(M)/C.$$

Teorema 1.18. *Sia M un modulo libero di base $\{x_i\}_{i \in I}$. Si ha l'isomorfismo canonico*

$$\mathrm{Sym}(M) \cong k[(x_i)_{i \in I}],$$

dove $k[(x_i)_{i \in I}]$ è l'algebra dei polinomi su M.
Per questa ragione $\mathrm{Sym}(M)$ *è anche detta* algebra polinomiale.

Dimostrazione. Questo risultato è provato in dettaglio nella sezione I.III.3 di [6]. □

Definizione 1.19. Sia M un modulo, e C l'ideale di $T(M)$ così generato:

$$C = \langle m \otimes n + n \otimes m | m, n \in M \rangle.$$

Definiamo *algebra esterna di* M:

$$\bigwedge(M) = T(M)/C.$$

Definizione 1.20. Sia A un'algebra associativa. A si dice *graduata* se per ogni intero $n \geq 0$ esiste un sottomodulo A_n di A tale che

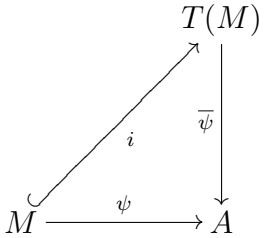
- $1 \in A_0$ e $A = \bigoplus_{n \geq 0} A_n$;
- $A_m A_n \subseteq A_{m+n}$ per ogni $m, n \geq 0$.

Gli elementi di A_n si chiamano *omogenei di grado n.*

Proposizione 1.21. *L'algebra tensoriale $T(M)$ è algebra graduata e tale graduazione è ereditata sia dall'algebra polinomiale che dall'algebra esterna.*

Dimostrazione. Per definizione, si ha $T(M) = \bigoplus_{n \geq 0} M^{\otimes n}$. Ponendo $A_n = M^{\otimes n}, \forall n \in \mathbb{N}$, si ha evidentemente A_n sottomodulo di $T(M)$, $A_0 = k$ e $A_i A_j \subseteq A_{i+j}\, \forall i, j \in \mathbb{N}$ e quindi l'algebra tensoriale verifica la definizione di algebra graduata.
Visto che l'ideale $C = \langle m \otimes n - n \otimes m | m, n \in M \rangle$ rispetto al quale quozientiamo l'algebra tensoriale quando andiamo a creare l'algebra polinomiale è generato da elementi omogenei di grado 2, la gradazione passa al quoziente; un discorso analogo vale per l'algebra esterna. Una prova dettagliata del fatto in questione si può trovare su [7], alle pagine 164-166. □

Teorema 1.22 (Proprietà universale dell'algebra tensoriale)**.** *Sia A un'algebra associativa e $\psi : M \longrightarrow A$ un morfismo di moduli.*
Allora esiste ed è unico il morfismo di algebre associative $\overline{\psi} : T(M) \longrightarrow A$ che rende commutativo il diagramma che segue:



Conseguentemente vi è una corrispondenza biunivoca

$$\mathrm{Hom}_{\mathrm{Mod}}(M, A) \cong \mathrm{Hom}_{\mathrm{Alg}}(T(M), A)$$

fra i morfismi di moduli da M ad A ed i morfismi di algebre associative da $T(M)$ ad A.

Dimostrazione. Data ψ, definiamo $\overline{\psi}$ come morfismo lineare nel seguente modo:

$$v_{i_1} \otimes \ldots \otimes v_{i_r} \quad \overset{\overline{\psi}}{\longmapsto} \quad \psi(v_{i_1}) \ldots \psi(v_{i_r}).$$

$\overline{\psi}$ è morfismo di algebre, difatti è morfismo di moduli perchè lo è ψ ed è morfismo di anelli perchè $\overline{\psi}(v_{i_1} \otimes \ldots \otimes v_{i_r}) = \psi(v_{i_1}) \ldots \psi(v_{i_r}) = \overline{\psi}(v_{i_1}) \ldots \overline{\psi}(v_{i_r})$. Inoltre $\overline{\psi} \circ i = \psi$ e questo prova la tesi. □

$T(M)$ è anche detta *algebra polinomiale libera* perchè non essendovi alcuna relazione fra i suoi elementi, costruiti a partire da quelli di M, essa è l'anello più generale possibile che possiamo associare al modulo M.

Definizione 1.23. Sia $\mathfrak{g}$ un'algebra di Lie. Consideriamo l'ideale bilatero di $T(\mathfrak{g})$

$$I = \langle x \otimes y - y \otimes x - [x, y] | x, y \in \mathfrak{g} \rangle.$$

Definiamo quindi il quoziente:

$$U(\mathfrak{g}) = T(\mathfrak{g})/I$$

e il morfismo di moduli

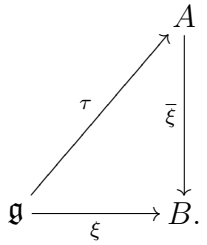
$$\tau : \mathfrak{g} \longrightarrow U(\mathfrak{g}), \quad \tau = p \circ i,$$

dove $i : \mathfrak{g} \hookrightarrow T(\mathfrak{g})$ è l'immersione e $p : T(\mathfrak{g}) \longrightarrow U(\mathfrak{g})$ è la proiezione sul quoziente. La coppia $(U(\mathfrak{g}), \tau)$ è detta *algebra inviluppante universale* (A.I.U.).

E' possibile dare una definizione alternativa ed equivalente tramite proprietà universale.

Teorema 1.24 (Proprietà Universale dell'Algebra Inviluppante)**.** *Sia $\mathfrak{g}$ un'algebra di Lie e sia (A, τ) una coppia con A algebra associativa, $\tau : \mathfrak{g} \longrightarrow A$ morfismo di moduli tale che*

1. *$\tau(\mathfrak{g})$ genera A;*
2. *$\forall x, y \in \mathfrak{g}, \quad \tau([x, y]) = \tau(x)\tau(y) - \tau(y)\tau(x)$;*
3. *Per ogni algebra di Lie associativa B, per ogni morfismo di algebre di Lie $\xi : \mathfrak{g} \longrightarrow B$ esiste ed è unico un morfismo di algebre associative $\overline{\xi} : A \longrightarrow B$ che rende commutativo il seguente diagramma:*



Allora

$$A \cong U(\mathfrak{g}), \quad \tau \cong p \circ i$$

dove $i : \mathfrak{g} \hookrightarrow T(\mathfrak{g})$ *è l'immersione e* $p : T(\mathfrak{g}) \longrightarrow U(\mathfrak{g})$ *è la proiezione sul quoziente.*

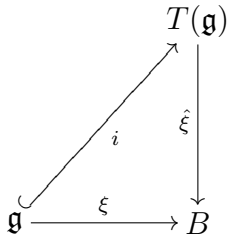
L'A.I.U. è allora l'unica algebra associativa (a meno di isomorfismo, anch'esso unico) ad avere la proprietà universale descritta.

Dimostrazione. Siano $U(\mathfrak{g})$ e τ come nella definizione 1.23, vediamo ora che esse verificano le condizioni (1), (2), (3).

1. $i(\mathfrak{g})$ genera l'algebra tensoriale $T(\mathfrak{g})$ e dunque anche il suo quoziente $U(\mathfrak{g})$.

2. $$\tau([x,y]) - \tau(x)\tau(y) + \tau(y)\tau(x) =$$
$$= p(i([x,y])) - p(i(x))p(i(y)) + p(i(y))p(i(x)) =$$
$$= p(i([x,y]) - i(x)i(y) + i(y)i(x)) = p([x,y] - x \otimes y + y \otimes x).$$
Ora $[x,y] - x \otimes y + y \otimes x \in I$ dove I è il sottomodulo per cui quozientiamo nella definizione di $U(\mathfrak{g})$, dunque la proiezione si annulla su di esso e vale perciò (2).

3. Dato $\xi : \mathfrak{g} \longrightarrow B$ come da ipotesi, per la proprietà universale del prodotto tensoriale esiste unico $\hat{\xi} : T(\mathfrak{g}) \longrightarrow B$ morfismo di algebre associative che rende commutativo il seguente diagramma:



Dato che ξ è morfismo di algebre di Lie, per ogni $x, y \in \mathfrak{g}$ si ha

$$\xi([x,y]) = [\xi(x), \xi(y)] = \xi(x)\xi(y) - \xi(y)\xi(x).$$

Per la commutatività del diagramma, $\xi([x,y]) = \hat{\xi}(i([x,y])) = \hat{\xi}([x,y])$ ed analogamente $\xi(x)\xi(y) - \xi(y)\xi(x) = \hat{\xi}(i(x))\hat{\xi}(i(y)) - \hat{\xi}(i(y))\hat{\xi}(i(x)) = \hat{\xi}(x)\hat{\xi}(y) - \hat{\xi}(y)\hat{\xi}(x)$. Dunque vale

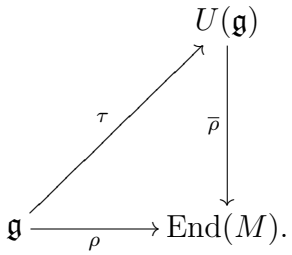
$$\hat{\xi}([x,y]) = \hat{\xi}(x)\hat{\xi}(y) - \hat{\xi}(y)\hat{\xi}(x) = \hat{\xi}(x \otimes y) - \hat{\xi}(y \otimes x),$$

cioè $\hat{\xi}$ si annulla sui generatori di $I = \langle x \otimes y - y \otimes x - [x,y] \rangle$ e conseguentemente si annulla su tutto I inducendo perciò un morfismo $\overline{\xi} : T(\mathfrak{g})/I \longrightarrow B$ che è evidentemente quello cercato.
Sia ora (A', τ') una coppia che soddisfa le condizioni (1), (2), (3).
Ponendo dapprima $B = U(\mathfrak{g})$ e poi $B = A'$ si ottengono dalla proprietà universale due morfismi $\overline{\tau'} : A' \longrightarrow U(\mathfrak{g})$ e $\overline{\tau} : U(\mathfrak{g}) \longrightarrow A'$ che sono uno inverso dell'altro e questo prova l'unicità a meno di un unico isomorfismo.

□

Teorema 1.25. *Le rappresentazioni di* $\mathfrak{g}$ *sono in corrispondenza biunivoca con le rappresentazioni di* $U(\mathfrak{g})$*, ossia* $\rho : \mathfrak{g} \longrightarrow \mathrm{End}(M)$ *è rappresentazione di* $\mathfrak{g}$ *in* M *se e solo se* $\overline{\rho} : U(\mathfrak{g}) \longrightarrow \mathrm{End}(M)$ *è rappresentazione di* $U(\mathfrak{g})$ *in* M*.*

Questa è conseguenza immediata della proprietà universale, come è evidente dal diagramma



Capitolo 2

Diamond Lemma

Il Diamond Lemma che andremo ad introdurre in questo capitolo e che useremo per dimostrare il Teorema di Poincaré-Birkhoff-Witt è dovuto a George Bergman (vedi [2]).
In questo risultato si vogliono definire delle condizioni necessarie affinchè vi sia una forma canonica degli elementi di un'algebra associativa di cui è data una presentazione. In altre parole si tratta di verificare se, riducendo un elemento iniziale in modi diversi, si giunge con un numero finito di passaggi ad uno stesso elemento finale e sotto quali condizioni ciò accade.

Sia k anello commutativo unitario.

Tutti i moduli e i morfismi, ove non diversamente specificato, sono da intendersi rispettivamente come k-moduli e k-morfismi.

2.1 Riduzioni

Definizione 2.1. Sia X un insieme non vuoto. Denotiamo con $\langle X\rangle$ il monoide libero su X con l'operazione di giustapposizione e con $k\langle X\rangle$ la k-algebra associativa libera su X. Chiameremo *monomi* gli elementi di $\langle X\rangle$ e *polinomi* gli elementi di $k\langle X\rangle$.

Queste sono le strutture algebriche su cui andremo a lavorare, introduciamo ora alcuni strumenti e alcune notazioni che useremo in seguito.

Definizione 2.2. Sia S un insieme di coppie della forma:

$$\sigma = (W_\sigma, f_\sigma) \text{ con } W_\sigma \in \langle X\rangle,\ f_\sigma \in k\langle X\rangle.$$

S è detto *sistema di riduzione.*
Per ogni $\sigma \in S$ e $A, B \in \langle X\rangle$ denotiamo con $r_{A\sigma B}$ l'endomorfismo di $k\langle X\rangle$

che fissa tutti gli elementi di $\langle X\rangle$ e manda

$$r_{A\sigma B} : AW_\sigma B \mapsto Af_\sigma B.$$

Le mappe $r_{A\sigma B}$ sono dette *riduzioni*.

Esempio 2.3. Sia $A = \{X, Y, H\}$. Tenendo anche in considerazione l'esempio 1.7 definiamo il sistema di riduzione su A:

$$S = \{\alpha = (XY, YX + H), \beta = (HX, XH + 2X), \gamma = (HY, YH - 2Y)\}.$$

Vediamo ora come agiscono le riduzioni sul monomio HXY per il quale possiamo evidentemente cominciare a ridurre in due modi diversi:

$$HXY \xrightarrow{r_{1\beta Y}} (XH+2X)Y = XHY+2XY \xrightarrow{r_{X\gamma 1}} X(YH-2Y)+2XY = XYH$$

$$\xrightarrow{r_{1\alpha H}} (YX + H)H = YXH + H^2.$$

$$HXY \xrightarrow{r_{H\alpha 1}} H(YX + H) = HYX + H^2 \xrightarrow{r_{1\gamma X}} (YH - 2Y)X + H^2 = YHX-$$

$$-2YX + H^2 \xrightarrow{r_{Y\beta 1}} Y(XH + 2X) - 2YX + H^2 = YXH + H^2.$$

Si arriva dunque allo stesso polinomio $YXH + H^2$ pur compiendo scelte diverse di riduzioni. Non è sempre così, ma in questo caso particolare non sorgono problemi in quanto stiamo in effetti lavorando su $U(\mathfrak{sl}_2)$ e il teorema di PBW che studieremo nel prossimo capitolo ci assicura che il sistema di riduzione associato ad un'algebra di Lie ha sempre questa proprietà. In altre parole, partendo da uno stesso elemento (in questo caso HXY) e fissato un ordine (in questo caso $Y < X < H$) è sempre possibile ridurre, anche in modi diversi, l'elemento dato in modo da ottenere alla fine un polinomio in cui i monomi sono ordinati nell'ordine stabilito. Tale polinomio è inoltre unico.

Definizione 2.4. Diciamo che una riduzione $r_{A\sigma B}$ *agisce banalmente* su un elemento $a \in k\langle X\rangle$ se il coefficiente di $AW_\sigma B$ in a è zero, ossia se $r_{A\sigma B}$ fissa a.
Un elemento $a \in k\langle X\rangle$ è detto *irriducibile secondo il sistema di riduzione* S se ogni riduzione di S agisce banalmente su a, ossia se in a non compare nessuno dei monomi $AW_\sigma B$ al variare di A, B in $\langle X\rangle$ e di σ in S.
Denotiamo con $k\langle X\rangle_{irr}$ l'insieme degli elementi irriducibili secondo S.

Osservazione 2.5. *L'insieme $k\langle X\rangle_{irr}$ è sottomodulo di $k\langle X\rangle$.*

Dimostrazione. Siano $a, b \in k\langle X\rangle_{irr}$ e $\alpha \in k$. Per definizione, in a e in b non compare nessuno dei monomi $AW_\sigma B$, dunque questi non compaiono nè nella somma $a + b$ nè in αa, perciò $\alpha a, a + b \in k\langle X\rangle_{irr}$. □

Definizione 2.6. Sia $a \in k\langle X\rangle$, una successione finita di riduzioni $r_1, \ldots, r_n$ dove $r_i = r_{A_i\sigma_i B_i}$ si dice *finale su* a se $r_n \ldots r_1(a) \in k\langle X\rangle_{irr}$.
Un'elemento $a \in k\langle X\rangle$ è detto *a riduzione finita* se per ogni successione infinita $(r_n)_{n\in\mathbb{N}}$ di riduzioni, r_i agisce banalmente su $r_{i-1} \ldots r_1(a)$ per ogni i sufficientemente grande.

Proposizione 2.7. *Sia $a \in k\langle X\rangle$ un elemento a riduzione finita. Ogni successione massimale di riduzioni (r_i), tale che ogni r_i agisce non banalmente su $r_{i-1} \ldots r_1(a)$, è finita e dunque è una successione finale su a. Inoltre, gli elementi a riduzione finita di $k\langle X\rangle$ formano un sottomodulo.*

Dimostrazione. L'esistenza di una successione *infinita* di riduzioni $(r_i)_{i\in\mathbb{N}}$ dove r_i agisce non banalmente su $r_{i-1} \ldots r_1(a)$ per ogni i va contro la definizione di elemento a riduzione finita, quindi ogni successione massimale è finita.
Le successioni massimali finite risultano allora essere successioni finali, perchè se $r_i \ldots r_1(a) \notin k\langle X\rangle_{irr}$ è possibile scegliere una riduzione r_{i+1} che agisce non banalmente su $r_i \ldots r_1(a)$ ma questo è assurdo per la massimalità della successione.
Ora verifichiamo che gli elementi a riduzione finita formano un sottomodulo di $k\langle X\rangle$. Ricordiamo che le riduzioni sono lineari, siano dunque $a, b \in k\langle X\rangle$ a riduzione finita. Per ogni successione infinita $(r_n)_{n\in\mathbb{N}}$ di riduzioni, r_{i_a} agisce banalmente su $r_{i_a-1} \ldots r_1(a)$ e r_{i_b} agisce banalmente su $r_{i_b-1} \ldots r_1(a)$ per ogni i_a, i_b sufficientemente grandi. Scelto perciò $i = \max\{i_a, i_b\}$, si ha che r_i agisce banalmente su $r_{i-1} \ldots r_1(a) + r_{i-1} \ldots r_1(b) = r_{i-1} \ldots r_1(a+b)$ e quindi $a + b$ è a riduzione finita.
Dato $\alpha \in k$, se la successione infinita di riduzioni (r_i) agisce banalmente su $r_{i-1} \ldots r_1(a)$ per i sufficientemente grande, allora essa agisce banalmente su $\alpha r_{i-1} \ldots r_1(a) = r_{i-1} \ldots r_1(\alpha a)$ per i sufficientemente grande, quindi αa è a riduzione finita e questo prova la seconda parte della proposizione. □

Definizione 2.8. Un'elemento $a \in k\langle X\rangle$ si dice *a riduzione unica* se è a riduzione finita e le sue immagini tramite ogni successione finale di riduzioni coincidono. Questa immagine comune, detta *forma ridotta di* a, è indicata con $r_s(a)$ e vale evidentemente $r_s(a) \in k\langle X\rangle_{irr}$.

Per il momento, r_s è solo la notazione con cui indichiamo la forma ridotta; nel lemma 2.14(1) vedremo che restringendo opportunamente dominio e codominio r_s risulta essere morfismo.

Osservazione 2.9. *Il fatto di essere a riduzione finita non implica affatto di essere a riduzione unica.*

Esempio 2.10. Modifichiamo il sistema dell'esempio 2.3. Sia $A = \{X, Y, H\}$, definiamo il sistema di riduzione su A:

$$S' = \{\alpha' = (XY, YX + H + X), \beta = (HX, XH + 2X), \gamma = (HY, YH - 2Y)\}.$$

Si ottiene allora

$$HXY \xrightarrow{r_{1\beta Y}} (XH+2X)Y = XHY + 2XY \xrightarrow{r_{X\gamma 1}} X(YH-2Y)+2XY = XYH$$

$$\xrightarrow{r_{1\alpha' H}} (YX + H + X)H = YXH + H^2 + XH.$$

$$HXY \xrightarrow{r_{H\alpha' 1}} H(YX + H + X) = HYX + H^2 + HX \xrightarrow{r_{1\gamma X}} (YH - 2Y)X + H^2 +$$

$$+HX = YHX - 2YX + H^2 + HX \xrightarrow{r_{Y\beta 1}} (Y + 1)(XH + 2X) - 2YX + H^2 =$$

$$= YXH + H^2 + XH + 2X.$$

perciò evidentemente HXY è a riduzione finita ma non a riduzione unica.

Definizione 2.11. Un insieme parzialmente ordinato X soddisfa la *condizione della catena discendente* se ogni catena discendente

$$x_1 \geq x_2 \geq \ldots \geq x_n \geq \ldots$$

è stazionaria, ossia esiste $n \in \mathbb{N}$ tale che $x_i = x_n$, $\forall i \geq n$.

Questo concetto fondamentale si ripresenta in forme un poco differenti in diverse branche dell'algebra.

Definizione 2.12. Un insieme parzialmente ordinato X soddisfa la *condizione minimale* se ogni sottoinsieme non vuoto A ha elemento minimale x_0, per cui vale cioè $x_0 \leq x$, $\forall x \in A$.

Teorema 2.13. *Un insieme parzialmente ordinato X soddisfa la condizione della catena discendente se e solo se soddisfa la condizione minimale.*

Dimostrazione. Vedi [1], pagina 74. □

Vogliamo ora spiegare brevemente quanto accade utilizzando la teoria dei grafi, prendendo spunto da [5].
Costruiamo, per un dato $k\langle X\rangle$ e un sistema di riduzione S, un grafo orientato G avente per vertici i polinomi e per lati le riduzioni: i vertici $a, b \in k\langle X\rangle$ sono collegati da un lato orientato da a verso b se e solo se esiste una riduzione semplice $r_{A\sigma B}$ tale che $r_{A\sigma B}(a) = b$. In questo modo possiamo definire una relazione sui vertici di G: $a \geq b$ se esiste una sequenza finita di lati $(a_0 = a, a_1), (a_1, a_2) \ldots (a_{n-1}, b = a_n)$ orientati da a_i verso a_{i+1} che congiungono a con b: poniamo che questa relazione sia d'ordine parziale, in particolare che verifichi la proprietà antisimmetrica (riflessiva e transitiva sono ovvie).
Supponiamo ora che il grafo così costruito soddisfi le condizioni riportate in [2] a pagina 179, vale a dire:

1. Il grafo orientato G soddisfa la condizione della catena discendente, ossia ogni percorso orientato in G termina;

2. il grafo orientato G soddisfa la condizione di confluenza (detta anche condizione a diamante), ossia ogni volta che due lati e, e' escono da un vertice a di G, esistono percorsi orientati p, p' in G che partono dai vertici di arrivo b, b' rispettivamente di e, e' e terminano in un vertice comune c.

Allora, come provato in [5], ogni componente connessa del grafo ha uno e un solo vertice minimale $m \in k\langle X\rangle$, dove con *vertice minimale* intendiamo un vertice m i cui lati afferenti sono tutti orientati verso di esso.
In tal caso, la componente connessa rappresenta una classe di equivalenza della relazione di equivalenza indotta dalle riduzioni: tutti i vertici della componente connessa hanno la medesima forma ridotta, che risulta essere l'elemento identificato dal vertice minimale m. Questo equivale, nel nostro linguaggio, al fatto che la procedura di riduzione sia ben definita e finita.
Si noti che è necessario richiedere sia l'unicità che l'esistenza del vertice minimale: infatti esso potrebbe essere non unico nel caso che da un elemento si possa, tramite sequenze diverse di riduzioni, giungere a due diversi elementi $m', m'' \in k\langle X\rangle_{irr}$; d'altronde il vertice minimale potrebbe anche non esistere se ad esempio la componente connessa risultasse essere un ciclo.

Ciò che vogliamo fare è tradurre le condizioni (1), (2) studiate da Newman in [5] per i grafi orientati nel contesto che abbiamo introdotto, in modo da stabilire le condizioni affinchè in $k\langle X\rangle$, con sistema di riduzione S, ogni monomio si possa opportunamente ridurre fino a raggiungere una forma irriducibile unica (la forma ridotta).

2.2 Ambiguità

Proviamo ora alcune proprietà delle riduzioni e della forma ridotta r_s.

Lemma 2.14. *1. L'insieme degli elementi a riduzione unica è un sottomodulo di $k\langle X\rangle$, inoltre r_s è un morfismo di questo sottomodulo in $k\langle X\rangle_{irr}$.*

2. Siano $a, b, c \in k\langle X\rangle$ tali che per ogni scelta di monomi A, B, C che compaiono rispettivamente in a, b, c il prodotto ABC sia a riduzione unica. Allora il prodotto abc è a riduzione unica.

3. Siano a, b, c come in (2) e r una qualsiasi composizione finita di riduzioni. Allora $ar(b)c$ è a riduzione unica e vale $r_s(ar(b)c) = r_s(abc)$.

Dimostrazione. 1. Siano $a, b \in k\langle X\rangle$ a riduzione unica, $\alpha \in k$. Per la proposizione 2.7 $\alpha a + b$ è a riduzione finita, proviamo ora che è a riduzione unica.
Sia r una successione finale di riduzioni su $\alpha a + b$, dato che a è a riduzione unica, esiste una composizione finita di riduzioni r' tale che $r'r(a) = r_s(a)$ ed analogamente per b esiste una composizione finita di riduzioni r'' tale che $r''r'r(b) = r_s(b)$.
Poichè r è morfismo e $r(\alpha a + b)$ è irriducibile, vale

$$r(\alpha a + b) = r''r'(r(\alpha a + b)) = \alpha r''r'r(a) + r''r'r(b) = \alpha r_s(a) + r_s(b).$$

Dunque una qualsiasi immagine di $\alpha a+b$ tramite una successione finale di riduzioni dà il medesimo elemento, perciò $\alpha a+b$ è a riduzione unica e vale

$$r_s(\alpha a + b) = r(\alpha a + b) = \alpha r_s(a) + r_s(b)$$

ossia r_s è morfismo.

2. Siano $a = \sum_i A_i, b = \sum_j B_j, c = \sum_k C_k$ con A_i, B_j, C_k monomi. Allora $abc = \sum_{i,j,k} A_i B_j C_k$ e per il punto (1) abc è a riduzione unica perchè somma di monomi a riduzione unica.

3. Per il punto (1) è sufficiente provare la tesi per $a = A, b = B, c = C \in \langle X\rangle$ monomi, in quanto r_s è morfismo; inoltre dato che r è composizione *finita* di riduzioni, possiamo limitarci a considerare il caso semplice in cui $r = r_{D\sigma E}$ con $D, E \in \langle X\rangle$ e $\sigma \in S$ qualsiasi.
Ora, se $B \neq DW_\sigma E$ sarà $r_{D\sigma E}(B) = B$ ed anche $ABC \neq ADW_\sigma EC \Rightarrow r_{AD\sigma EC}(ABC) = ABC \Rightarrow Ar_{D\sigma E}(B)C = ABC = r_{AD\sigma EC}(ABC)$.
Se invece $B = DW_\sigma E$ sarà $r_{D\sigma E}(B) = Df_\sigma E$ ed anche $ABC = ADW_\sigma EC \Rightarrow r_{AD\sigma EC}(ABC) = ADf_\sigma EC \Rightarrow Ar_{D\sigma E}(B)C = ADf_\sigma EC = r_{AD\sigma EC}(ABC)$. Perciò in ogni caso $Ar_{D\sigma E}(B)C = r_{AD\sigma EC}(ABC)$.
Poichè ABC è a riduzione unica, anche $r_{AD\sigma EC}(ABC)$ è a riduzione unica e

$$r_s(ABC) = r_s(r_{AD\sigma EC}(ABC)) = r_s(Ar_{D\sigma E}(B)C)$$

il che prova la tesi.

□

Andiamo ora a formalizzare la condizione (2) per i grafi nel nostro linguaggio. I due lati e, e' uscenti dal medesimo vertice a sono dunque due riduzioni diverse che agiscono non banalmente sullo stesso monomio: questo dà origine ad un'*ambiguità*.

Definizione 2.15. Una quintupla (σ, τ, A, B, C) con $\sigma, \tau \in S$ e $A, B, C \in \langle X \rangle - \{1\}$ tale che

$$W_\sigma = AB, \quad W_\tau = BC$$

si dice *ambiguità sovrapposta* di S. In altre parole, in S vi sono le coppie (W_σ, f_σ), (W_τ, f_σ).

Se incontriamo il monomio ABC siamo allora di fronte ad una scelta: applicare la riduzione $r_{1\sigma C}$ oppure la riduzione $r_{A\tau 1}$ ottenendo i due polinomi distinti $f_\sigma C$ o Af_τ. Sorge allora il problema di capire se i due elementi $f_\sigma C$ e Af_τ confluiscono in un unico elemento, si vuole cioè stabilire se esistono riduzioni che trasformano questi due elementi nello stesso polinomio.

Definizione 2.16. L'ambiguità sovrapposta (σ, τ, A, B, C) si dice *risolubile* se esistono due composizioni di riduzioni r, r' tali che

$$r(f_\sigma C) = r'(Af_\tau).$$

Questa è evidentemente la *condizione di confluenza* o *condizione a diamante* e traduce la condizione (2) enunciata a pagina 16.

L'esempio 2.3 e il successivo esempio 2.10 mostrano che non tutte le ambiguità sono risolubili. Per esempio nell'esempio 2.3 vediamo che l'ambiguità sovrapposta (β, α, H, X, Y) è risolubile per S, mentre in 2.10 $(\beta, \alpha', H, X, Y)$ non è risolubile per S'.

Vediamo ora un altro tipo di ambiguità.

Definizione 2.17. Una quintupla (σ, τ, A, B, C) con $\sigma \neq \tau \in S$ e $A, B, C \in \langle X \rangle$ tale che

$$W_\sigma = B, \quad W_\tau = ABC$$

si dice *ambiguità inclusiva* di S.

In questo caso, quando incontriamo ABC dobbiamo scegliere se applicare la riduzione $r_{A\sigma C}$ o la riduzione $r_{1\tau 1}$ ed è dunque necessario capire se $Af_\sigma C$ e f_τ confluiscono in un unico polinomio.

Definizione 2.18. L'ambiguità inclusiva (σ, τ, A, B, C) si dice *risolubile* se esistono due composizioni di riduzioni r, r' tali che

$$r(Af_\sigma C) = r'(f_\tau).$$

Questa è nuovamente la condizione di confluenza.

Si noti che nella definizione di ambiguità sovrapposta richiediamo che A, B, C siano diversi dall'elemento identico, difatti per $A = 1$ o $C = 1$ ricadiamo in un'ambiguità inclusiva, mentre per $B = 1$ non si ha in effetti alcuna ambiguità. Questo spiega anche perchè nella definizione di ambiguità inclusiva è invece necessario permettere i casi in cui uno degli elementi A, B, C sia 1.

Definizione 2.19. Si definisce *ordinamento parziale su* $\langle X \rangle$ un ordine parziale $\leq$ tale che

$$B < B' \Rightarrow ABC < AB'C \quad \forall A, B, B', C \in \langle X \rangle - \{1\}.$$

Questo ordinamento parziale si dice *compatibile col sistema di riduzione* S se per ogni $\sigma \in S$, f_σ è combinazione lineare di monomi ognuno dei quali $< W_\sigma$.
In particolare questo significa che per un qualsiasi polinomio $a \in k\langle X \rangle$ e una qualsiasi riduzione $r_{A\sigma B}$ non esistono monomi di $r_{A\sigma B}(a)$ che siano $>$ di ogni monomio di a.

Tornando agli esempi 2.3 e 2.10 vediamo che sia S che S' sono compatibili con l'ordine parziale $Y < X < H$ esteso lessicograficamente sui monomi dello stesso grado, cioè $Z_{i_1} \dots Z_{i_k} < Z_{j_1} \dots Z_{j_k}$ se e solo se $Z_{i_1} < Z_{j_1}$ oppure $Z_{i_1} \dots Z_{i_h} = Z_{j_1} \dots Z_{j_h}$ e $Z_{i_{h+1}} < Z_{j_{h+1}}$. Ovviamente, di due monomi di grado diverso il maggiore è quello di grado più alto.

Definizione 2.20. Sia I l'ideale bilatero di $k\langle X \rangle$ così generato:

$$I := \langle W_\sigma - f_\sigma \rangle_{\sigma \in S}.$$

Come sottomodulo, I è evidentemente generato dai prodotti $A(W_\sigma - f_\sigma)B$ al variare di $\sigma \in S$ e di $A, B \in \langle X \rangle$.
Dato $\leq$ ordinamento parziale compatibile con S, per ogni $A \in \langle X \rangle$ si definisce il sottomodulo di $k\langle X \rangle$

$$I_A := \big(B(W_\sigma - f_\sigma)C\big)$$

al variare di $B, C \in \langle X \rangle$, $\sigma \in S$ tali che

$$BW_\sigma C < A.$$

Definizione 2.21. Sia $\leq$ ordinamento parziale compatibile con S.
L'ambiguità sovrapposta (σ, τ, A, B, C) si dice *risolubile relativamente a* $\leq$ se $f_\sigma C - Af_\tau \in I_{ABC}$, vale a dire se $f_\sigma C - Af_\tau$ è somma di termini del tipo $D(W_\upsilon - f_\upsilon)E$ per ognuno dei quali si ha $DW_\upsilon E < ABC$.
L'ambiguità inclusiva (σ, τ, A, B, C) si dice *risolubile relativamente a* $\leq$ se $Af_\sigma C - f_\tau \in I_{ABC}$, vale a dire se $Af_\sigma C - f_\tau$ è somma di termimi del tipo $D(W_\upsilon - f_\upsilon)E$ per ognuno dei quali si ha $DW_\upsilon E < ABC$.

Proposizione 2.22. *Sia $\leq$ ordinamento parziale compatibile con S. Allora ogni ambiguità risolubile è risolubile relativamente a $\leq$.*

Dimostrazione. Sia $f \in k\langle X\rangle$ qualsiasi e $r_{A\sigma C}$ non banale su f; poniamo $g = r_{A\sigma C}(f)$. Allora

$$f - g = cA(W_\sigma - f_\sigma)C \tag{2.1}$$

con $c \in k - \{0\}$, in quanto ogni monomio diverso da $AW_\sigma C$ compare identico in f e g. Useremo spesso questo fatto.
Sia ora (σ, τ, A, B, C) un'ambiguità sovrapposta risolubile, dunque esistono le riduzioni r, r' tali che $r(f_\sigma C) = r'(Af_\tau)$ per le quali si ha $r = r_1 \dots r_n$, $r' = r'_1 \dots r'_m$ con $r_i = r_{A_i\sigma_i C_i}$, $r'_j = r'_{A'_j\sigma'_j C'_j}$ riduzioni semplici.
Vale quindi

$$f_\sigma C = f_n \overset{r_n}{\longmapsto} f_{n-1} \overset{r_{n-1}}{\longmapsto} \dots \overset{r_1}{\longmapsto} f_0,$$

$$Af_\tau = g_m \overset{r'_m}{\longmapsto} g_{m-1} \overset{r'_{m-1}}{\longmapsto} \dots \overset{r'_1}{\longmapsto} g_0 = f_0.$$

Perciò, per il risultato (2.1)

$$f_\sigma C - f_0 = \sum_{h=1}^{n} f_h - f_{h-1} = \sum_{h=1}^{n} f_h - r_h(f_h) = \sum_{h=1}^{n} c_h A_h(W_{\sigma_h} - f_{\sigma_h})C_h,$$

$$Af_\tau - g_0 = \sum_{k=1}^{m} g_k - g_{k-1} = \sum_{k=1}^{m} g_k - r'_k(g_k) = \sum_{k=1}^{m} c'_k A'_k(W_{\sigma'_k} - f_{\sigma'_k})C'_k,$$

con $c_h, c'_k \in k - \{0\}$.
Dato che $\leq$ è ordinamento parziale compatibile con S, $f_\sigma C$ e Af_τ sono combinazioni lineari di monomi $< W_\sigma C = ABC = AW_\tau$, e perciò anche f_h, g_k al variare di $h \in \{0 \dots n\}, k \in \{0 \dots m\}$ sono combinazioni lineari di monomi $< ABC$; fra questi monomi compaiono anche $A_h W_{\sigma_h} C_h$ elemento di f_h e $A'_k W_{\sigma'_k} C'_k$ elemento di g_k.
Si ha allora

$$f_\sigma C - Af_\tau = \sum_{h=1}^{n} c_h A_h(W_{\sigma_h} - f_{\sigma_h})C_h - \sum_{k=1}^{m} c'_k A'_k(W_{\sigma'_k} - f_{\sigma'_k})C'_k \in I_{ABC}$$

perchè come appena visto $A_h W_{\sigma_h} C_h, A'_k W_{\sigma'_k} C'_k < ABC$ e quindi ognuno degli addendi fa parte del sottomodulo I_{ABC}.
Questo prova allora che (σ, τ, A, B, C) è risolubile relativamente a $\leq$.

La prova per l'ambiguità inclusiva procede in maniera assolutamente analoga: sia (σ, τ, A, B, C) un'ambiguità inclusiva risolubile, dunque esistono le riduzioni r, r' tali che $r(Af_\sigma C) = r'(f_\tau)$ per le quali si ha $r = r_1 \dots r_n$,

$r' = r'_1 \dots r'_m$ con $r_i = r_{A_i\sigma_i C_i}$, $r'_j = r'_{A'_j\sigma'_j C'_j}$ riduzioni semplici.
Vale quindi

$$Af_\sigma C = f_n \overset{r_n}{\longmapsto} f_{n-1} \overset{r_{n-1}}{\longmapsto} \dots \overset{r_1}{\longmapsto} f_0,$$

$$f_\tau = g_m \overset{r'_m}{\longmapsto} g_{m-1} \overset{r'_{m-1}}{\longmapsto} \dots \overset{r'_1}{\longmapsto} g_0 = f_0.$$

Perciò, per il risultato (2.1)

$$Af_\sigma C - f_0 = \sum_{h=1}^{n} f_h - f_{h-1} = \sum_{h=1}^{n} f_h - r_h(f_h) = \sum_{h=1}^{n} c_h A_h (W_{\sigma_h} - f_{\sigma_h}) C_h,$$

$$f_\tau - f_0 = \sum_{k=1}^{m} g_k - g_{k-1} = \sum_{k=1}^{m} g_k - r'_k(g_k) = \sum_{k=1}^{m} c'_k A'_k (W_{\sigma'_k} - f_{\sigma'_k}) C'_k,$$

con $c_h, c'_k \in k - \{0\}$.
Dato che $\leq$ è ordinamento parziale compatibile con S, $Af_\sigma C$ e f_τ sono combinazioni lineari di monomi $< AW_\sigma C = ABC = W_\tau$, e perciò anche f_h, g_k al variare di $h \in \{0 \dots n\}, k \in \{0 \dots m\}$ sono combinazioni lineari di monomi $< ABC$; fra questi monomi compaiono anche $A_h W_{\sigma_h} C_h$ elemento di f_h e $A'_k W_{\sigma'_k} C'_k$ elemento di g_k.
Si ha allora

$$Af_\sigma C - f_\tau = \sum_{h=1}^{n} c_h A_h (W_{\sigma_h} - f_{\sigma_h}) C_h - \sum_{k=1}^{m} c'_k A'_k (W_{\sigma'_k} - f_{\sigma'_k}) C'_k \in I_{ABC}$$

perchè come appena visto $A_h W_{\sigma_h} C_h, A'_k W_{\sigma'_k} C'_k < ABC$ e quindi ognuno degli addendi fa parte del sottomodulo I_{ABC}.
Questo prova allora che (σ, τ, A, B, C) è risolubile relativamente a $\leq$.

□

2.3 Il Diamond Lemma di Bergman

In questa sezione conclusiva, enunciamo e dimostriamo il Diamond Lemma, nella formulazione di Bergman.

Lemma 2.23. *Siano S un sistema di riduzione su una k-algebra associativa libera $k\langle X\rangle$ e $\leq$ un ordinamento parziale su $\langle X\rangle$ compatibile con S che rispetta la condizione della catena discendente (DCC). Allora ogni elemento di $k\langle X\rangle$ è a riduzione finita.*

Dimostrazione. Per assurdo sia

$$N = \{A \in \langle X \rangle \text{ tali che } A \text{ non è a riduzione finita } \} \neq \emptyset.$$

Dato che $\leq$ soddisfa la DCC, N possiede un elemento minimale M_0 che non è a riduzione finita, sia allora $r_1 = r_{A\sigma B}$ una riduzione non banale su di esso con $AW_\sigma B = M_0$ tale che $r_{A\sigma B}(M_0) = Af_\sigma B$.
Per la compatibilità di $\leq$ con S i monomi di $Af_\sigma B$ sono $< M_0$ e dunque sono a riduzione finita per la minimalità di M_0 in N.
Perciò anche M_0 è a riduzione finita, in quanto per una qualsiasi successione infinita di riduzioni $(r_n)_{n\in\mathbb{N}}$ si ha r_i che agisce banalmente su $r_{i-1} \dots r_2(r_1(M_0))$ con i sufficientemente grande. Questo è assurdo, dunque $N = \emptyset$, perciò ogni elemento di $\langle X \rangle$ è a riduzione finita, ma per la proposizione 2.7 gli elementi a riduzione finita formano un sottomodulo, dunque ogni elemento di $k\langle X \rangle$ è a riduzione finita. □

Teorema 2.24 (Diamond Lemma)**.** *Siano S un sistema di riduzione su una k-algebra associativa libera $k\langle X \rangle$ e $\leq$ un ordinamento parziale su $\langle X \rangle$ compatibile con S che rispetta la condizione della catena discendente (DCC).*
Le seguenti condizioni sono allora equivalenti:

1. *Ogni ambiguità di S è risolubile.*
2. *Ogni ambiguità di S è risolubile relativamente a $\leq$.*
3. *Ogni elemento di $k\langle X \rangle$ è a riduzione unica.*
4. *Un insieme di rappresentanti in $k\langle X \rangle$ per gli elementi dell'algebra $R = k\langle X \rangle / I$, dove $I = (W_\sigma - f_\sigma)_{\sigma \in S}$ è l'ideale bilatero della definizione 2.20, è dato dal sottomodulo $k\langle X \rangle_{irr}$ generato dai monomi di $\langle X \rangle$ irriducibili secondo S.*

Dimostrazione. Per il lemma, ogni elemento di $k\langle X \rangle$ è a riduzione finita.
Osserviamo subito che poichè $k\langle X \rangle_{irr} \cap I = \{0\}$ in quanto ogni elemento non banale di I è riducibile, abbiamo

$$(4) \iff k\langle X \rangle = k\langle X \rangle_{irr} \oplus I.$$

- (3) $\Rightarrow$ (4). Proviamo ora che $ker(r_s) = I$ da cui seguirà (4) per il teorema fondamentale di omomorfismo in quanto $r_s : k\langle X \rangle \longrightarrow k\langle X \rangle_{irr}$ è morfismo suriettivo.
  Siano $A, B \in \langle X \rangle$ e $\sigma \in S$ qualsiasi, per il lemma 2.14.(1) si ha $r_s(A(W_\sigma - f_\sigma)B) = r_s(AW_\sigma B) - r_s(Af_\sigma B)$ e per il lemma 2.14.(3) vale $r_s(AW_\sigma B) = r_s(r_{A\sigma B}(AW_\sigma B)) = r_s(Af_\sigma B)$ dunque r_s si annulla

su ogni generatore del sottomodulo I, perciò $I \subseteq ker(r_s)$.
Viceversa, sia $f \in ker(r_s)$, esiste dunque una successione finita di riduzioni semplici $r_1, \ldots, r_n$ tale che $r_n \ldots r_1 = r_s$ e si ha perciò la somma telescopica

$$f - r_s(f) = f - r_n \ldots r_1(f) =$$

$$= \big[f - r_1(f)\big] + \big[r_1(f) - r_2(r_1(f))\big] + \ldots +$$
$$+\big[r_{n-1}(\ldots(r_1(f))) - r_n(r_{n-1}(\ldots(r_1(f))))\big]$$

dove ogni termine fra parentesi $\big[\,\big]$ appartiene a I per il risultato (2.1). Dato che $r_s(f) = 0$, si ha $f \in I$ perciò $ker(r_s) \subseteq I$ e questo prova la tesi.

- (4) $\Rightarrow$ (3). Sia $f \in k\langle X\rangle$ tale che esistano successioni finali di riduzioni semplici $r_1, \ldots, r_m$ e $r_{m+1}, \ldots, r_{m+n}$ per le quali $r_m \ldots r_1(f) = f_1$ e $r_{m+n} \ldots r_{m+1}(f) = f_2$, con $f_1, f_2 \in k\langle X\rangle_{irr}$. Vogliamo mostrare che $f_1 = f_2$.
Si hanno dunque le seguenti somme telescopiche:

$$f_1 - f_2 = r_m \ldots r_1(f) - r_{m+n} \ldots r_{m+1}(f) =$$

$$= \big[r_m(r_{m-1} \ldots r_1(f)) - r_{m-1} \ldots r_1(f)\big] + \ldots + \big[r_1(f) - f\big] +$$
$$+\big[f - r_{m+1}(f)\big] + \ldots + \big[r_{m+n-1} \ldots r_{m+1}(f) - r_{m+n}(r_{m+n-1} \ldots r_{m+1}(f))\big]$$

dove ogni termine fra parentesi $\big[\,\big]$ appartiene a I per il risultato (2.1). Perciò $f_1 - f_2 \in k\langle X\rangle_{irr} \cap I = \{0\}$ dunque $f_1 = f_2$ e perciò f è a riduzione unica.

- (3) $\Rightarrow$ (1). Sia (σ, τ, A, B, C) un'ambiguità sovrapposta di S; dato che ogni elemento è a riduzione unica, per il lemma 2.14.(3) si ha

$$r_s(f_\sigma C) = r_s(r_{1\sigma 1}(W_\sigma)C) = r_s(r_{1\sigma 1}(AB)C) = r_s(ABC).$$

D'altra parte:

$$r_s(Af_\tau) = r_s(Ar_{1\tau 1}(W_\tau)) = r_s(Ar_{1\tau 1}(BC)) = r_s(ABC).$$

e dunque l'ambiguita è risolubile, perchè $r_s(f_\sigma C) = r_s(Af_\tau)$ ed r_s è una composizione finita di riduzioni.
Sia (σ, τ, A, B, C) un'ambiguità inclusiva di S; dato che ogni elemento è a riduzione unica, per il lemma 2.14.(3) si ha

$$r_s(Af_\sigma C) = r_s(Ar_{1\sigma 1}(W_\sigma)C) = r_s(Ar_{1\sigma 1}(B)C) = r_s(ABC).$$

D'altra parte:

$$r_s(f_\tau) = r_s(r_{1\tau 1}(W_\tau)) = r_s(r_{1\tau 1}(ABC)) = r_s(ABC)$$

e dunque l'ambiguita è risolubile, perchè $r_s(Af_\sigma C) = r_s(f_\tau)$ ed r_s è una composizione finita di riduzioni.

- (1) $\Rightarrow$ (2) per la proposizione 2.22.
- (2) $\Rightarrow$ (3). Per assurdo sia

$$N = \{A \in \langle X \rangle \text{ tali che } A \text{ non è a riduzione unica } \} \neq \emptyset.$$

Dato che $\leq$ soddisfa la DCC, N possiede un elemento minimale M_0 che non è a riduzione unica, mentre ogni elemento $< M_0$ è a riduzione unica. In particolare, poichè il dominio di r_s è formato dagli elementi a riduzione unica, esso contiene tutti questi monomi e perciò il sottomodulo da essi generato.
I generatori del sottomodulo $I_{M_0} = \langle A(W_\sigma - f_\sigma)B \rangle$, $AW_\sigma B < M_0$ (definito come in 2.20) sono nuovamente tutti a riduzione unica, perchè f_σ è formato da monomi ognuno $< W_\sigma$ e per costruzione $AW_\sigma B < M_0$, dunque anche $Af_\sigma B$ è formato da monomi ognuno $< M_0$ per la compatibilità di $\leq$ con S.
Per ognuno di questi generatori vale però

$$r_s(A(W_\sigma - f_\sigma)B) = r_s(AW_\sigma B) - r_s(Af_\sigma B) =$$

$$= r_s(AW_\sigma B) - r_s(Ar_{1\sigma 1}(W_\sigma)B) = 0$$

grazie al lemma 2.14.(3), e dunque $I_{M_0} \subset ker(r_s)$.
Proviamo ora che due qualsiasi riduzioni $r_{L\sigma M'}, r_{L'\tau M}$ che agiscono non banalmente su M_0 (e dunque tali che $LW_\sigma M' = M_0 = L'W_\tau M$) danno $r_s(r_{L\sigma M'}(M_0)) = r_s(r_{L'\tau M}(M_0))$ e dunque M_0 è a riduzione unica.
Supponendo, senza perdere in generalità, che la copia di W_σ in M_0 cominci prima, più a sinistra, della copia di W_τ, è necessario distinguere tre casi:

- W_σ e W_τ si sovrappongono in M_0 ma nessun monomio contiene l'altro, ossia $M_0 = LABCM$ con $AB = W_\sigma, BC = W_\tau$, dove (σ, τ, A, B, C) è una ambiguità sovrapposta di S. In questo caso $M' = CM$, $L' = LA$. Vale evidentemente

$$r_{L\sigma M'}(M_0) - r_{L'\tau M}(M_0) = Lf_\sigma CM - LAf_\tau M = L(f_\sigma C - Af_\tau)M.$$

Per l'ipotesi (2) l'ambiguità (σ, τ, A, B, C) è risolubile relativamente a $\leq$, cioè $f_\sigma C - A f_\tau \in I_{ABC}$ e dunque

$$L(f_\sigma C - A f_\tau)M \in I_{LABCM} = I_{M_0} \subset ker(r_s)$$

perciò $r_s(r_{L\sigma M'}(M_0)) = r_s(r_{L'\tau M}(M_0))$ come volevamo.

– W_σ e W_τ si sovrappongono in M_0 e W_σ contiene W_τ, ossia $M_0 = LABCM'$ con $ABC = W_\sigma, B = W_\tau$ dove (τ, σ, A, B, C) è una ambiguità inclusiva di S. In questo caso $L' = LA$, $M = CM'$. Vale evidentemente

$$r_{L\sigma M'}(M_0) - r_{L'\tau M}(M_0) = L f_\sigma M' - LA f_\tau C M' = L(f_\sigma - A f_\tau C)M'.$$

Per l'ipotesi (2) l'ambiguità (τ, σ, A, B, C) è risolubile relativamente a $\leq$, cioè $f_\sigma - A f_\tau C \in I_{ABC}$ e dunque

$$L(f_\sigma - A f_\tau C)M' \in I_{LABCM'} = I_{M_0} \subset ker(r_s)$$

perciò $r_s(r_{L\sigma M'}(M_0)) = r_s(r_{L'\tau M}(M_0))$ come volevamo.

– W_σ e W_τ sono disgiunti in M_0, cioè $M_0 = A W_\sigma B W_\tau C$ dove $L = A$, $M' = B W_\tau C$, $L' = A W_\sigma B$, $M = C$. Questo è il caso più semplice. Per il lemma 2.14.(3) si ha allora

$$r_s(r_{L\sigma M'}(M_0)) = r_s(A f_\sigma B W_\tau C) = r_s(A f_\sigma r_{B\tau C}(B W_\tau C)) =$$

$$= r_s(A f_\sigma B f_\tau C) = r_s(r_{A\sigma B}(A W_\sigma B) f_\tau C) =$$

$$= r_s(A W_\sigma B f_\tau C) = r_s(r_{L'\tau M}(M_0))$$

come volevasi dimostrare.

Abbiamo quindi provato che due qualsiasi riduzioni non banali su M_0 hanno la stessa forma ridotta, dunque M_0 è a riduzione unica, ma questo è assurdo perciò $N = \emptyset$.
Ogni monomio di $\langle X \rangle$ è allora a riduzione unica, ma gli elementi a riduzione unica formano un sottomodulo per il lemma 2.14.(1), dunque ogni elemento di $k\langle X \rangle$ è a riduzione unica.

□

Mostriamo per completezza un esempio in cui vi sono ambiguità *non* risolubili.

Esempio 2.25. In $k\langle x, y, z\rangle$ definiamo il sistema di riduzione

$$S = \{\sigma = (xy, yx), \tau = (x^2y, xy + y)\}.$$

Vediamo come queste agiscono su x^2y, si ha:

$$x^2y \xrightarrow{r_{1\tau 1}} xy + y \xrightarrow{r_{1\sigma 1}} yx + y.$$

Ed anche

$$x^2y \xrightarrow{r_{x\sigma 1}} xyx \xrightarrow{r_{1\sigma x}} yx^2.$$

Evidentemente $yx + y, yx^2 \in k\langle x, y, z\rangle_{irr}$ e sono polinomi diversi, dunque l'ambiguità inclusiva $(\sigma, \tau, x, xy, 1)$ non è risolubile.

Riprendiamo ora gli esempi 2.3 e 2.10. In questi due esempi si esaminavano due sistemi di riduzione molto simili sull'alfabeto H, X, Y:

$$S = \{\alpha = (XY, YX + H), \beta = (HX, XH + 2X), \gamma = (HY, YH - 2Y)\},$$

$$S' = \{\alpha' = (XY, YX + H + X), \beta = (HX, XH + 2X), \gamma = (HY, YH - 2Y)\}.$$

Dagli esempi appare chiaro che l'elemento HXY è a riduzione unica per S ma non per S'; dal teorema precedente ciò equivale al fatto che l'ambiguità $(\beta, \alpha', H, X, Y)$ non è risolubile per S'.
Mentre è chiaro che entrambi i sistemi di riduzione S ed S' ci permettono di riordinare qualunque polinomio nell'alfabeto dato, secondo qualunque ordine scelto, non è chiaro invece quale sia la condizione che un sistema S generico, preposto ad ordinare polinomi nell'alfabeto H, X, Y, deve soddisfare affinchè tutte le ambiguità siano risolubili.
Come vedremo nel prossimo capitolo, per i sistemi S che provengono da algebre di Lie tale condizione è rappresentata dall'*identità di Jacobi*.

Capitolo 3

Il Teorema di Poincaré-Birkhoff-Witt

In questo capitolo enunciamo e dimostriamo il Teorema di Poincaré-Birkhoff-Witt (d'ora in avanti anche indicato come PBW), presentando due trattazioni diverse ed indipendenti: nel primo caso assumeremo l'algebra di Lie $\mathfrak{g}$ definita su di un anello, mentre nel secondo caso supporremo di trovarci su di un campo.

3.1 Il Teorema PBW sugli anelli

Sia k anello commutativo unitario con 2, 3 invertibili.
Tutti i moduli, i morfismi e le algebre, ove non diversamente specificato, sono da intendersi rispettivamente come k-moduli, k-morfismi e k-algebre.
Sia X un insieme e consideriamo $\mathfrak{g}$ modulo libero di base X. Supponiamo inoltre che $\mathfrak{g}$ sia un'algebra di Lie.
Come nella definizione 1.23, definiamo l'ideale bilatero I dell'algebra tensoriale $T(\mathfrak{g})$ come:

$$I = \langle x \otimes y - y \otimes x - [x, y] | x, y \in \mathfrak{g} \rangle.$$

Si ha allora $U(\mathfrak{g}) = T(\mathfrak{g})/I$ ed è definito il morfismo

$$\tau : \mathfrak{g} \longrightarrow U(\mathfrak{g}), \quad \tau = p \circ i,$$

dove $i : \mathfrak{g} \hookrightarrow T(\mathfrak{g})$ è l'immersione e $p : T(\mathfrak{g}) \longrightarrow U(\mathfrak{g})$ è la proiezione sul quoziente. Useremo nel seguito la notazione:

$$x' = \tau(x) \quad \forall x \in \mathfrak{g}.$$

Dalle definizioni abbiamo immediatamente la seguente osservazione.

Osservazione 3.1. *Si ha l'isomorfismo di algebre associative $k\langle X\rangle \cong T(\mathfrak{g})$ secondo l'identificazione che sui generatori (come moduli) dà:*

$$x_1 \dots x_n \cong x_1 \otimes \dots \otimes x_n$$

per $x_1, \dots, x_n \in X$.

Dimostrazione. Definiamo $f : k\langle X\rangle \longrightarrow T(\mathfrak{g})$ come mappa lineare che sui generatori di $k\langle X\rangle$ agisce nel seguente modo: $x_1 \dots x_n \mapsto x_1 \otimes \dots \otimes x_n$, proviamo che è morfismo di algebre. Vale $f(ax_1 \dots x_n) = ax_1 \otimes \dots \otimes x_n = af(x_1 \dots x_n)$ per ogni $a \in k$, ed anche $f(x_1 \dots x_n + y_1 \dots y_m) = x_1 \otimes \dots \otimes x_n + y_1 \otimes \dots \otimes y_m = f(x_1 \dots x_n) + f(y_1 \dots y_m)$ grazie alla linearità, dunque f è morfismo di moduli. Inoltre $f(x_1 \dots x_n y_1 \dots y_m) = x_1 \otimes \dots \otimes x_n \otimes y_1 \otimes \dots \otimes y_m = f(x_1 \dots x_n) \otimes f(y_1 \dots y_m)$ per l'associatività del prodotto tensoriale, quindi f è morfismo di algebre. Definiamo ora $g : T(\mathfrak{g}) \longrightarrow k\langle X\rangle$ come mappa lineare che agisce sui generatori di $T(\mathfrak{g})$ come segue: $x_1 \otimes \dots \otimes x_n \mapsto x_1 \dots x_n$, è evidente che g ed f sono una l'inversa dell'altra, proviamo che g è morfismo di algebre da cui seguirà la tesi. Si ha dunque $g(ax_1 \otimes \dots \otimes x_n) = ax_1 \dots x_n = ag(x_1 \otimes x_n)$ per ogni $a \in k$, ed anche $g(x_1 \otimes \dots \otimes x_n + y_1 \otimes \dots \otimes y_m) = x_1 \dots x_n + y_1 \dots y_m = g(x_1 \otimes \dots \otimes x_n) + g(y_1 \otimes \dots \otimes y_m)$ grazie alla linearità, quindi g è morfismo di moduli. Inoltre $g(x_1 \otimes \dots \otimes x_n \otimes y_1 \otimes \dots \otimes y_m) = x_1 \dots x_n y_1 \dots y_m = g(x_1 \otimes \dots \otimes x_n)g(y_1 \otimes \dots \otimes y_m)$, quindi g è morfismo di algebre e segue la tesi. □

Identificheremo perciò nel seguito $k\langle X\rangle$ con $T(\mathfrak{g})$ e cioè trascureremo il simbolo di prodotto tensoriale sostituendolo con quello di prodotto in $k\langle X\rangle$.

Proposizione 3.2. *Sia $\leq$ un ordine totale su X. Allora si ha*

$$I = \langle xy - yx - [x, y] | x < y \text{ con } x, y \in X\rangle.$$

ossia I è generato dai soli elementi della base X, scelti in modo tale che $x < y$.

Dimostrazione. Si noti che abbiamo scritto I secondo la convenzione introdotta dalla proposizione precedente, cioè $I \subseteq k\langle X\rangle \cong T(\mathfrak{g})$.
Sia $I' := \langle xy - yx - [x, y] | x < y \in X\rangle$, vogliamo provare che $I' = I$.
Evidentemente $I' \subseteq I$ perchè fra i generatori di I vi sono anche quelli di I'. Notiamo poi che per $y < x$ si ha $yx - xy - [y, x] = yx - xy + [x, y] = -(xy - yx - [x, y]) \in I'$ dunque $xy - yx - [x, y] \in I'$ anche per $x > y$.
Siano ora $a, b \in \mathfrak{g}$, esistono allora $x_i, y_j \in X$ tali che $a = \sum_i a_i x_i$, $b = \sum_j b_j y_j$ con $a_i, b_j \in k$. Per il generico generatore di I si ha dunque:

$$ab - ba - [a, b] = (\sum_i a_i x_i)(\sum_j b_j y_j) - (\sum_j b_j y_j)(\sum_i a_i x_i) -$$

$$-[\sum_i a_i x_i, \sum_j b_j y_j] = \sum_{i,j} a_i b_j (x_i y_j) - \sum_{j,i} b_j a_i (y_j x_i) - \sum_{i,j} a_i b_j [x_i, y_j] =$$

$$= \sum_{i,j} a_i b_j (x_i y_j - y_j x_i - [x_i, y_j])$$

grazie alla bilinearità della bracket. Tale generatore è combinazione lineare di elementi di I', dunque $ab - ba - [a,b] \in I'$ e perciò $I \subseteq I'$ da cui la tesi. □

Teorema 3.3 (Teorema PBW sugli anelli)**.** *Sia $\mathfrak{g}$ algebra di Lie e modulo libero con base X. Sia $\leq$ un ordine totale su X. Allora l'algebra inviluppante universale $U(\mathfrak{g})$ è modulo libero di base $1 \cup \{x'_1 \dots x'_n\}$ al variare di $n \in \mathbb{N}$ e di $x_1, \dots, x_n \in X$, tali che $x_1 \leq \dots \leq x_n$.*

Ricordiamo che $x' = \tau(x)$ per ogni $x \in \mathfrak{g}$, dove $\tau : \mathfrak{g} \longrightarrow U(\mathfrak{g})$, dunque $\{x'_j\}_j$ non sono altro che le immagini in $U(\mathfrak{g})$ della base X.

Dimostrazione. Sia S il sistema di riduzione su $k\langle X\rangle$ formato dalle coppie $\sigma_{xy} = (yx, xy - [x,y])$ per ogni $y > x$ con $x, y \in X$. Per l'ideale bilatero di $k\langle X\rangle$ della definizione 2.20 si ha allora

$$J = \langle W_\sigma - f_\sigma \rangle_{\sigma \in S} = \langle yx - xy + [x,y] \rangle_{y>x} = \langle xy - yx - [x,y] \rangle_{y>x} = I$$

per la proposizione precedente. Inoltre, l'immagine secondo $p : k\langle X\rangle \longrightarrow U(\mathfrak{g})$ di $k\langle X\rangle_{irr}$ ha per sistema di generatori la base nella tesi, in quanto, per come sono state costruite le riduzioni, un elemento $f = \sum a_{i_1 \dots i_n} x_{i_1} \dots x_{i_n}$ di $k\langle X\rangle$ è irriducibile se e solo se per ogni suo multindice $(i_1, \dots, i_n)$ si ha $x_{i_1} \leq \dots \leq x_{i_n}$.
Proviamo ora che in questo sistema di riduzione ogni elemento di $k\langle X\rangle$ è a riduzione finita. Faremo ciò utilizzando il lemma 2.23, dunque prima definiremo un ordine parziale su $k\langle X\rangle$ compatibile con S, poi mostreremo che soddisfa la DCC.
Definiamo l'*indice di disordine* di un monomio $A = x_1 \dots x_n \in \langle X\rangle$ come il numero i_A delle coppie (i,j) tali che $i < j$ ma $x_i > x_j$, ad esempio 0 se $x_1 \leq \dots \leq x_n$ oppure $\frac{n(n-1)}{2}$ se $x_1 > \dots > x_n$, e la sua *lunghezza* l_A come il numero di termini di X da cui è formato. Consideriamo ora la seguente relazione su $\langle X\rangle$: $A < B$ se $l_A < l_B$ oppure se $l_A = l_B$ e A è permutazione dei termini x_i di B ma $i_A < i_B$. Si noti in particolare che $A < B \Rightarrow l_A \leq l_B$ in quanto se $l_A \not< l_B$, $A < B$ implica che A e B sono permutazione degli stessi termini ed hanno perciò la stessa lunghezza.
Questa relazione risulta essere una relazione d'ordine stretto e parziale: infatti $A \not< A$, $A < B \Rightarrow B \not< A$ sono ovvie, e per $A < B$, $B < C$ se A non ha lunghezza minore di C, A, B, C sono permutazioni dei medesimi termini

e vale $i_A < i_B < i_C \Rightarrow A < C$.
Proviamo ora che $<$ è un ordinamento parziale su $\langle X \rangle$ compatibile con S secondo la definizione 2.19. In altre parole vogliamo mostrare che $<$ è ordinamento parziale su X, cioè se $B < B'$ allora $ABC < AB'C$, ed inoltre che è compatibile con S, cioè $\forall \sigma \in S$ f_σ è combinazione lineare di monomi minori di W_σ. Siano $B < B'$ e $A, C \in \langle X \rangle$ qualsiasi. Se $l_B < l_{B'}$ allora $l_{ABC} < l_{AB'C} \Rightarrow ABC < AB'C$. Se $l_B = l_{B'}$ allora B e B' sono permutazioni dei medesimi termini con $i_B < i_{B'}$; in tal caso, anche ABC e $AB'C$ sono permutazioni dei medesimi termini, confrontiamone gli indici di disordine scegliendo una coppia di termini che dà un'inversione in ABC, per cui cioè $i < j$ ma $x_i > x_j$:

- se x_i, x_j sono entrambi in $A \cup C$, la coppia (x_i, x_j) dà un'inversione anche in $AB'C$;
- se x_i, x_j sono entrambi in B, permutazione di B', si ha $i_B < i_{B'}$ per ipotesi;
- se x_i è presente in A e x_j compare in B, si ha x_j presente anche in B' e dunque tale coppia dà inversione anche in $AB'C$;
- se x_i è presente in B e x_j compare in C, si ha x_i presente anche in B' e dunque tale coppia dà inversione anche in $AB'C$.

Questo prova che $i_{ABC} < i_{AB'C} \Rightarrow ABC < AB'C$ e dunque $<$ è ordinamento parziale su $\langle X \rangle$.
Ora è evidente che per una qualsiasi riduzione σ_{xy} si ha $l_{yx} = 2$, $i_{yx} = 1$, d'altra parte $i_{xy} = 0 \Rightarrow yx > xy$ e $[x, y] = \sum_i a_i x_i$, con $l_{x_i} = 1 \Rightarrow yx > x_i$, dunque W_σ è maggiore di ogni monomio che appare in f_σ e questo prova che l'ordine $<$ è compatibile con S.
Mostriamo ora che $<$ soddisfa, su $\langle X \rangle$, la descending chain condition (vedi definizione 2.11). Sia $A_1 \geq \ldots \geq A_n \geq \ldots$ una catena discendente, con $A_j \in \langle X \rangle$; per come è definito $<$ vale:

$$A_j > A_{j+1} \Rightarrow l_{A_j} > l_{A_{j+1}} \text{ oppure } l_{A_j} = l_{A_{j+1}} \text{ e } i_{A_j} > i_{A_{j+1}}$$

per ogni j.
Supponiamo per assurdo che la catena sia non stazionaria, cioè che vi sia un numero infinito di disuguglianze strette. Fissato A_j, il suo numero massimo di coppie disordinate è $i_j := \frac{l_{A_j}(l_{A_j}-1)}{2}$, che sarà quindi il massimo indice di disordine che può assumere A_j. Consideriamo il monomio A_k che si trova $i_j + 1$ disuguaglianze strette dopo A_j, per il quale esistono i_j monomi A_i tutti diversi fra loro tali che $A_j > A_i > A_k$. A_k avrà allora lunghezza

strettamente minore di A_j, perchè se per assurdo fosse $l_{A_k} = l_{A_j}$, si avrebbe $i_{A_k} \leq i_{A_j} - (i_j + 1) \leq i_j - (i_j + 1) = -1$, assurdo in quanto l'indice di disordine è un numero non negativo. Dunque in un numero finito di disuguaglianze strette la lunghezza diminuisce almeno di 1, d'altronde l_{A_1} è fissato, perciò in un numero finito di disuguaglianze la lunghezza calerà fino ad essere $l_{A_N} = 1$; da questo momento in poi la catena non può che essere stazionaria perchè il monomio A_N è formato da un solo elemento: ciò è assurdo perchè avevamo supposto che la catena fosse non stazionaria. Questo prova la validità della DCC e ciò conclude la dimostrazione del fatto che ogni elemento di $k\langle X\rangle$ è a riduzione finita rispettivamente al sistema di riduzione considerato.
Si noti ora che per ogni $a, b \in \mathfrak{g}$ con $a = \sum_i a_i x_i$, $b = \sum_j b_j y_j$, si ha

$$ab - ba - [a, b] = \sum_{i,j} a_i b_j (x_i y_j - y_j x_i - [x_i, y_j])$$

dove per ogni i, j sarà $x_i > y_j$ oppure $x_i < y_j$. Per ogni addendo vale allora

$$(x_i y_j - y_j x_i - [x_i, y_j]) = (x_i y_j - y_j x_i + [y_j, x_i]) = (W_{\sigma_{y_j x_i}} - f_{\sigma_{y_j x_i}}) \text{ per } x_i > y_j$$

oppure

$$(x_i y_j - y_j x_i - [x_i, y_j]) = -(y_j x_i - x_i y_j + [x_i, y_j]) = -(W_{\sigma_{x_i y_j}} - f_{\sigma_{x_i y_j}}) \text{ per } x_i < y_j.$$

Dunque

$$\text{per ogni monomio } C,\ l_C > 2 \text{ si ha } ab - ba - [a, b] \in I_C$$

in quanto (vedi definizione 2.20) $ab - ba - [a, b]$ è somma di termini del tipo $c(W_\sigma - f_\sigma)$ con $c \in k$ e $W_\sigma < C$ perchè $l_{W_\sigma} = 2 < l_C$.
Mostriamo ora che ogni ambiguità di S è risolubile relativamente a $<$. Si ha un'ambiguità quando all'interno di un monomio vi è un termine del tipo $\ldots zyx \ldots$ con $z > y > x$, ossia le ambiguità sono tutte e sole le quintuple $(\sigma_{zy}, \sigma_{yx}, z, y, x)$ al variare di $x, y, z \in \langle X\rangle$ con $z > y > x$. Vale allora

$$r_{1\sigma_{zy}x}(zyx) - r_{z\sigma_{yx}1}(zyx) = f_{\sigma_{zy}} x - z f_{\sigma_{yx}} = (yzx - [y, z]x) - (zxy - z[x, y]).$$

Per ridurre il termine yzx applichiamo dapprima $r_{y\sigma_{zx}1}$ e quindi $r_{1\sigma_{yx}z}$ ottenendo

$$yzx - [y, z]x \overset{r_{y\sigma_{zx}1}}{\mapsto} yxz - y[x, z] - [y, z]x \overset{r_{1\sigma_{yx}z}}{\mapsto} xyz - [x, y]z - y[x, z] - [y, z]x.$$

Visto che $yzx < zyx$, si ha $y(zx - (xz - [x, z])) \in I_{zyx}$ secondo la definizione 2.20 e quindi applicare la riduzione $r_{y\sigma_{zx}1}$ equivale effettivamente a sottrarre

l'elemento $yzx - yxz + y[x,z]$ di I_{zyx}, analogamente $yxz < zyx$ dunque $(yx - (xy - [x,y]))z \in I_{zyx}$ e applicare $r_{1\sigma_{yx}z}$ equivale a sottrarre l'elemento $yxz - xyz + [x,y]z$ di I_{zyx}. Complessivamente vale

$$r_{1\sigma_{zy}x}(zyx) = xyz - [x,y]z - y[x,z] - [y,z]x + i$$

con $i \in I_{zyx}$. In maniera analoga, per ridurre zxy applichiamo dapprima $r_{1\sigma_{zx}y}$ e quindi $r_{x\sigma_{zy}1}$ ottenendo

$$zxy - z[x,y] \overset{r_{1\sigma_{zx}y}}{\mapsto} xzy - [x,z]y - z[x,y] \overset{r_{x\sigma_{zy}1}}{\mapsto} xyz - x[y,z] - [x,z]y - z[x,y].$$

Dato che $zxy < zyx$, applicare la riduzione $r_{1\sigma_{zx}y}$ equivale a sotrarre l'elemento $zxy - xzy + [x,z]y$ di I_{zyx}, mentre per $xzy < zyx$ applicare la riduzione $r_{x\sigma_{zy}1}$ equivale a sottrarre l'elemento $xzy - xyz + x[y,z]$ di I_{zyx}. Complessivamente vale

$$r_{z\sigma_{yx}1}(zyx) = xyz - x[y,z] - [x,z]y - z[x,y] + i'$$

con $i' \in I_{zyx}$. Dunque per un qualche $j \in I_{zyx}$ si ha

$$r_{1\sigma_{zy}x}(zyx) - r_{z\sigma_{yx}1}(zyx) =$$

$$= -[x,y]z - y[x,z] - [y,z]x + x[y,z] + [x,z]y + z[x,y] + j =$$

$$= (x[y,z] - [y,z]x) + ([x,z]y - y[x,z]) + (z[x,y] - [x,y]z) + j.$$

Per quanto visto sopra, $x[y,z] - [y,z]x = [x,[y,z]] + j_1$, $[x,z]y - y[x,z] = [[x,z],y] + j_2$, $z[x,y] - [x,y]z = [z,[x,y]] + j_3$ con $j_1, j_2, j_3 \in I_{zyx}$ in quanto $l_{zyx} > 2$. In conclusione si ha, per un certo $j' \in I_{zyx}$,

$$r_{1\sigma_{zy}x}(zyx) - r_{z\sigma_{yx}1}(zyx) = [x,[y,z]] + [[x,z],y] + [z,[x,y]] + j' = 0 + j'$$

per l'identità di Jacobi, dunque $f_{\sigma_{zy}}x - zf_{\sigma_{yx}} \in I_{zyx}$, cioè ogni ambiguità è risolubile relativamente a $\leq$ e quindi per il Diamond Lemma (2) $\Rightarrow$ (4) si ha la tesi.
Infatti la definizione del sistema di riduzione S fà sì che l'ideale $\langle W_\sigma - f_\sigma \rangle_{\sigma \in S}$ della dimostrazione coincida con l'ideale I che dà $T(\mathfrak{g})/I = U(\mathfrak{g})$, grazie anche alla proposizione precedente. Dunque per la condizione (4) un sistema di rappresentanti in $k\langle X \rangle = T(\mathfrak{g})$ per gli elementi dell'algebra $T(\mathfrak{g})/I = U(\mathfrak{g})$ è dato dal sottomodulo di $T(\mathfrak{g})$ generato dai monomi irriducibili secondo S, che per costruzione del sistema di riduzione risultano essere tutti e soli i monomi ordinati $x'_1 \dots x'_n$ tali che $x_1 \leq \dots \leq x_n$. □

Corollario 3.4. *La mappa $\tau : \mathfrak{g} \longrightarrow U(\mathfrak{g})$ è iniettiva, cioè $\mathfrak{g} \hookrightarrow U(\mathfrak{g})$.*

Infatti l'immagine di una base è formata da elementi linearmente indipendenti, quindi il morfismo τ è iniettivo.

3.2 Il Teorema PBW sui campi

In questa sezione k indicherà un campo con *char* $k \neq 2, 3$.
Vogliamo dare una dimostrazione del teorema di Poincaré-Birkhoff-Witt sui campi senza far uso del Diamond Lemma.

Teorema 3.5 (Teorema PBW sui campi)**.** *Sia* $\mathfrak{g}$ *un'algebra di Lie con base linearmente ordinata* $\{x_1, \ldots, x_n, \ldots\}$*. Allora* $U(\mathfrak{g})$ *ha base*

$$1 \cup \{\psi(x_{i_1}) \ldots \psi(x_{i_s})\}$$

al variare di $1 \leq i_1 \leq \ldots \leq i_s$ *dove* $\psi : \mathfrak{g} \longrightarrow U(\mathfrak{g})$ *è la mappa* τ *della definizione 1.23.*

Dimostrazione. Definiamo, per ogni $p \in \mathbb{N}$, il sottomodulo di $T(\mathfrak{g})$ formato da monomi ordinati di grado p:

$$T(\mathfrak{g})^0_p := span\{x_{i_1} \otimes \ldots \otimes x_{i_p} | 1 \leq i_1 \leq \ldots \leq i_p\}.$$

Sia anche $T(\mathfrak{g})^0 = \bigoplus_{p \geq 0} T(\mathfrak{g})^0_p$. La tesi è conseguenza dell'affermazione

$$T(\mathfrak{g}) = T(\mathfrak{g})^0 \oplus I$$

dove $I = \langle x \otimes y - y \otimes x - [x, y] | x, y \in \mathfrak{g} \rangle$ è l'ideale bilatero di $T(\mathfrak{g})$ della definizione 1.23. In particolare $T(\mathfrak{g}) = T(\mathfrak{g})^0 + I$ implica che i monomi ordinati siano generatori di $U(\mathfrak{g})$, mentre $T(\mathfrak{g})^0 \cap I = \{0\}$ assicura che tali generatori siano linearmente indipendenti.
Definiamo ora:

$$T(\mathfrak{g})^d_p := span\{t = x_{i_1} \otimes \ldots \otimes x_{i_p} | ind(t) = d\}$$

dove $ind(t)$ è il numero delle coppie *disordinate* (r, s) tali che $1 \leq r < s \leq p$ ma $i_r > i_s$ ed è in effetti una misura del disordine di t: infatti in tal caso $x_{i_r} > x_{i_s}$ ma x_{i_r} viene più a sinistra di x_{i_s} in t. Per questo chiamiamo tali coppie (r, s) *coppie disordinate*. Si noti inoltre che questa definizione di $T(\mathfrak{g})^d_p$ si accorda con quella di $T(\mathfrak{g})^0_p$, che è in effetti l'insieme dei tensori di grado p aventi indice di disordine nullo. Sia inoltre

$$T(\mathfrak{g})_p := \bigoplus_{d \geq 0} T(\mathfrak{g})^d_p$$

l'insieme dei tensori di grado p.
Per dimostrare che $I + T(\mathfrak{g})^0 = T(\mathfrak{g})$, è sufficiente mostrare che

$$T(\mathfrak{g})_r \subseteq I + \sum_{q=0}^{r} T(\mathfrak{g})^0_q$$

per ogni $r \geq 0$. Dimostriamolo per induzione su r. Per $r = 0$ e $r = 1$ è ovvio, perchè ogni monomio è già ordinato. Assumiamo dunque che il risultato sia vero per $r-1$ e proviamolo per r.
E' sufficiente mostrare che

$$T(\mathfrak{g})_r^d \subseteq I + \bigoplus_{q=0}^{r} T(\mathfrak{g})_q^0$$

per ogni $d \geq 0$, in quanto ogni monomio di $T(\mathfrak{g})_r$ sta in $T(\mathfrak{g})_r^d$ per un certo d: mostriamo anche questo per induzione, stavolta su d. Per $d = 0$ è ovviamente vero, assumiamo dunque che $T(\mathfrak{g})_r^e \subseteq I + \bigoplus_{q=0}^{r} T(\mathfrak{g})_q^0$ per ogni e tale che $0 \leq e \leq d-1$.
Sia ora $t = x_{i_1} \otimes \ldots \otimes x_{i_r} \in T(\mathfrak{g})_r^d$, dato che $d \geq 1$ deve esistere una coppia disordinata ed inoltre è possibile trovare una *coppia disordinata consecutiva*, cioè del tipo $(u, u+1)$ con $i_u > i_{u+1}$, ossia una coppia $(u, u+1)$ tale che $x_{i_u} > x_{i_{u+1}}$, infatti se tale coppia consecutiva non esistesse il monomio sarebbe già ordinato. Consideriamo dunque $t' := x_{i_1} \otimes \ldots \otimes x_{i_{u+1}} \otimes x_{i_u} \otimes \ldots \otimes x_{i_r}$ dove $t' \in T(\mathfrak{g})_r^{d-1}$ poichè abbiamo ridotto di uno il numero delle coppie disordinate. Si ha dunque:

$$t - t' = x_{i_1} \otimes \ldots \otimes x_{i_{u-1}} \otimes (x_{i_u} \otimes x_{i_{u+1}} - x_{i_{u+1}} \otimes x_{i_u}) \otimes x_{i_{u+2}} \otimes \ldots \otimes x_{i_r}.$$

Dato $a := x_{i_u} \otimes x_{i_{u+1}} - x_{i_{u+1}} \otimes x_{i_u} - [x_{i_u}, x_{i_{u+1}}] \in I$, vale perciò

$$t - t' = x_{i_1} \otimes \ldots \otimes x_{i_{u-1}} \otimes [x_{i_u}, x_{i_{u+1}}] \otimes x_{i_{u+2}} \otimes \ldots \otimes x_{i_r} +$$

$$+ x_{i_1} \otimes \ldots \otimes x_{i_{u-1}} \otimes a \otimes x_{i_{u+2}} \otimes \ldots \otimes x_{i_r} \in T(\mathfrak{g})_{r-1} + I$$

dunque

$$t = t - t' + t' \in I + T(\mathfrak{g})_{r-1} + T(\mathfrak{g})_r^{d-1} \subseteq I + T(\mathfrak{g})_{r-1} + (I + \bigoplus_{q=0}^{r} T(\mathfrak{g})_q^0)$$

per ipotesi induttiva su d, ed anche

$$t \in I + T(\mathfrak{g})_{r-1} + \bigoplus_{q=0}^{r} T(\mathfrak{g})_q^0 \subseteq I + (I + \sum_{q=0}^{r} T(\mathfrak{g})_q^0) + \bigoplus_{q=0}^{r} T(\mathfrak{g})_q^0$$

per ipotesi induttiva su r, quindi

$$t \in I + \sum_{q=0}^{r} T(\mathfrak{g})_q^0 = I + \bigoplus_{q=0}^{r} T(\mathfrak{g})_q^0$$

in quanto la somma di tensori ordinati di gradi diversi è ovviamente diretta; abbiamo provato la tesi per d, dunque per doppia induzione segue la tesi.
Proviamo ora che $I \cap T(\mathfrak{g})^0 = 0$. Per fare ciò costruiamo un endomorfismo lineare $L : T(\mathfrak{g}) \longrightarrow T(\mathfrak{g})$ tale che $L|_{T(\mathfrak{g})^0} = id$ e $L(I) = 0$, da cui seguirà ovviamente la tesi.
Vogliamo definire ricorsivamente sul grado p dei tensori un endomorfismo L tale che:

1. $L(t) = t$ per ogni tensore ordinato $t \in T(\mathfrak{g})^0$;

2. se $p \geq 2$ e $(s, s+1)$ è una coppia disordinata consecutiva tale che $x_{i_s} > x_{i_{s+1}}$, per $t = x_{i_1} \otimes \ldots \otimes x_{i_p} \in T(\mathfrak{g})^d_p$, $d \geq 2$ valga

$$L(x_{i_1} \otimes \ldots \otimes x_{i_s} \otimes x_{i_{s+1}} \otimes \ldots \otimes x_{i_p}) =$$

$$= L(x_{i_1} \otimes \ldots \otimes x_{i_{s+1}} \otimes x_{i_s} \otimes \ldots \otimes x_{i_p}) + L(x_{i_1} \otimes \ldots \otimes [x_{i_s}, x_{i_{s+1}}] \otimes \ldots \otimes x_{i_p}).$$

Una volta trovato L ed estesolo per linearità a tutto $T(\mathfrak{g})$ è chiaro che avrà le proprietà richieste. Ad esempio verificiamo che si annulla su I. Per la linearità di L, basta verificare che si annulla sugli elementi del tipo $t_1 \otimes (x_i \otimes x_j - x_j \otimes x_i - [x_i, x_j]) \otimes t_2$ con t_1, t_2 tensori prodotto dei soli monomi x_k della base e x_i, x_j elementi della base di $\mathfrak{g}$. Si hanno allora tre possibili casi:

- $x_i = x_j$, allora $x_i \otimes x_j - x_j \otimes x_i - [x_i, x_j] = 0 - 0 + 0$ per la proprietà (1) della definizione di bracket, dunque $L(0) = 0$;

- $x_i < x_j$, allora applicando la definizione

$$L(t_1 \otimes x_j \otimes x_i \otimes t_2) = L(t_1 \otimes x_i \otimes x_j \otimes t_2) + L(t_1 \otimes [x_j, x_i] \otimes t_2) =$$

$$= L(t_1 \otimes x_i \otimes x_j \otimes t_2) + L(t_1 \otimes -[x_i, x_j] \otimes t_2)$$

  e nuovamente

$$L(t_1 \otimes (x_i \otimes x_j - x_j \otimes x_i - [x_i, x_j]) \otimes t_2) = 0;$$

- $x_i > x_j$, allora applicando la definizione

$$L(t_1 \otimes x_i \otimes x_j \otimes t_2) = L(t_1 \otimes x_j \otimes x_i \otimes t_2) + L(t_1 \otimes [x_i, x_j] \otimes t_2)$$

  e nuovamente

$$L(t_1 \otimes (x_i \otimes x_j - x_j \otimes x_i - [x_i, x_j]) \otimes t_2) = 0.$$

L si annulla in ognuno dei casi sopracitati, dunque si annulla su I, cioè $L(I) = 0$.
Definiamo L come l'identità su $T(\mathfrak{g})_0$ e $T(\mathfrak{g})_1$ e fissato $p \geq 2$ supponiamo che L sia un endomorfismo di $\sum_{0\leq q\leq p-1} T(\mathfrak{g})_q$. Vogliamo estendere ricorsivamente L ai tensori t di grado p in modo che le condizioni (1) e (2) siano soddisfatte e lo facciamo per induzione su $d = ind(t)$. Per $d = 0$ poniamo $L(t) = t$, per ipotesi induttiva supponiamo di avere definito L che soddisfa le condizioni (1), (2) per ogni monomio di grado p e indice $\leq d-1$, sia dunque $t = x_{i_1} \otimes \ldots \otimes x_{i_p} \in T(\mathfrak{g})^d_p$.
Scegliamo ora una coppia disordinata consecutiva $(r, r+1)$ tale che $x_{i_r} > x_{i_{r+1}}$ e poniamo $L(t) := L(x_{i_1} \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + L(x_{i_1} \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p})$. Una volta provato che L è ben definito, cioè che la sua definizione non dipende dalla scelta della coppia disordinata consecutiva (che non è in generale unica), visto che $x_{i_1} \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p} \in T(\mathfrak{g})^{d-1}_p$ e $x_{i_1} \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p} \in T(\mathfrak{g})_{p-1}$, per induzione su d avremo provato che L è ben definito su ogni tensore di grado p e quindi avremo definito L ricorsivamente sull'intero $T(\mathfrak{g})$, come volevamo.
Supponiamo dunque di avere due coppie disordinate consecutive $(l, l+1)$ e $(r, r+1)$, con $l < r \leq p-1$, possono presentarsi due casi:

- $(l, l+1)$ e $(r, r+1)$ sono disgiunte, cioè $l + 1 < r$ e si ha

$$t = x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes x_{i_{l+1}} \otimes \ldots \otimes x_{i_r} \otimes x_{i_{r+1}} \otimes \ldots \otimes x_{i_p}.$$

  Definiamo allora:

$$u = L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes \ldots \otimes x_{i_p}) + L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_p}),$$

$$v = L(x_{i_1} \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + L(x_{i_1} \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p})$$

  le due espressioni che si ottengono scegliendo per prima $(l, l+1)$ oppure $(r, r+1)$ come coppia disordinata. Si noti che i termini a cui applichiamo L fanno parte di $\sum_{0\leq e\leq d-1} T(\mathfrak{g})^e_p + \sum_{0\leq q\leq p-1} T(\mathfrak{g})_q$ e dunque per ipotesi induttiva L è ben definito su di essi. Applicando la definizione

di L ricorsivamente otteniamo rispettivamente:

$$\begin{aligned} u &= L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes \ldots \otimes x_{i_r} \otimes x_{i_{r+1}} \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_r} \otimes x_{i_{r+1}} \otimes \ldots \otimes x_{i_p}) = \\ &= L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p}) \end{aligned}$$

ed analogamente

$$\begin{aligned} v &= L(x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes x_{i_{l+1}} \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes x_{i_{l+1}} \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p}) = \\ &= L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_{r+1}} \otimes x_{i_r} \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p}) + \\ &+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes [x_{i_r}, x_{i_{r+1}}] \otimes \ldots \otimes x_{i_p}). \end{aligned}$$

Dunque $u = v$ e questo prova che in questo caso L è ben definito.

- $(l, l+1)$ e $(r, r+1)$ non sono disgiunte, ossia $r = l+1$ e $t = x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes x_{i_{l+1}} \otimes x_{i_{r+1}} \otimes \ldots \otimes x_{i_p}$. Per comodità di notazione, scriviamo $x_{i_{r+1}} = x_{i_{l+2}}$, vale dunque $i_l > i_{l+1} > i_{l+2}$, cioè $x_{i_l} > x_{i_{l+1}} > x_{i_{l+2}}$.
  Per le due espressioni

$$\begin{aligned} u = L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes x_{i_l} \otimes x_{i_{l+2}} \otimes \ldots \otimes x_{i_p}) + \\ + L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes x_{i_{l+2}} \otimes \ldots \otimes x_{i_p}), \\ v = L(x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes x_{i_{l+2}} \otimes x_{i_{l+1}} \otimes \ldots \otimes x_{i_p}) + \\ + L(x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p}) \end{aligned}$$

  che si ottengono scegliendo per prima $(l, l+1)$ oppure $(l+1, l+2)$ come coppia disordinata, si ha che i termini a cui applichiamo L fanno parte

di $\sum_{0\leq q\leq p-1} T(\mathfrak{g})_q + \sum_{0\leq e\leq d-1} T(\mathfrak{g})_p^e$ e dunque per ipotesi induttiva L è ben definito su di essi. Applicando ripetutamente la definizione otteniamo

$$\begin{aligned}
u &= L(x_{i_1}\otimes\ldots\otimes x_{i_{l+1}}\otimes x_{i_l}\otimes x_{i_{l+2}}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes [x_{i_l},x_{i_{l+1}}]\otimes x_{i_{l+2}}\otimes\ldots\otimes x_{i_p}) \;=\\
&= L(x_{i_1}\otimes\ldots\otimes x_{i_{l+1}}\otimes x_{i_{l+2}}\otimes x_{i_l}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes x_{i_{l+1}}\otimes [x_{i_l},x_{i_{l+2}}]\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes [x_{i_l},x_{i_{l+1}}]\otimes x_{i_{l+2}}\otimes\ldots\otimes x_{i_p}) \;=\\
&= L(x_{i_1}\otimes\ldots\otimes x_{i_{l+2}}\otimes x_{i_{l+1}}\otimes x_{i_l}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes [x_{i_{l+1}},x_{i_{l+2}}]\otimes x_{i_l}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes x_{i_{l+1}}\otimes [x_{i_l},x_{i_{l+2}}]\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes [x_{i_l},x_{i_{l+1}}]\otimes x_{i_{l+2}}\otimes\ldots\otimes x_{i_p})
\end{aligned}$$

ed analogamente

$$\begin{aligned}
v &= L(x_{i_1}\otimes\ldots\otimes x_{i_l}\otimes x_{i_{l+2}}\otimes x_{i_{l+1}}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes x_{i_l}\otimes [x_{i_{l+1}},x_{i_{l+2}}]\otimes\ldots\otimes x_{i_p}) \;=\\
&= L(x_{i_1}\otimes\ldots\otimes x_{i_{l+2}}\otimes x_{i_l}\otimes x_{i_{l+1}}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes [x_{i_l},x_{i_{l+2}}]\otimes x_{i_{l+1}}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes x_{i_l}\otimes [x_{i_{l+1}},x_{i_{l+2}}]\otimes\ldots\otimes x_{i_p}) \;=\\
&= L(x_{i_1}\otimes\ldots\otimes x_{i_{l+2}}\otimes x_{i_{l+1}}\otimes x_{i_l}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes x_{i_{l+2}}\otimes [x_{i_l},x_{i_{l+1}}]\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes [x_{i_l},x_{i_{l+2}}]\otimes x_{i_{l+1}}\otimes\ldots\otimes x_{i_p}) \;+\\
&+ L(x_{i_1}\otimes\ldots\otimes x_{i_l}\otimes [x_{i_{l+1}},x_{i_{l+2}}]\otimes\ldots\otimes x_{i_p}).
\end{aligned}$$

Ora si ha

$$\begin{aligned}
u - v &= L(x_{i_1} \otimes \ldots \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes x_{i_l} \otimes \ldots \otimes x_{i_p}) \quad + \\
&+ L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes [x_{i_l}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p}) \quad + \\
&+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes x_{i_{l+2}} \otimes \ldots \otimes x_{i_p}) \quad - \\
&- L(x_{i_1} \otimes \ldots \otimes x_{i_{l+2}} \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_p}) \quad - \\
&- L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+2}}] \otimes x_{i_{l+1}} \otimes \ldots \otimes x_{i_p}) \quad - \\
&- L(x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p}) \quad = \\
&= L(x_{i_1} \otimes \ldots \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes x_{i_l} \otimes \ldots \otimes x_{i_p} \quad - \\
&- x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p}) \quad + \\
&+ L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes [x_{i_l}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p} \quad - \\
&- x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+2}}] \otimes x_{i_{l+1}} \otimes \ldots \otimes x_{i_p}) \quad + \\
&+ L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes x_{i_{l+2}} \otimes \ldots \otimes x_{i_p} \quad - \\
&- x_{i_1} \otimes \ldots \otimes x_{i_{l+2}} \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_p}).
\end{aligned}$$

Ognuno dei termini a cui applichiamo L fa parte di $T(\mathfrak{g})_{p-1}$ sul quale L è endomorfismo per ipotesi induttiva, dunque L si annulla su I. Pertanto:

$$\begin{aligned}
& L(x_{i_1} \otimes \ldots \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes x_{i_l} \otimes \ldots \otimes x_{i_p} \quad - \\
- \; & x_{i_1} \otimes \ldots \otimes x_{i_l} \otimes [x_{i_{l+1}}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p}) \quad = \\
= \; & L(x_{i_1} \otimes \ldots \otimes [[x_{i_{l+1}}, x_{i_{l+2}}], x_{i_l}] \otimes \ldots \otimes x_{i_p}),
\end{aligned}$$

$$\begin{aligned}
& L(x_{i_1} \otimes \ldots \otimes x_{i_{l+1}} \otimes [x_{i_l}, x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p} \quad - \\
- \; & x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+2}}] \otimes x_{i_{l+1}} \otimes \ldots \otimes x_{i_p}) \quad = \\
= \; & L(x_{i_1} \otimes \ldots \otimes [x_{i_{l+1}}, [x_{i_l}, x_{i_{l+2}}]] \otimes \ldots \otimes x_{i_p}),
\end{aligned}$$

$$
\begin{aligned}
& L(x_{i_1} \otimes \ldots \otimes [x_{i_l}, x_{i_{l+1}}] \otimes x_{i_{l+2}} \otimes \ldots \otimes x_{i_p} \quad - \\
- \quad & x_{i_1} \otimes \ldots \otimes x_{i_{l+2}} \otimes [x_{i_l}, x_{i_{l+1}}] \otimes \ldots \otimes x_{i_p}) \quad = \\
= \quad & L(x_{i_1} \otimes \ldots \otimes [[x_{i_l}, x_{i_{l+1}}], x_{i_{l+2}}] \otimes \ldots \otimes x_{i_p}).
\end{aligned}
$$

In conclusione,

$$
\begin{aligned}
u - v \quad = \quad & L(x_{i_1} \otimes \ldots \otimes ([[x_{i_{l+1}}, x_{i_{l+2}}], x_{i_l}] + [x_{i_{l+1}}, [x_{i_l}, x_{i_{l+2}}]] + \\
& + [[x_{i_l}, x_{i_{l+1}}], x_{i_{l+2}}]) \otimes \ldots \otimes x_{i_p}) \quad = \\
= \quad & L(x_{i_1} \otimes \ldots \otimes ([x_{i_l}, [x_{i_{l+2}}, x_{i_{l+1}}]] + [x_{i_{l+1}}, [x_{i_l}, x_{i_{l+2}}]] + \\
& + [x_{i_{l+2}}, [x_{i_{l+1}}, x_{i_l}]]) \otimes \ldots \otimes x_{i_p}) \quad = \\
= \quad & L(x_{i_1} \otimes \ldots \otimes 0 \otimes \ldots \otimes x_{i_p}) \quad = \\
= \quad & 0
\end{aligned}
$$

per l'identità di Jacobi, quindi L è ben definito anche in questo caso e ciò completa la dimostrazione.

□

Corollario 3.6. *La mappa $\psi : \mathfrak{g} \longrightarrow U(\mathfrak{g})$ è iniettiva, cioè $\mathfrak{g} \hookrightarrow U(\mathfrak{g})$.*

Infatti l'immagine di una base è formata da elementi linearmente indipendenti, quindi il morfismo ψ è iniettivo.

Capitolo 4

Applicazioni

In questo capitolo spiegheremo l'importanza della generalizzazione agli anelli delle costruzioni che abbiamo esaminato ed in particolare del teorema di Poincaré-Birkhoff-Witt. Vogliamo in particolare spiegare perchè è necessario, nell'ambito della geometria algebrica, considerare anche algebre di Lie definite su anelli.
Non scenderemo nei dettagli tecnici delle dimostrazioni e delle verifiche, in quanto ci porterebbe lontano dagli scopi della tesi. Quello che ci interessa è l'idea generale di queste costruzioni e la loro motivazione.

4.1 Gruppi Algebrici Lineari

Sia k un campo, con *char* $k \neq 2, 3$ e consideriamo r equazioni polinomiali a coefficienti in k

$$\begin{cases} f_1(x_1, \ldots, x_m) = 0 \\ \qquad \ldots \\ f_r(x_1, \ldots, x_m) = 0. \end{cases}$$

Sia

$$G = \{(x_1, \ldots, x_m) \in k^m | f_i(x_1, \ldots, x_m) = 0 \text{ per ogni } 1 \leq i \leq r\}$$

l'insieme degli zeri dei polinomi $f_1, \ldots, f_r$ in k, e supponiamo di avere una funzione polinomiale $G \times G \xrightarrow{*} G$ che rende $(G, *)$ un gruppo; in altre parole abbiamo definito una struttura di gruppo sull'insieme algebrico G.

Definizione 4.1. Sia

$$I := \langle f_1, \ldots, f_r \rangle \subseteq k[x_1, \ldots, x_m]$$

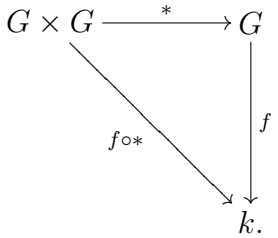
l'ideale generato dai polinomi $f_1, \ldots, f_r$. Si dice *anello delle coordinate* oppure *algebra affine di* G il quoziente

$$k[G] := k[x_1, \ldots, x_m]/I.$$

$k[G]$ può essere interpretata come l'algebra delle funzioni regolari su G, cioè funzioni $f : G \longrightarrow k$ che possono essere scritte come quozienti di polinomi con denominatore mai nullo su G, ossia

$$k[G] := \{f : G \longrightarrow k | f = \frac{g(x_1, \ldots, x_m)}{h(x_1, \ldots, x_m)} \text{ con } h(x_1, \ldots, x_m) \neq 0 \text{ su } G\}.$$

Definizione 4.2. Il fatto che l'insieme algebrico G abbia una struttura di gruppo ci permette di definire un'operazione Δ, duale della moltiplicazione, detta *comoltiplicazione*. Definiamo $\Delta : k[G] \longrightarrow k[G \times G] \cong k[G] \otimes k[G]$ nel modo seguente: sia $f \in k[G]$, allora $\Delta(f) := f \circ *$ come risulta dal diagramma



Siano anche $\epsilon : k[G] \longrightarrow k$, $f \mapsto f(1_G)$ e $\delta : k[G] \longrightarrow k[G]$, $f \mapsto f \circ i$ con $i : G \longrightarrow G$, $i(g) = g^{-1}$. $k[G]$ munita della comoltiplicazione Δ, counità ϵ e antipodo δ che soddisfano i diagrammi commutativi duali rispetto alla moltiplicazione, unità e inverso in G, si dice *algebra di Hopf*.

Vogliamo ora cercare gli zeri dei polinomi $f_1, \ldots, f_r$ su anelli diversi da k.

Definizione 4.3. Sia R una k-algebra, definiamo

$$G(R) := \{(x_1, \ldots, x_m) \in R^m | f_i(x_1, \ldots, x_m) = 0 \text{ per ogni } 1 \leq i \leq r\} \quad (4.1)$$

ossia $G(R)$ consiste negli zeri dei polinomi $f_1, \ldots, f_r$ in R, e non in k. Si noti che essendo R una k-algebra, esiste un'immersione naturale $k \hookrightarrow R$, dunque le funzioni polinomiali sono effettivamente ben definite anche in R.

Tramite le mappe dell'algebra di Hopf possiamo indurre una struttura di gruppo su $G(R)$.
E' possibile descrivere $G(R)$ in maniera equivalente come insieme di morfismi, nel seguente modo:

$$G(R) = \{\phi : k[x_1, \ldots, x_m]/\langle f_1, \ldots, f_r \rangle \longrightarrow R\},$$

sfruttando la corrispondenza biunivoca

$$\begin{array}{lcl} (r_1,\dots,r_m)\in R^m & \leftrightarrow & \phi : k[x_1,\dots,x_m]/\langle f_i\rangle_{i=1\dots r} \longrightarrow R \\ f_i(x_1,\dots,x_m)=0 & & x_i \mapsto r_i. \end{array}$$

Infatti se $(x_1,\dots,x_m)\in R^m$ annulla ogni polinomio f_i possiamo associargli

$$\begin{array}{rcl} \phi' : k[x_1,\dots,x_m] & \longrightarrow & R \\ x_i & \mapsto & r_i. \end{array}$$

Questo morfismo passa al quoziente in quanto $f_i(x_1,\dots,x_m)=f_i(r_1,\dots,r_m)=0$ per ogni i e dunque induce

$$\phi : k[x_1,\dots,x_m]/I \longrightarrow R.$$

Viceversa dato

$$\phi : k[x_1,\dots,x_m]/I \longrightarrow R,$$

siano $r_j := \phi(x_j)$, allora

$$(r_1,\dots,r_m)$$

annulla ogni polinomio f_i ed è l'elemento di $G(R)$ corrispondente.
Definiamo ora l'operazione di gruppo $\star$ su $G(R)$ nel seguente modo. Siano $\phi,\psi : k[G]\longrightarrow R$, si ha allora la mappa

$$k[G]\xrightarrow{\Delta} k[G]\otimes k[G]\xrightarrow{\phi\otimes\psi} R\otimes R\xrightarrow{m} R$$

dove m è la moltiplicazione dell'algebra R.
Poniamo quindi

$$\phi\star\psi := (m)\circ(\phi\otimes\psi)\circ(\Delta)$$

la mappa descritta sopra. L'inverso risulta essere

$$\phi^{-1} : k[G]\xrightarrow{\delta} k[G]\xrightarrow{\phi} R$$

e l'unità

$$\varepsilon : k[G]\xrightarrow{\epsilon} k\hookrightarrow R.$$

Abbiamo allora definito un funtore

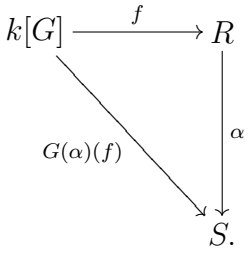
$$G:\ (k\text{-}alg)\ \longrightarrow\ (Gr)$$

tra gli oggetti della categoria delle k-algebre agli oggetti quella dei gruppi. Sui morfismi, G è definito in modo naturale come è riassunto dai diagrammi

seguenti:

$$\begin{array}{ccc} R & \xrightarrow{G} & G(R) \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle G(\alpha)} \\ S & \xrightarrow{G} & G(S). \end{array} \tag{4.2}$$

dove $G(\alpha)$ agisce come da diagramma:



Abbiamo dunque visto che a partire da un insieme algebrico in k^m, cioè ottenuto come zeri di polinomi $f_1, \ldots, f_r$ in k^m, è possibile ottenere un funtore dalla categoria delle k-algebre alla categoria dei gruppi:

$$G(R) = \mathrm{Hom}(k[G], R), \qquad G(\alpha)(f) = \alpha \circ f.$$

Funtori del tipo $\mathrm{Hom}(k[G], -)$ si dicono *rappresentabili* e hanno un'importanza fondamentale nella teoria dei gruppi algebrici. E' interessante notare che l'informazione contenuta nel gruppo G e nel funtore ad esso associato (comunemente indicato con la stessa lettera) è completamente equivalente a $k[G]$.

4.2 Esempi

In questa sezione vogliamo esaminare alcuni esempi.
Consideriamo il gruppo speciale lineare sul campo k:

$$SL_n(k) = \{A \in M_n(k) | detA = 1\}.$$

L'equazione $detA = 1$ è un'equazione polinomiale di grado n nelle n^2 variabili a_{ij}. Dunque

$$k[SL_n] = k[a_{ij}]/\langle det(a_{ij}) - 1\rangle.$$

L'operazione di gruppo è rappresentata dalla moltiplicazione fra matrici

$$SL_n(k) \times SL_n(k) \longrightarrow SL_n(k)$$

che può anche essere descritta in termini di polinomi: ognuna delle n^2 componenti di questa funzione è in effetti una funzione quadratica di $2n^2$ variabili. Data una k-algebra R, possiamo allora considerare $SL_n(R)$ gruppo delle matrici a coefficienti in R con determinante 1. Dunque

$$SL_n : \ (k\text{-}alg) \ \longrightarrow \ (Gr), \qquad R \mapsto SL_n(R)$$

è un funtore in gruppi come descritto nella sezione precedente.

Consideriamo ora un altro esempio molto importante: il gruppo generale lineare

$$GL_n(k) := \{A \in M_n(k) | det(A) \neq 0\}.$$

Evidentemente questa non è un'equazione, possiamo però pensare $GL_n(k) \subseteq SL_{n+1}(k)$ considerando matrici del tipo

$$\begin{pmatrix} a_{11} & \dots & a_{n1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & 0 \\ 0 & \dots & 0 & t \end{pmatrix} \in SL_{n+1}(k)$$

e definendo il gruppo $GL_n(k)$ entro $SL_{n+1}(k)$ tramite l'equazione

$$det(A)t = 1 \text{ in } k[x_{11}, \dots, x_{nn}, t].$$

Dunque

$$k[GL_n] = k[x_{ij}, t]/\langle det(x_{ij})t = 1\rangle.$$

Nello stesso modo, possiamo dare una struttura di gruppo algebrico anche a $GL_n(R)$ ottenendo nuovamente un funtore di gruppi.

Vediamo un ultimo esempio. La circonferenza

$$S^1 = \{(x, y) \in \mathbb{R}^2 \text{ tale che } x^2 + y^2 = 1\}$$

ha struttura di gruppo algebrico con la moltiplicazione fra numeri complessi

$$(x, y) \cdot (x', y') = (xx' - yy', xy' + x'y).$$

Tuttavia, questa può anche essere vista come sottogruppo di $SL_2(\mathbb{R})$ formato dalle matrici del tipo

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Così possiamo nuovamente ampliare il campo in cui consideriamo le equazioni e considerare il *gruppo di rotazione* su R, ove R è una $\mathbb{R}$-algebra, formato dalle matrici del tipo

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in SL_2(R).$$

4.3 Algebre di Lie su di un gruppo algebrico

Riprendiamo ora la costruzione funtoriale di $G(R)$ ed in particolare il diagramma (4.2). Vogliamo associare ad ogni funtore rappresentabile dalle k-algebre ai gruppi un funtore dalle k-algebre a valore nelle algebre di Lie sugli anelli

Definizione 4.4. Consideriamo dunque le due k-algebre R e

$$R(\varepsilon) := R[\varepsilon]/(\varepsilon^2)$$

e siano anche $i : R \hookrightarrow R(\varepsilon)$, $p : R(\varepsilon) \longrightarrow R$ con $a + b\varepsilon \xrightarrow{p} a$. Funtorialmente è allora definito il morfismo di gruppi $G(p)$:

$$\begin{array}{ccc} R(\varepsilon) & \xrightarrow{\quad G \quad} & G(R(\varepsilon)) \\ \Big\downarrow{\scriptstyle p} & & \Big\downarrow{\scriptstyle G(p)} \\ R & \xrightarrow{\quad G \quad} & G(R). \end{array}$$

Poniamo ora

$$\mathrm{Lie}(G)(R) := ker(G(p)) \triangleleft G(R(\varepsilon))$$

in quanto $ker(G(p))$ è sottogruppo normale di $G(R(\varepsilon))$.
Su $\mathrm{Lie}(G)(R)$ è possibile definire una struttura di R-algebra di Lie e identificarla con lo spazio tangente alla varietà algebrica G, tuttavia anzichè studiare questo procedimento da un punto di vista teorico, vedremo invece qualche esempio interessante.

Esempio 4.5. Consideriamo $G = GL_n$ il gruppo generale lineare e andiamo a calcolare $Lie(G)$. Si ha

$$\begin{array}{ccccc} G(p) & : & GL_n(R(\varepsilon)) & \longrightarrow & GL_n(R) \\ & & \begin{pmatrix} a_{11}+\varepsilon b_{11} & \dots & a_{1n}+\varepsilon b_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+\varepsilon b_{n1} & \dots & a_{nn}+\varepsilon b_{nn} \end{pmatrix} & \mapsto & \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}. \end{array}$$

Dunque

$$ker(G(p)) \cong \left\{ \begin{pmatrix} 1+\varepsilon b_{11} & \varepsilon b_{12} & \dots & \varepsilon b_{1n} \\ \varepsilon b_{21} & 1+\varepsilon b_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ \varepsilon b_{n1} & \dots & \dots & 1+\varepsilon b_{nn} \end{pmatrix} \right\}$$

e perciò possiamo identificare $ker(G(p))$ con le matrici $n \times n$ a coefficienti in R. Quindi

$$Lie(G)(R) = \mathfrak{M}_n(R)$$

dove $\mathfrak{M}_n(R)$ è un'algebra di Lie con bracket $[X, Y] = XY - YX$ come da esempio 1.4.

Esempio 4.6. Poniamo ora $G = SL_n$ e andiamo a costruire Lie$(SL_n)(R)$. Si ha allora

$$SL_n(R) = \{\text{matrici in } R \text{ con } det = 1\} = \{\text{morfismi } f : k[SL_n] \longrightarrow R\}$$

e

$$SL_n(R(\varepsilon)) = \{\text{matrici in } R(\varepsilon) \text{ con } det = 1\} = \{\text{morfismi } f : k[SL_n] \longrightarrow R(\varepsilon)\}$$

dove $k[SL_n] = k[x_{11}, \ldots, x_{nn}]/ < det = 1 >$ è l'algebra affine della definizione 4.1.
Dato $p : R(\varepsilon) \longrightarrow R$, vediamo come agisce $SL_n(p)$. Sia $A = (a_{ij} + \varepsilon b_{ij}) \in SL_n(R(\varepsilon))$, cioè A si può identificare con il morfismo $x_{ij} \mapsto a_{ij} + \varepsilon b_{ij}$ secondo la corrispondenza biunivoca descritta in precedenza, con la condizione $det(a_{ij} + \varepsilon b_{ij}) = 1$. Componendo con la mappa p, si ha

$$x_{ij} \overset{A}{\mapsto} a_{ij} + \varepsilon b_{ij} \overset{p}{\mapsto} a_{ij},$$

dunque $SL_n(p)(A)$ è il morfismo $x_{ij} \mapsto a_{ij}$, cioè la matrice $SL_n(p)(A) = (a_{ij}) \in SL_n(R)$.
Cerchiamo dunque il nucleo di $SL_n(p)$, ossia gli elementi $A \in SL_n(R(\varepsilon))$ che hanno per immagine la matrice identica $(a_{ij}) = (\delta_{ij}) \in SL_n(R)$ dove δ_{ij} è il delta di Kronecker. Dev'essere

$$A = \begin{pmatrix} 1 + \varepsilon b_{11} & \varepsilon b_{12} & \ldots & \varepsilon b_{1n} \\ \varepsilon b_{21} & 1 + \varepsilon b_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ \varepsilon b_{n1} & \ldots & \ldots & 1 + \varepsilon b_{nn} \end{pmatrix}.$$

Vediamo come si esplicita la condizione $det(A) = 1$. Visto che in $R(\varepsilon)$ vale $\varepsilon^2 = 0$, sviluppando secondo Laplace lungo le righe, gli unici termini che non si annullano sono quelli sulla diagonale principale, e si ottiene infine

$$det(A) = (1 + \varepsilon b_{11}) \ldots (1 + \varepsilon b_{nn}) = 1 + \varepsilon(b_{11} + \ldots + b_{nn}) = 1 + \varepsilon tr(A)$$

dove $tr(A)$ indica la traccia di A.
Otteniamo perciò le condizioni equivalenti

$$det(A) = 1 \iff tr(A) = 0$$

e quindi

$$ker(SL_n(p)) \cong \{\text{matrici } (b_{ij}) \text{ a traccia nulla}\}$$

ed in conclusione

$$\text{Lie}(SL_n)(R) = \mathfrak{sl}_n(R).$$

$\mathfrak{sl}_n(R)$ è algebra di Lie, infatti eredita la bracket da $\mathfrak{gl}_n(R) = M_n(R)$ dell'esempio precedente.

Bibliografia

[1] M. F. Atiyah, I. G. McDonald *Introduction to commutative algebra*, Perseus books, 1969.

[2] G. Bergman, *The diamond lemma for ring theory*, Advances in Mathematics 29, 178-218, 1978.

[3] K. S. Brown, *Buildings*, Springer-Verlag, 1989.

[4] T. W. Hungerford, *Algebra*, Springer, 2003.

[5] M. H. A. Newman, *On theories with a combinatorial definition of equivalence*, Annals of Mathematics 43, 223-243, 1942.

[6] J. P. Serre, *Lie Algebra and Lie Groups*, Springer, 1992.

[7] V. S. Varadarajan, *Lie groups, Lie Algebras, and Their Representations*, Springer, 1984.

[8] Wenfeng Ge, *Grobner Bases Theory and The Diamond Lemma*, thesis for the degree of Master of Mathematics in Pure Mathematics, University of Waterloo, Ontario, Canada, 2006.