7. Nilpotent matrices and systems

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **29 (1983)**

Heft 1-2: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **21.07.2024**

define $K(\mu, \nu)$ as the number of semistandard μ-tableaux of type ν for any sequence of nonnegative integers $\nu = (\nu_1, ..., \nu_s)$ such that $|\nu| = n$. Let $\bar{\nu} = (\bar{\nu}_1, ..., \bar{\nu}_s)$ denote the rearrangement of the ν_i such that $\bar{\nu}_1 \geqslant \bar{\nu}_2 \geqslant ... \geqslant \bar{\nu}_s$. Then $K(\mu, \nu) = K(\mu, \bar{\nu})$ and from this (non trivial) fact combined with lemma 6.7 it is easy to see that $K(\mu, \kappa) \leqslant K(\mu, \lambda)$ if $\kappa < \lambda$. (Assume λ covers κ and rearrange both so that the two changing entries are the first two.) We owe these remarks (indirectly) to A. Lascoux.

6.11. *Nilpotent Matrices and Representations* [11]. Let N_κ be the set of nilpotent matrices labelled by the partition κ, cf. 2.11 above. Let $\bar{N}_\kappa$ be its closure and let C be the set of diagonal matrices. Now take the scheme theoretic intersection of the closed subvarieties $\bar{N}_\kappa$ and C of the scheme of $n \times n$ matrices over **C**. This is a finite **C**-algebra with an obvious S_n-action. This turns out to be the permutation representation $\rho(\kappa)$ and using results from [39] a proof of the Snapper, Liebler-Vitale, Lam, Young theorem can be deduced. One very nice thing about this construction is that it also makes sense for the other classical simple Lie algebras and their Weyl groups. There are also relations with the so-called Springer representations of Weyl groups, [40-42].

7. Nilpotent matrices and systems

As was remarked in section 5 above the connection A in the diagram above essentially consists of an almost identical proof of the two theorems. We start with a proof of the Gerstenhaber-Hesselink theorem. The first ingredient which we shall also need for the feedback orbits theorem is the following elementary remark on ranks of matrices.

7.1. *Lemma.* Let $A(t)$ be a family of matrices depending polynomially on a complex or real parameter t. Suppose that rank $A(t) \leqslant$ rank $A(t_0)$ for all t. Then rank $A(t)$ = rank $A(t_0)$ for all but finitely many t.

This follows immediately from the fact that a polynomial in t has only finitely many zeros.

Let A be a nilpotent matrix. Then of course the similarity type of A is determined by the sequence of numbers.

$$n_i = \dim \operatorname{Ker} A^i .$$

The numbers $e_i = n_{i+1} - n_i$ form a partition of n and are dual to the partition formed by the sizes of the Jordan blocks.

The key to a simple proof of the Gerstenhaber-Hesselink theorem is in exploiting this filtration instead of the Jordan form. The following elementary lemma is the key observation.

7.2. *Lemma.* Let A be a nilpotent $n \times n$ matrix and let F be such that

$$F(\operatorname{Ker} A^i) \subset \operatorname{Ker} A^{i-1}, i = 1, 2, \ldots, n . \tag{7.3}$$

Then $tA + (1-t)F$ is similar to A for all but finitely many t.

Proof. We show first that

$$\operatorname{Ker}(tA + (1-t)F)^i \supset \operatorname{Ker} A^i \tag{7.4}$$

for all t. Indeed from (7.3) with $i = 1$ we see that $F(\operatorname{Ker} A) = 0$ and it follows that $(tA + (1-t)F)(\operatorname{Ker} A) = 0$ which proves (7.4) for $i = 1$. Assume with induction that (7.4) holds for all $i < s$. Then

$$\begin{aligned}(tA + (1-t)F)^s \operatorname{Ker} A^s &= (tA + (1-t)F)^{s-1}(tA + (1-t)F)\operatorname{Ker} A^s \\ &\subset (tA + (1-t)F)^{s-1} \operatorname{Ker} A^{s-1} = 0\end{aligned}$$

because $A \operatorname{Ker} A^s \subset \operatorname{Ker} A^{s-1}$ and $F(\operatorname{Ker} A^s) \subset \operatorname{Ker} A^{s-1}$ by (7.3). This proves (7.4). Using (7.4) we know by (7.1) that for almost all t (take $t_0 = 1$)

$$\operatorname{rank}(tA + (1-t)F)^i = \operatorname{rank}(A^i) \tag{7.5}$$

and because $tA + (1-t)F$ and A are both nilpotent it then follows that the conclusion of the lemma is satisfied.

Now let A be a nilpotent matrix. We say that A is of type $\kappa = (\kappa_1, \ldots, \kappa_m)$ if the Jordan normal form of A consists of m Jordan blocks of sizes $\kappa_i \times \kappa_i$, $i = 1, \ldots, m$. E.g. A is of type (4, 2) iff its Jordan form is

$$\left[\begin{array}{cccc|cc} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Consider $\operatorname{Ker} A$, $\operatorname{Ker} A^2, \ldots, \operatorname{Ker} A^n$. Then A is of type κ iff

$$\dim(\operatorname{Ker} A^i) = \kappa_1^* + \ldots + \kappa_i^*, i = 1, \ldots, n$$

where κ^* is the dual partition of κ. Thus in the example the kernel spaces $\operatorname{Ker} A^i$ are spanned by the basis vectors

$$\{e_1, e_5\}, \{e_1, e_2, e_5, e_6\}, \{e_1, e_2, e_3, e_5, e_6\}, \{e_1, e_2, e_3, e_4, e_5, e_6\} .$$

7.6. *Semistandard Tableaux and Nilpotent Matrices.* Let A be a nilpotent matrix of type κ. Let μ be another partition of n and suppose that there is a μ^*-tableaux of type κ^*. Then there is nilpotent matrix F of type μ such that $F(\operatorname{Ker} A^i) \subset \operatorname{Ker} A^{i-1}$ for all i. This matrix F is constructed as follows. First choose a basis $e_1, \ldots, e_n$ of $\mathbf{R}^n$ such that the first $\kappa_1^* + \ldots + \kappa_i^*$ elements of this basis form a basis for Ker A^i, $i = 1, \ldots, n$. Now consider a semistandard μ^*-tableaux T of type κ^*. Take the Young diagram of μ^* and lable the boxes of it by the basis vectors $e_1, \ldots, e_n$ in such a way that the boxes marked with i in the semistandard tableaux T are filled with the basis vectors

$$e_{\kappa_1^* + \cdots + \kappa_{i-1}^* + 1}, \ldots, e_{\kappa_1^* + \cdots + \kappa_i^*}$$

This can be done because T is of type κ^* so that there are precisely κ_i^* boxes labelled i in T. Call this new μ^*-tableaux T'. Now define F by $F(e_i) = e_j$, if e_j is just above e_i in the μ^*-tableaux T' and $F(e_j) = 0$ if e_j occurs in the first row of T'. Then obviously

$$\dim \operatorname{Ker} F^i = \mu_1^* + \ldots + \mu_i^*{}_*$$

so that F is of type μ and $F(\operatorname{Ker} A^i) \subset \operatorname{Ker} A^{i-1}$ because the μ^*-tableaux T was semistandard which implies that the labels are strictly increasing along columns.

An example may illustrate things. Let $\kappa^* = (2, 2, 2)$, $\mu^* = (4, 1, 1)$. A μ^*-tableaux of type κ^* is then

$$\begin{matrix} 1 & 1 & 2 & 3 \\ 2 & & & \\ 3 & & & \end{matrix}$$

Inserting $e_1, \ldots, e_6$ in such a way that e_1, e_2 are put into boxes marked with 1, e_3, e_4 in boxes marked with 2 and e_5, e_6 in boxes marked with 3 gives for example

$$\begin{matrix} e_1 & e_2 & e_3 & e_5 \\ e_4 & & & \\ e_6 & & & \end{matrix}$$

which yields an F defined by $F(e_6) = e_4$, $F(e_4) = e_1$,

$$F(e_1) = F(e_2) = F(e_3) = F(e_5) = 0\,.$$

7.7. *Proof of the Gerstenhaber-Hesselink Theorem.* (Cf. 2.11 above for a statement of the theorem.)

The implication $\rightarrow$ is immediate. Indeed if $A_t \in 0(\kappa)$ converges to $A_0 \in 0(\lambda)$ as $t \rightarrow 0$ then rank $(A_t^i) \geqslant$ rank (A_0^i) for small t and all $i = 1, ..., n$. Hence

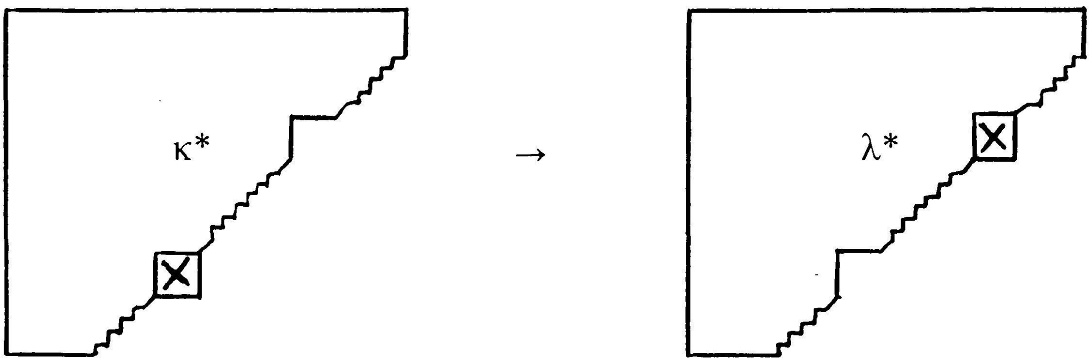
$$\dim(\operatorname{Ker} A_t^i) \leqslant \dim(\operatorname{Ker} A_0^i)$$

for small t so that

$$\kappa_1^* + ... + \kappa_i^* \leqslant \lambda_1^* + ... + \lambda_i^*$$

for all i, hence $\kappa^* > \lambda^*$ and $\kappa < \lambda$. To prove the opposite implication it suffices to show this in case that κ is obtained from λ by a transformation of the type described in lemma 6.7. (Because if $\overline{0(\kappa)} \supset 0(\lambda)$ and $\overline{0(\lambda)} \supset 0(\mu)$, then $\overline{0(\kappa)} \supset \overline{0(\lambda)}$, and hence $\overline{0(\kappa)} \supset 0(\mu)$.) Then λ^* is obtained from κ^* by a similar transformation.

Recall the picture



Now take the unique semistandard κ^*-tableau of type κ^* and transform the box $\boxtimes$ together with its label. The result is obviously a semistandard λ^*-tableau of type κ^*. Let A be a nilpotent matrix of type κ. Then by the construction of 7.6 above there is an F of type λ such that $F(\operatorname{Ker} A^i) \subset \operatorname{Ker} A^{i-1}$. Then $tA + (1-t)F$ is similar to A for almost all t by lemma 7.2 so that there is a sequence of A's in $0(\kappa)$ converging to $F \in 0(\lambda)$, proving that $0(\lambda) \subset \overline{0(\kappa)}$, which finishes the proof of the theorem.

Incidentally it is quite easy to describe F directly without resorting to semistandard tableaux [7].

7.10. *Kronecker Indices of Systems.* Let $(A, B) \in L_{m,n}^{cr}$ be a completely reachable pair of matrices. Recall that this means the matrix $R(A, B) = (B\ AB\ ...\ A^nB)$ has rank n. Recall that the Kronecker indices $\kappa(A, B)$ of the pair (A, B) are defined as follows. Let for $i = 1, ..., n$

(7.11) $\quad V_i(A, B) =$ space spanned by the column vectors of

$$A^jB, j = 0, ..., i-1 .$$

Let

$$d_i = \dim V_i(A, B), e_i = d_i - d_{i-1}, d_0 = 0 .$$

Then

$$e_i \leqslant e_{i-1}, i = 1, \dots, n-1 ,$$

and $\kappa(A, B)$ is defined as the dual partition of n.

$$\kappa(A, B) = e(A, B)^* \tag{7.12}$$

where $e(A, B) = (e_1, \dots, e_n)$.

The orbits of the feedback group (cf. 2.6 above) acting on $L^{cr}_{m,n}$ are precisely the subsets of $L^{cr}_{m,n}$ with constant $\kappa(A, B)$. Let $U(\kappa)$ be this orbit. The "degeneration of systems theorem" now says

7.13. *Theorem.* $\overline{U(\lambda)} \supset U(\kappa) \leftrightarrow \lambda > \kappa$.

Here follows a proof which is virtually identical with the proof of the Gerstenhaber-Hesselink theorem given above. First if $(A_t, B_t) \to (A_0, B_0)$ as $t \to 0$,

$$(A_t, B_t) \in U(\lambda), (A_0, B_0) \in U(\kappa) ,$$

then

$$\operatorname{rank}(A_t^{i-1}B_t; \dots; A_tB_t; B_t) \geqslant \operatorname{rank}(A_0^{i-1}B_0; \dots; A_0B_0; B_0)$$

for small t. Hence

$$\dim V_i(A_t, B_t) \geqslant \dim V_i(A_0, B_0)$$

for small t. Hence $e(A_t, B_t) < e(A_0, B_0)$ for small t and $\kappa(A_t, B_t) > \kappa(A_0, B_0)$ for small t which proves the implication $\Rightarrow$.

To prove the inverse implication it suffices to prove this in the case κ is obtained from λ by a transformation as described in lemma 6.7 (exactly as in the case of the Gerstenhaber-Hesselink theorem). Now let $(A, B) \in U(\lambda)$. Choose a basis $e_1, \dots, e_n$ for $\mathbf{R}^n$ such that the first $\lambda_1^* + \dots + \lambda_i^*$ elements of $e_1, \dots, e_n$ form a basis for $V_i(A, B)$, $i = 1, \dots, n$. Now write in the $e_1, \dots, e_n$ in λ^* in the standard way and transform λ^* backwards to κ^*, moving $\boxtimes$ together with its label, cf. the picture in section 7.7 above. E.g. if $\kappa^* = (4, 3, 2, 2, 1)$ and $\lambda^* = (4, 4, 2, 1, 1)$ then this would give

$$\begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \\ e_5 & e_6 & e_7 & \\ e_9 & e_{10} & & \\ e_{11} & e_8 & & \\ e_{12} & & & \end{array} \qquad \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \\ e_5 & e_6 & e_7 & e_8 \\ e_9 & e_{10} & & \\ e_{11} & & & \\ e_{12} & & & \end{array}$$

The vectors in the first i rows of λ^* are a basis for $V_i(A, B)$. Now define a pair (F, G) in terms of κ^* as follows. G consists of the vectors in the first row of κ^* (plus a zero vector in case $\kappa_1^* < \lambda_1^*$), and F is defined by $F(e_i) = e_{i'}$ if $e_{i'}$ occurs just below e_i in κ^* and $F(e_i) = 0$ otherwise. Note the similarity with the construction in 7.6. One could put this in "Young tableaux" terms too. The relevant "Young tableaux" are then the inverse semistandard ones with labels strictly decreasing from left to right along rows and decreasing from top to bottom along columns. Then (F, G) has the following properties (all immediate)

(i) $(F, G) \in U(\kappa) \subset L_{m,n}^{cr}$

(ii) $V_i(F, G) \subset V_i(A, B)$

(iii) $F V_i(A, B) \subset V_{i+1}(A, B)$

(of course (ii) follows from (iii) together with $V_1(F, G) \subset V_1(A, B)$). Now consider $A_t = tA + (1-t)F$, $B_t = tB + (1-t)G$. Then

$$V_i(A_t, B_t) \subset V_i(A, B) \qquad \text{for all } t \tag{7.14}$$

$$V_i(A_t, B_t) = V_i(A, B) \qquad \text{for all but finitely many } t. \tag{7.15}$$

Indeed obviously $V_1(A_t, B_t) \subset V_1(A, B)$ because of (ii) above for $i = 1$. Now assume that (7.14) holds for all $i < r$. Then

$$\begin{aligned} V_r(A_t, B_t) &= \big(tA + (1-t)F\big)V_{r-1}(A_t, B_t) + V_{r-1}(A_t, B_t) \\ &\subset tAV_{r-1}(A, B) + (1-t)FV_{r-1}(A, B) + V_{r-1}(A, B) \\ &\subset V_r(A, B) + V_r(A, B) + V_{r-1}(A, B) = V_r(A, B) \end{aligned}$$

This proves (7.14) and (7.15) follows by means of lemma 7.1 (with $t_0 = 1$) because

$$\dim V_i(A_t, B_t) = \operatorname{rank}\big(A_t^{i-1}B_t; \ldots; B_t\big)$$

Now $(A_t, B_t) \to (F, G) \in U(\kappa)$ as $t \to 0$ and by (7.15) (and the theorem that the orbits under the feedback group are classified by the Kronecker indices) all but finitely many of the (A_t, B_t) are feedback equivalent to (A, B). Thus $(F, G) \in U(\kappa)$ and $(F, G) \in \overline{U(\lambda)}$ proving the theorem.

7.16. *Remarks.* The two proofs are very similar (up to duality in a certain sense). This can be given more precise form as follows. For a nilpotent matrix $N \in N_n$ let

$$\underline{s}(N) = \{(A, B) \in L_{m,n}^{cr} \mid N^i A^{i-1} B = 0, i = 1, \ldots, n\}$$

and for $(A, B) \in L_{m,n}^{cr}$ let

$$\underline{t}(A, B) = \{N \in N_n \mid N^i A^{i-1} B = 0, i = 1, \ldots, n\}.$$

Then using the results above one shows that

$$\underline{t}\ \underline{s}(\overline{0(\kappa)}) = \overline{0(\kappa)},\ \underline{s}\ \underline{t}(\overline{U(\kappa)}) = \overline{U(\kappa)}$$

so that $\underline{t}$ and $\underline{s}$ set up a bijective correspondence between the closures of orbits in the two cases and hence induce a bijective order preserving correspondence between the sets of orbits themselves.

8. Vectorbundles and systems

This section contains a modified version of the construction of Hermann-Martin [14] associating a vectorbundle $E(\Sigma)$ over the Riemann sphere $\mathbf{P}^1(\mathbf{C})$ to every $\Sigma = (A, B) \in L^{cr}_{m,n}$. This version makes it almost trivial to see that $E(\Sigma)$ splits as a direct sum of line bundles $L(\kappa_i)$, $i = 1, ..., m$ where $\kappa = (\kappa_1, ..., \kappa_m)$ is the set of Kronecker indices of Σ.

The first thing needed is some more information on the universal bundle ξ_m.

8.1. *On the Universal Bundle* $\xi_m \to \mathbf{G}_n(\mathbf{C}^{n+m})$. Let V be a complex $n + m$ dimensional vector space and $V^* = \mathrm{Hom}_{\mathbf{C}}(V, \mathbf{C})$ its dual vector space. Given $x \in \mathbf{G}_n(\mathbf{C}^{n+m})$ define $x^* = \{y \in V^* \mid < y, v > = 0 \text{ for all } x \in V\}$ where $< , >$ denotes the usual pairing $V^* \times V \to \mathbf{C}$. Then x^* is m-dimensional and $x \mapsto x^*$ defines a holomorphic isomorphism

$$d : \mathbf{G}_n(V) \to \mathbf{G}_m(V^*)\,. \tag{8.2}$$

Now $v \in V/x$ defines a unique homomorphism $v^T : x^* \to \mathbf{C}$ as follows:

$v^T(a) = < a, \tilde{v} >$ for all $a \in x^*$, where $\tilde{v} \in V$ is any representative of v. This is well defined because $< a, b > = 0$ for all $b \in x$ if $a \in x^*$. This defines an isomorphism between the pullback $(d^{-1})\ \xi_m$ and the dual of the subbundle η_m on $G_m(V^*)$ defined by

$$\eta_m = \{(x^*, w) \in \mathbf{G}_m(V^*) \times V^* \mid w \in x^*\}$$

It follows that ξ_m is a subbundle of an $n + m$ dimensional trivial bundle $\mathbf{G}_n(\mathbf{C}^{n+m}) \times \mathbf{C}^{n+m}$. Because $\mathbf{G}_n(\mathbf{C}^{n+m})$ is projective (compact) all holomorphic maps $\mathbf{G}_n(\mathbf{C}^{n+m}) \to \mathbf{C}$ are constant so that the space of holomorphic sections $\Gamma(\mathbf{G}_n(\mathbf{C}^{n+m}) \times \mathbf{C}^{n+m}, \mathbf{G}_n(\mathbf{C}^{n+m}))$ is of dimension $n + m$. As a subbundle of a trivial $(n+m)$-dimensional bundle ξ_m can therefore have at most $(n+m)$ linearly