Nilpotent conjugacy classes in the classical groups

Bill Casselman
University of British Columbia
cass@math.ubc.ca

When I began writing this, my tentative title was 'What you always wanted to know about nilpotence but were afraid to ask'. That title is unfortunately not likely to be an accurate description of the present version. Instead, this essay is my unfinished attempt to understand and explain nilpotent conjugacy classes in the classical complex semi-simple Lie algebras (and therefore also, through the exponential map, of the unipotent classes in the associated groups). There are many accounts of this material in the literature, but I try here something a bit different from what is available. For one thing, I have written much of this essay while writing programs to investigate nilpotent classes in all semi-simple complex Lie algebras. In such a task, problems arise that do not seem to be clearly dealt with in the literature.

I begin with the simplest case M_n, the Lie algebra of GL_n. Here there are no mysteries, and I discuss this case in more detail than the others. The Jordan decomposition theorem tells us that there is a bijection of nilpotent classes with partitions of n. The main result relates dominance of partitions to closures of conjugacy classes of nilpotent matrices. The original reference for thsi material seems to be [Gerstenhaber:1959], which I follow loosely, although my explicit algorithmic approach seems to be new. I have also used [Collingwood-McGovern:1993], which is probably the most thorough introduction to the subject.

I include for completeness, at the beginning, a discussion of partitions, particularly of Young's raising and lowering operators (which I call shifts). In this discussion I have taken some material from [Brylawski:1972] whch although elementary does not seem to be found elsewhere.

The basic question is simple to describe by an example. Every nilpotent 4×4 matrix has one of these Jordan forms:

$$\begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & 1 & \circ \\ \circ & \circ & \circ & 1 \\ \circ & \circ & \circ & \circ \end{bmatrix} \quad \begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & 1 & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix} \quad \begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & 1 \\ \circ & \circ & \circ & \circ \end{bmatrix} \quad \begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix} \quad \begin{bmatrix} \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix}$$

which are parametrized in an obvious way by partitions

$$(4),\ (3,1),\ (2,2),\ (2,1,1),\ (1,1,1,1)$$

of 4 that specify which blocks of regular nilpotents occur on the diagonal. To each of these corresponds the set all the matrices similar to it, which form a locally closed algebraic subvariety of M_4. The Zariski closure of one of these will be a union of some of the others. *Which occur in this closure?*

I also discuss in the case of GL_n the subvarieties $\mathfrak{B}_\nu$ defined by Steinberg and Spaltenstein.

In the last sections, I sketch to a less extent what happens in the symplectic and orthogonal groups, for which the results are very similar. In another essay, or perhaps just a later version of this one, I'll compare the relatively simple methods used here with the more systematic approach of [Carter:1985].

I wish to thank Erik Backelin for pointing out an error in an earlier version, in which I short-circuited an argument of Gerstenhaber.

Contents

Part I. Young diagrams

1. Partitions

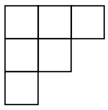
In this essay a **partition** will be a weakly decreasing array (λ_i) of non-negative numbers

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$$

eventually equal to 0. The **magnitude** of a partition λ is

$$|\lambda| = \sum \lambda_i \,.$$

I shall often identify a partition of magnitude n with an array of length n, filled in with zeroes if necessary—I shall assume $\lambda_k = 0$ where not explicitly specified. Each partition $(\lambda_1, \lambda_2, \ldots, \lambda_k)$ corresponds to a **Young diagram**, which has rows of lengths λ_1, λ_2, etc. arrayed under each other and aligned flush left. For example, the partition $6 = 3 + 2 + 1$ matches with this diagram:



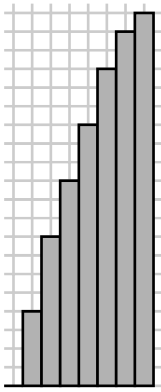
Multiplicity in partitions is commonly expressed by exponents. Thus

$$(4, 4, 3, 3, 3, 2, 1) = (4^2, 3^3, 2, 1)\,.$$

DOMINANCE. If (λ_k) is a partition, define the *summation array* $\sum^\lambda = \left(\sum_k^\lambda\right)$ with

$$\textstyle\sum_k^\lambda = \displaystyle\sum_1^k \lambda_i \,.$$

The map taking λ to $\sum^\lambda$ is a bijection between partitions of magnitude n with those arrays $\sum = \left(\sum_k\right)_{k=0}^n$ such that $\sum_0 = 0$, $\sum_n = n$, and the graph of the function $k \mapsto \sum_k$ is weakly concave:



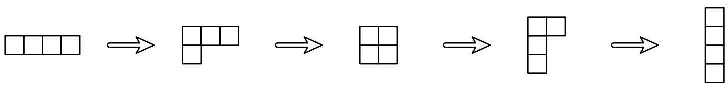
If λ, μ are partitions with $|\lambda| = |\mu|$, the partition λ is said to **dominate** μ if

$$\sum_k^\lambda = \lambda_1 + \cdots + \lambda_k \geq \sum_k^\mu = \mu_1 + \cdots + \mu_k$$

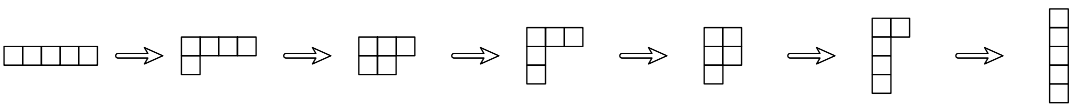
for all k. This happens if and only if the graph of $\sum^\lambda$ lies weakly above that of $\sum^\mu$. For example,

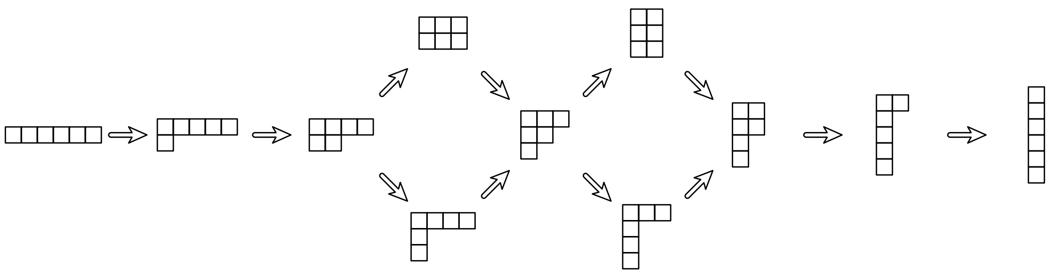
$$(4) \geq (3,1) \geq (2,2) \geq (2,1,1) \geq (1,1,1,1)\,.$$

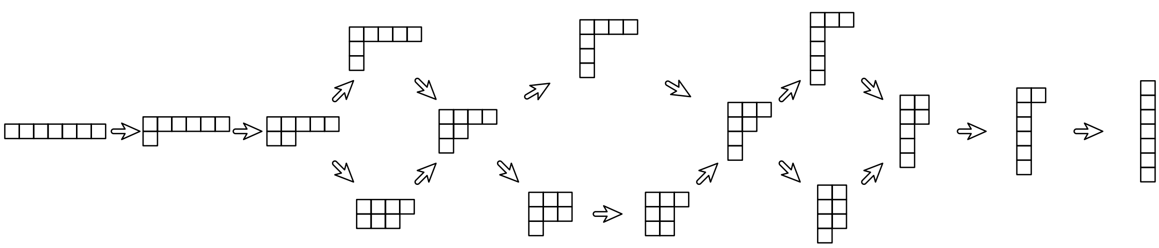
Or, in a diagram:



Here are a few more examples:







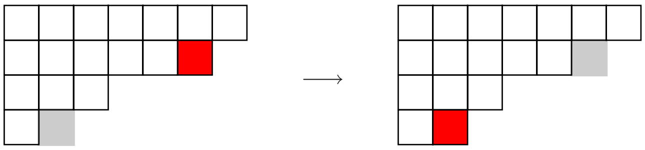
As we shall see, this is a very natural partial ordering, and one of great importance in representation theory. Unfortunately, beyond a few simple features, there does not seem to be much pattern to them. One of the simple features is horizontal symmetry, which I shall explain later.

SHIFTS. For each pair (i, j) with $i < j$ let $a = \alpha_{i,j}$ be the integral vector with coordinates $a_k = 0$ for all $k \neq i, j$, $a_i = 1$, $a_j = -1$. If λ is a partition, a partition μ is obtained from it by a **shift** if $\mu = \lambda \pm \alpha_{i,j}$ for some (i, j). The shift is said to be *down* if $\alpha_{i,j}$ is subtracted.

For example:

$$(4) \longmapsto (3, 1) = (4, 0) - (1, -1)\,.$$

In terms of Young diagrams, a shift down by $\alpha_{i,j}$ removes a box from the end of row i and places it on the end of row j somewhere below it. For example, a shift down by $\alpha_{2,4}$ changes the diagram on the left below to the one on the right:



In these circumstances

(1.1)
$$\Sigma_k^\mu = \begin{cases} \Sigma_k^\lambda & \text{if } k < i \\ \Sigma_k^\lambda - 1 & i \le k < j \\ \Sigma_k^\lambda & j \le k. \end{cases}$$

Hence:

1.2. Lemma. *Suppose λ to be a partition, with $\mu = \lambda - \alpha_{i,j}$.*

(a) *The array μ is a partition if and only if*

$$\begin{cases} \lambda_{i+1} + 1 \le \lambda_i - 1 & \text{if } j = i + 1 \\ \lambda_j < \lambda_{j-1} \le \lambda_{i+1} < \lambda_i & \text{if } j \ge i + 2. \end{cases}$$

(b) *In these circumstances, λ dominates μ.*

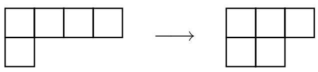
If $\lambda \ge \mu$ define the distance from μ to λ to be

$$\|\lambda - \mu\| = \sum_k \left(\Sigma_k^\lambda - \Sigma_k^\mu\right).$$

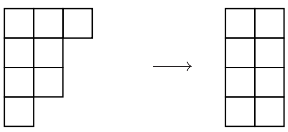
If $\lambda \ge \mu$ and $\|\lambda - \mu\| = 0$ then $\lambda = \mu$. If λ and $\mu = \lambda - \alpha_{i,j}$ are both partitions, it follows from (1.1) that $\|\lambda - \mu\| = j - i$.

Certain shifts are more elementary than others. If $\mu = \lambda - \alpha_{i,j}$ then it is said to be obtained by a **short shift** if either $(*)$ $j = i + 1$ (in which case $\lambda_{i+1} \le \lambda_i - 2$) or $(**)$ $\lambda_j = \lambda_i - 2$ and $\lambda_k = \lambda_i - 1$ for all $i > k > j$.

$*$:



$**$:



In these circumstances, μ is again a partition, and it should be clear that λ **covers** μ in the sense that there is no partition properly between them. Following [Brylawski:1973], in the first case I'll say that λ is a $*$ cover of μ, and in the second a $**$ cover.

The following is probably well known, but I have seen it formulated only by Brylawski (Proposition 2.3 in his paper). A weaker version is classical—found, for example, in [Hardy-Littlewood-Polya:1952] (Lemma 2, page 47), and also in [Gerstenhaber:1961].

1.3. Proposition. *The partition λ covers μ if and only if μ is obtained from λ by a short shift.*

Because any chain between two partitions can be saturated, here is an equivalent statement:

1.4. Corollary. *If λ and μ are partitions, then $\lambda \geq \mu$ if and only if μ can be obtained from λ by a sequence of zero or more short shifts.*

This enables one to construct easily the partially ordered set of all partitions of a given n. The process starts with (n), which dominates all others. It then uses a stack to produce all short shifts from a succession of partitions.

Proof. This proof will be constructive, and goes by induction on $\|\lambda - \mu\|$. When this is 0 there is nothing to prove. Otherwise, say $\lambda > \mu$. I claim that we can find a shift down taking λ to a partition $\kappa = \lambda - \alpha_{k,\ell}$ with

$$\mu \leq \kappa < \lambda\,.$$

If so, then

$$\|\kappa - \mu\| = \|\lambda - \mu\| - (\ell - k)$$

and we can apply the induction hypothesis.

I must now just prove the claim. There are several steps.

Step 1. *Suppose that $\sum_k^\lambda > \sum_k^\mu$ and that $\lambda_{k+1} \leq \lambda_k - 2$. Then $\kappa = \lambda - \alpha_{i,i+1}$ is still a partition, λ is a $\diamond 0$ cover of κ, and $\kappa \geq \mu$.*

This is an easy exercise. We'll apply it in a moment.

Step 2. Let m be least such that $\sum_m^\lambda > \sum_m^\mu$.

Then $\sum_i^\lambda = \sum_i^\mu$ for $i < m$ but $\sum_m^\lambda > \sum_m^\mu$. This implies as well that $\lambda_i = \mu_i$ for $i < m$ but $\lambda_m > \mu_m$. Furhermore $\mu_m > 0$, since otherwise $\sum_m^\mu = |\mu| = |\lambda| > \sum_m^\lambda$, a contradiction. Therefore $\lambda_m > \mu_m > 0$ and $\lambda_m \geq 2$.

Step 3. Let k be greatest such that $\lambda_k = \lambda_m$. Thus $\lambda_{k+1} < \lambda_k$.

We still have $\sum_k^\lambda > \sum_k^\mu$, and in fact

(1.5)
$$\textstyle\sum_k^\lambda > \sum_k^\mu + (k - m)$$

since

$$\begin{aligned}
\lambda_k &= \lambda_m > \mu_m \geq \mu_k \geq 0\,, \\
\textstyle\sum_k^\lambda &= \textstyle\sum_m^\lambda + (k-m)\lambda_m \\
&= \textstyle\sum_m^\lambda + (k-m)(\lambda_m - \mu_m) + (k-m)\mu_m \\
&> \textstyle\sum_m^\mu + (k-m)\mu_m + (k-m) \\
&\geq \textstyle\sum_m^\mu + \mu_{k+1} + \cdots + \mu_m + (k-m) \\
&\geq \textstyle\sum_k^\mu + (k-m)\,.
\end{aligned}$$

Step 4. If $\lambda_{k+1} \leq \lambda_k - 2$, set $\kappa = \lambda - \alpha_{m,m+1}$. Then λ covers κ and $\kappa \geq \mu$, according to what was shown in the first step.

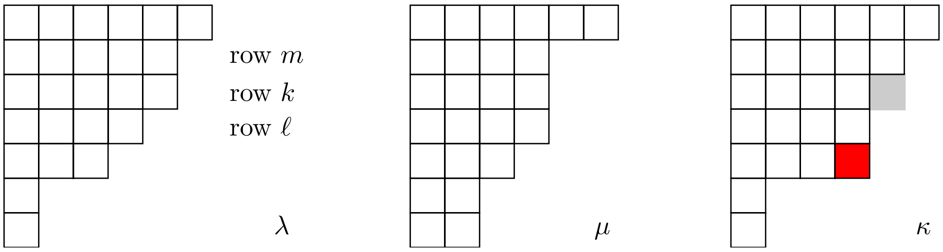
Step 5. Otherwise, we now have $\lambda_{k+1} = \lambda_k - 1$. Choose $\ell \geq k+1$ such that $\lambda_\ell = \lambda_{k+1}$ but $\lambda_{\ell+1} < \lambda_\ell$, which is certainly possible since eventually $\lambda_\ell = 0$. Then certainly $\lambda_i = \lambda_k - 1$ for $k > i \geq \ell$.

Furthermore, from (1.5) it follows that $\sum_\ell^\lambda > \sum_\ell^\mu$.

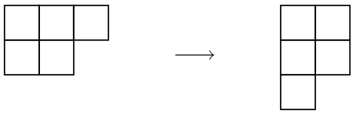
Step 6. If $\lambda_{\ell+1} \leq \lambda_\ell - 2$, set $\kappa = \lambda - \alpha_{\ell,\ell+1}$. Again apply the assertion of the first step.

Step 7. Otherwise, we now have $\lambda_{\ell+1} = \lambda_\ell - 1 = \lambda_k - 2$. Set $\kappa = \lambda - \alpha_{k,\ell+1}$. Then $\kappa_i = \kappa_j$ for $k \leq j \leq \ell+1$, so λ is a $**$ cover of κ. I leave it as exercise to see that $\kappa \geq \mu$.

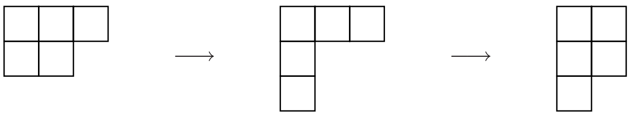
Here is an example of how it goes:



Remark. Proposition 6.2.4 of [Collingwood-McGovern:1993] contains an incorrect criterion for covering. Among other problems, there seems to be a simple typographical error, but even assuming the obvious correction, this Proposition tells us that the shift



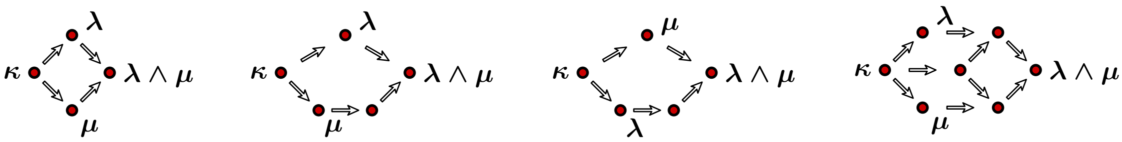
is a covering, whereas in fact it factors as



THE LATTICE OF PARTITIONS. The partially ordered set of partitions of a given magnitude is a lattice, which is to say that any pair of partitions λ, μ possesses a unique maximal $\nu = \lambda \wedge \mu$ such that $\nu \leq \lambda$, $\nu \leq \mu$, and also a minimal $\nu = \lambda \vee \mu$ dominating both. The summation array of $\lambda \wedge \mu$ is the infimum of $\sum^\lambda$ and $\sum^\mu$.

I find the structure of the set of partitions of n ordered by dominance rather mysterious. There does not seem to be any clear pattern in sight. There is, however, some local patterm that is pointed out in Proposition 3.2 of [Brylawski:1973], which I'll say something about here.

1.6. Proposition. *Suppose $\lambda \neq \mu$ are both covered by κ. The following diagrams exhaust the possible configurations for the intervals $[\kappa, \lambda \wedge \mu]$:*



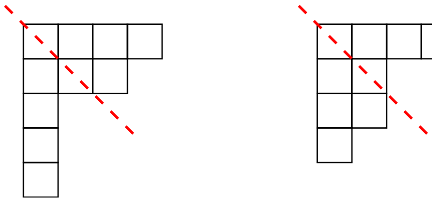
I leave the proof as an exercise.

[Brylawski:1973] goes on to describe the Möbius function on the lattice of partitions, which he shows to take only values $0, \pm 1$, but I'll not say anything about that here.

Remark. Dominance is just one example of closure relations among nilpotent conjugacy classes in semi-simple Lie algebras, and to tell the truth I find all of them more mysterious than not. Almost every argument involving them sooner or later goes through case by case, a curiously unsatisfactory state of affairs. As Nicolas Spaltenstein complains somewhere, there is no real theory, just a collection of examples. The situation is reminiscent of the classification of semi-simple Lie algebras, which was not really understood until Kac-Moody algebras were discovered. What is particularly frustrating is that there is, as far as I know, no uniform algorithm to produce a list of all classes for a Lie algebra with given root system, and along with it all important data.

2. Duality of diagrams

Suppose λ to be a partition of n. This means that λ_i is the length of the i-row in its Young diagram. The column lengths μ_i also give a partition of n, called the **dual partition**. Geometrically, this amounts to reflecting the diagram along the NW-SE axis. For example, $(4,3,1,1,1)$ changes to $(5,2,2,1)$:



Thus μ_1 is the number of non-zero rows in the diagram, μ_2 is the number of rows of length at least two, and in general μ_i is the number of $\lambda_j \geq i$. Equivalently

$$\mu_i - \mu_{i+1} = \text{ number of } \lambda_j \text{ equal to } i\ .$$

A good method for computing μ relies on the observation that the vector δ of differences $\delta_i = \mu_i - \mu_{i-1}$ is so easy to calculate from λ. The following program scans only the array of row lengths in the diagram:

```
r = number of non-zero λ
c = λ_1
δ = [0, . . . , 0]  (length c)
for i in [1, r]:
    δ_{λ_i} = δ_{λ_i} + 1
μ = [0, . . . , 0]
μ_1 = r
for j in [2, c]:
    μ_j = μ_{j−1} − δ_j
```

The following basic fact relating dominance and duality does not seem to be quite trivial.

2.1. Lemma. *If λ and μ are two partitions of the same integer n, then $\lambda \geq \mu$ if and only if $\overline{\mu} \geq \overline{\lambda}$.*

Proof. It is easy to check when one is obtained from the other by a shift, since the dual of a shift down is a shift up, as a diagram will make clear. Corollary 1.4 implies that this suffices. ▯

This duality explains the reflection symmetry in the diagrams of the dominance graphs in §1.

Part II. The special linear group

3. Classification

In this section I'll begin to explain the relationship between partitions and nilpotent matrices.

STANDARD NILPOTENT MATRICES. For every $\ell \geq 1$ define ν_ℓ to be the nilpotent $\ell \times \ell$ matrix $\nu = (n_{i,j})$ with

$$n_{i,j} = \begin{cases} 1 & \text{if } j = i+1 \\ 0 & \text{otherwise.} \end{cases}$$

For example

$$\nu_4 = \begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & 1 & \circ \\ \circ & \circ & \circ & 1 \\ \circ & \circ & \circ & \circ \end{bmatrix}.$$

For every partition λ of n define the nilpotent matrix ν_λ to be the direct sum of matrices ν_{λ_i}. For example

$$\nu_{3,2} = \left[\begin{array}{ccc|cc} \circ & 1 & \circ & \circ & \circ \\ \circ & \circ & 1 & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \hline \circ & \circ & \circ & \circ & 1 \\ \circ & \circ & \circ & \circ & \circ \end{array}\right].$$

3.1. Theorem. *Every nilpotent matrix is conjugate to a unique ν_λ.*

Proof. The proof comes in two parts, existence and uniqueness.

EXISTENCE. For this I'll give two versions. The first is more conceptual, but the second more direct.

(1) Suppose ν to be a nilpotent matrix. The map $x \mapsto \nu$ generates a ring homomorphism from the polynomial ring $R = F[x]$ to $\mathrm{M}_n(F)$, making F^n into a finite-dimensional module over $F[x]$. Let (e_i) be the standard basis of the vector space F^n, We now get a map from R^n to F^n, taking $\big(P_i(x)\big)$ to $\big(\sum_i P_i(\nu)e_i\big)$.

The kernel is an R-submodule K of R^n. This is finitely generated on general principles since R is Noetherian, but in fact we can find an explicit set of generators. Since ν is nilpotent, we know that $\nu^N = 0$ for some N. Therefore the map from R^n to F^n factors through the F-space $\big(R/(x^N)\big)^n$. Finding its kernel is just a matter of finding the null space of a given matrix.

Generators of the kernel give us an $n \times N$ matrix whose columns are in R^n. The elementary divisor theorem tells us how to reduce this to a diagonal matrix with non-increasing diagonal entries. Therefore this module is isomorphic to a direct sum of quotients $F[x]/(x^{\lambda_i})$ with $\lambda_1 \geq \ldots \geq \lambda_k$ and $\lambda_1 + \cdots + \lambda_k = n$. But if we choose basis $1, x, \ldots, x_{\ell-1}$ multiplication by x on $F[x]/(x^\ell)$ has matrix ν_ℓ. So a suitable choice of basis will give us ν_λ as the matrix of ν.

(2) Let ν be a nilpotent matrix acting on the vector space V. Calculate powers of ν to find k such that $\nu^k = 0$ but $\nu^{k-1} \neq 0$. We can do this most efficiently by conjugating ν to an upper triangular matrix first (as recalled in the appendix). At any rate, we get in this way a vector v such that $\nu^{k-1}v \neq 0$. Let U be the vector subspace spanned by the vectors $v_i = \nu^i v$ for $0 \leq i < k$. It has dimension k (i.e. the vectors v_i are linearly independent. Follow the method outlined in the appendix to extend the v_i to a basis. Applying induction, we may find a Jordan form of ν on V/U. The original matrix will now be similar to one of the form

$$\begin{bmatrix} \nu_k & m \\ 0 & \nu_* \end{bmatrix}$$

with ν_* in Jordan form. Progressing through the columns of m from left to right, using the fact that $\nu_*^k = 0$, we may change the basis to get a matrix in Jordan form.

UNIQUENESS. I shall interpret the partition λ in terms of explicitly available information about ν. If ν is any nilpotent $n \times n$ matrix, then by the previous results we know that $\nu^n = 0$, so we have a filtration

$$0 = \mathrm{Ker}(\nu^0) \subseteq \mathrm{Ker}(\nu^1) \subseteq \ldots \subseteq \mathrm{Ker}(\nu^n) = F^n ,$$

and the

$$\kappa_k = \dim \mathrm{Ker}(\nu^k) - \dim \mathrm{Ker}(\nu^{k-1}) \quad (k > 0)$$

satisfy the equation

$$\kappa_1 + \kappa_2 + \cdots = n .$$

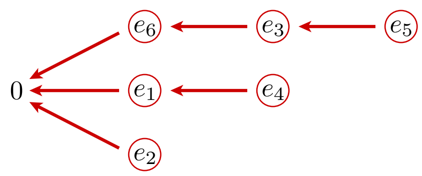
I'll call κ the **kernel partition** of ν. It clearly depends only on the conjugacy class of ν. Here is the basic fact relating the Jordan block partition of a nilpotent matrix with its kernel partition:

3.2. Proposition. *The kernel partition of ν_λ is the partition $\overline{\lambda}$ dual to λ.*

This is not difficult to prove, but there is a way to see it immediately, in terms of **tableaux**. A tableau in this note will be any numbering of boxes of size n in a Young diagram by integers in the range $[1, n]$, like this:

6	3	5
1	4	
2		

Each tableau corresponds not just to a class of similar matrices, but in fact a matrix. I illustrate this by an example—the matrix ν_λ with $\lambda = (3, 2, 1)$ can be represented graphically as



This can be represented more succinctly just by the tableau above.

Thus we can interpret a tableau as a nilpotent linear transformation according to the rule that it takes $e_{t_{i,j}}$ to $e_{t_{i,j-1}}$ if $j > 0$ and otherwise to 0. In other words, it shifts indices of basis elements to the left one column. If T is this transformation, then the kernel of T is spanned by the basis elements shifted completely off to the left, which is to say those whose indices are in the first column. And the kernel of T^k is spanned by those whose indices occur in one of the first k columns.

We can now read off Proposition 3.2 immediately.

The use of tableaux to illustrate nilpotent transformations is occasionally very useful. Any tableau of the same shape λ gives rise to a matrix conjugate to ν_λ—choosing a new tableau with the same diagram just amounts to renaming the basis elements.

3.3. Corollary. *If ν is a nilpotent linear transformation, the partition associated to its dual $\widehat{\nu}$ is the same as that associated to ν.*

Proof. Because the transpose of a Jordan matrix is clearly similar to it. After all, if ν takes e_i to e_j than $\widehat{\nu}$ takes $\widehat{e}_j$ to $\widehat{e}_i$. In other words, in the diagram above we can just reverse the arrows to get the diagram for $\widehat{\nu}$.

An immediate consequence of Proposition 3.2:

3.4. Proposition. *If A is nilpotent and $c \neq 0$ then A and cA are conjugate.*

4. Duality of classes

If λ is a partition of n, then so is its dual $\overline{\lambda}$. There is an elegant characterization of the nilpotent class $\nu_{\overline{\lambda}}$ in terms of ν_λ.

Every partition λ of n determines a parabolic subgroup P_λ of GL_n. Its Levi factor is made up of matrices of diagonal blocks of $\lambda_i \times \lambda_i$, and the Lie algebra of its unipotent radical is that of nilpotent matrices in the complement of these above the diagonal. Here, for example, is what this subgroup looks like for $(4, 2, 1)$:

$$\begin{bmatrix} & & & & * & * & * \\ & & & & * & * & * \\ & & & & * & * & * \\ & & & & * & * & * \\ \circ & \circ & \circ & \circ & & & * \\ \circ & \circ & \circ & \circ & & & * \\ \circ & \circ & \circ & \circ & \circ & \circ & \end{bmatrix}$$

Here the sign $*$ marks matrices in the nilpotent radical.

The matrix ν_λ is the sum of principal nilpotent matrices in $\mathfrak{m}_\lambda$.

4.1. Lemma. *If U is any affine subspace of a nilpotent Lie subalgebra of $\mathfrak{sl}_n$, then there exists a unique nilpotent conjugacy class in $\mathfrak{sl}_n$ whose intersection with U is open.*

This conjugacy class is called the **generic** class of U.

Proof. Because there are only a finite number of nilpotent conjugacy classes in $\mathfrak{g}$.

As a consequence, there exists a unique nilpotent conjugacy class n_λ that intersects $\mathfrak{n}_\lambda$ in a dense open subset.

4.2. Proposition. *The conjugacy class of n_λ is that of $\nu_{\overline{\lambda}}$.*

Proof. The kernel of n_λ^k is the same as the subspace annihilated by $\mathfrak{n}_\lambda^k$. But it is a straightforward exercise to see that the kernel of $\mathfrak{n}^k$ is the subspace spanned by e_j for $j \le \lambda_1 + \cdots + \lambda_k$. Therefore the kernel partition of n_λ is λ.

5. Closures

Define $\mathfrak{N}_\lambda$ to be the set of all matrices in $\mathrm{M}_n(F)$ conjugate to ν_λ. Here is the main theorem relating dominance to the closure of the algebraic varieties $\mathfrak{N}_\lambda$:

5.1. Theorem. *The set $\mathfrak{N}_\mu$ is contained in the closure of $\mathfrak{N}_\lambda$ if and only if λ dominates μ.*

There are some special cases of this that are easy to see. For the moment, if x is an array $(x_1, \ldots, x_{n-1})$ of complex numbers then let ν_x be the $n \times n$ matrix ν with

$$\nu_{i,j} = \begin{cases} x_i & \text{if } j = i+1 \\ 0 & \text{otherwise.} \end{cases}$$

If t is an invertible diagonal matrix then $t\nu_x t^{-1} = \nu_y$ with $y_i = \alpha_i(t) x_i$, where

$$\alpha_i(t) = t_{i,i}/t_{i+1,i+1} \,.$$

On the one hand, the null matrix 0 corresponds to the partition $1 + 1 + \cdots + 1$, which is dominated by every partition; and on the other conjugating by suitable t one can see that it is in the closure of every $\mathfrak{N}_\lambda$. The regular unipotent element corresponds to the singleton partition (n), which dominates all other partitions of n, and again conjugating by suitable t one can see that every $\mathfrak{N}_\lambda$ is contained in its closure.

Proof. One half of the Theorem, at least, is relatively straightforward. Any matrix in the closure of $\mathfrak{N}_\lambda$ satisfies any polynomial equation satisfied by all the matrices in $\mathfrak{N}_\lambda$. Any matrix ν in $\mathfrak{N}_\lambda$ has

$$\dim \mathrm{Ker}(\nu^k) = \kappa_k = \overline{\lambda}_1 + \cdots + \overline{\lambda}_k \,.$$

But the dimension of the kernel of a matrix is the complement of its rank, which is the largest integer m such that M possesses a non-singular $m \times m$ sub-matrix, or equivalently $\bigwedge^m M$ possesses a non-zero entry. Hence:

5.2. Lemma. *If M is any matrix then $\dim \mathrm{Ker}(M) \ge k$ if and only if $\bigwedge^{n-k+1} M = 0$.*

Therefore

$$\bigwedge^{n-\bar{\lambda}_k+1} \nu_\lambda = 0$$

for all k, and the same equation holds for all ν_μ in the Zariski closure of $\mathfrak{N}_\lambda$. But then $\overline{\mu}$ must dominate $\overline{\lambda}$ and by Lemma 2.1 λ dominates μ.

For the other half of the proof I follow, if only roughly, [Gerstenhaber:1959], which seems to be the original published source of this by now well known but little-discussed result. Gerstenhaber's argument is technically a bit obsolete, since his basis for algebraic geometry is unfortunately Weil's **Foundations**, rather than (say) Serre's *Fasceaux algébriques cohérents*. I use one observation that Gerstenhaber does not—if $\mathfrak{N}_\mu$ is contained in the closure of $\mathfrak{N}_\lambda$ and $\mathfrak{N}_\lambda$ is contained in the closure of $\mathfrak{N}_\kappa$ then $\mathfrak{N}_\mu$ is contained in the closure of $\mathfrak{N}_\kappa$, so it suffices to prove that if μ is obtained from λ by a simple shift, then some conjugate of ν_μ is contained in the closure of $\mathfrak{N}_\lambda$. Nonetheless my argument is essentially the same as his, in so far as it relies on two Lemmas:

5.3. Lemma. *If $A = \nu_\lambda$ and $\mu < \lambda$ is obtained from λ by a single shift, then there exists a nilpotent matrix B conjugate to ν_μ such that $B \cdot \mathrm{Ker}(A^k) \subseteq \mathrm{Ker}(A^{k-1})$ for all $k > 0$.*

Gerstenhaber doesn't rely on the one-step shifts, and proves in place of this Lemma the more general claim valid on the assumption merely that $\mu \le \lambda$.

5.4. Lemma. *If A and B are nilpotent matrices and $B \cdot \mathrm{Ker}(A^k) \subseteq \mathrm{Ker}(A^{k-1})$ for all $k > 0$ then A belongs to the generic class of the linear subspace spanned by A and B.*

Proof of Lemma 5.3. We just move the numbered box in accord with the shift, getting a matrix B.

1	5	8	11	13	15	17
2	6	9	12	14	16	
3	7	10				
4						

1	5	8	11	13	15	17
2	6	9	12	14		
3	7	10				
4	16					

Suppose we move box n, which occurs in column c. For all i the subspace $\mathrm{Ker}(A^i)$ is spanned by the e_m with m in columns $\le i$ of the original tableau. The two linear transformations A and B disagree only for e_n, which belongs to $\mathrm{Ker}(A^i)$ for $i \ge c$, so we must show only that $B(e_n)$ lies in $\mathrm{Ker}(A^{c-1})$. The shift moves n to an earlier column, so of course this is true. ∎

Proof of Lemma 5.4. Suppose $B \cdot \mathrm{Ker}(A^k) \subseteq \mathrm{Ker}(A^{k-1})$ for all $k > 0$. If $C = \lambda A + \mu B$ then $C^j \cdot \mathrm{Ker}(A^k) \subseteq \mathrm{Ker}(A^{k-j})$ for all $j \ge 0$, and in particular C^k annihilates $\mathrm{Ker}(A^k)$, so that $\mathrm{Ker}(C^k) \supseteq \mathrm{Ker}(A^k)$ for all $k \ge 0$. Therefore $\dim \mathrm{Ker}(A^k)$ is the minimum value of $\dim \mathrm{Ker}(C^k)$ as C ranges over the span of A and B. This is also true of the generic class of this span, so by Theorem 3.1 and Proposition 3.2 the conjugacy class of A is the same as the generic class.

But B is in the closure of the generic class, hence the proof of the Theorem is concluded. ∎

6. Flags and nilpotents

Let $\mathfrak{n}$ be the Lie algebra of upper triangular nilpotent $n \times n$ matrices. Suppose ν to be a nilpotent matrix and $\mathcal{C}_\nu$ its conjugacy class in GL_n. From Theorem 3.1 it follows that $\mathcal{C}_\nu \cap \mathfrak{n}$, whose closure is by Proposition 3.4 an algebraic cone in $\mathfrak{n}$, is not empty. Its structure is quite interesting, as we shall see. This cone is stable with respect to conjugation by elements of the Borel group B of upper triangular matrices, and it therefore gives rise to a fibre bundle over $B\backslash G$. This plays an important role in representation theory.

In this essay, a tableau is an arbitrary numbering of the boxes in a Young diagram. A **Young tableau** is a tableau in which the indices increase from left to right in any row and from top to bottom in any column. As we have seen earlier, any tableau defines a nilpotent linear transformation of $\mathbb{C}^n$, and this will be upper triangular if and only if it is a Young tableau. Thus each Young tableau T of shape λ defines an element ν_T of $\mathcal{C}_\lambda \cap \mathfrak{n}$.

6.1. Theorem. *Each irreducible component C of the algebraic variety $\mathcal{C}_\lambda \cap \mathfrak{n}$ contains a unique ν_T. The map taking C to T is a bijection of the set of irreducible components of $\mathcal{C}_\lambda$ with the Young tableaux of shape λ.*

The proof will require a serious digression. In compensation for its length, other interesting matters will come to light.

LISTING YOUNG TABLEAUX. In order for Theorem 6.1 to be practical, one must know how to list Young tableaux of a given shape. This not difficult. In a Young tableau of size n, removing the box labeled n will leave a Young tableau of size $n-1$. This gives rise to an inductive process if one knows where a box labeled n can appear. But this is very simple. A corner of a Young diagram is a box that is at the right end of a row and the bottom of a column.

6.2. Lemma. *A box labeled n in a Young tableau of size n has to appear at a corner.*

This is a basic if trivial fact about the way Young tableaux grow, and will be used later on.

So to list all tableaux of a given shape, we maintain a stack of pairs (T, m), where T is a Young diagram that is partially filled in and m is the index next to be put into the diagram. At the beginning, this stack contains only the empty diagram of the given shape and index n.

While the stack is not empty, we pop a partial tableau T and an index m off it, then list all the partial tableaux we obtain by putting m in at corners. If $m > 1$, we put all those back on the stack, but if $m = 1$ we add them to our list of completed tableaux.

FILTRATIONS OF N. For the moment, suppose the algebraic group N to be an arbitrary unipotent group, say of dimension d. It has a filtration

$$N_d = \{1\} \subset N_{d-1} \subset \ldots \subset N_1 \subset N_0 = N$$

with these properties: (a) each N_i is normal in N; (b) $(N, N_{i-1}) \subseteq N_i$; (c) each N_{i-1}/N_i is isomorphic to the additive group. Suppose that we are given homomorphisms

$$u_i\colon \mathbb{G}_a \longrightarrow N_{i-1}$$

inducing an isomorphism of $\mathbb{G}_a$ with N_{i-1}/N_i. Let U_i be the image of u_i. It can be seen easily by induction that every u in N can be expressed uniquely as a product of elements of the U_i, written in increasing order of i:

$$u = u_1(x_1)u_2(x_2)\ldots u_d(x_d)\,. \tag{6.3}$$

If one is given explicit commutation formulas

$$(u_i(s), u_j(t)) = \prod_{k>i,j} u_k(P(s,t)) \qquad (P \text{ a polynomial})$$

then it is not too difficult to come up with an algorithm that multiplies two elements of N expressed in the form (6.3). The best technique for this is *collection from the left*, as explained in [Cohen-Murray-Taylor:2005].

The filtration (N_i) is rarely canonical, and the expression in (6.3) is somewhat arbitrary. In fact:

6.4. Lemma. *Every element of N can be written as a product of elements in the U_i, in an arbitrary order.*

Proof. By induction on d. The case $d = 1$ is completely trivial. For $d = 2$, there is only one non-trivial permutation possible, and $u_2(x_2)u_1(x_1) = u_1(x_1)u_2(x_2)$ since $U_2 = N_1$ is in the center of N.

For $d > 2$, suppose a permutation σ given. Express it as the array $(\sigma(i))$. Let $\overline{\sigma}$ be the permutation of $[1, n-1]$ obtained by simply deleting n from the array, say at place i. Given u in N, apply induction to the image of u in N/N_{d-1} and $\overline{\sigma}$. This gives us an expression $u = \overline{u}\,u_d$ with u_d in U_d. But then u_d can be inserted into the expression for $\overline{u}$ at place i, since it is central.

This seems to be pretty much all one can do in arbitrary semi-simple groups, as also explained in [Cohen-Murray-Taylor:2005]. The classical semi-simple groups are those with representations of small dimension, and for those the general process can be replaced by relatively easy matrix computations. The relevant matrix computations are especially simple for SL_n, and for this group there is a curious relationship between matrix computations and the type of computation described here, as I'll now explain. I'll set this up in a slightly more general manner than I actually need, in order to relate what happens for SL_n to what happens for other groups.

ROOTS. Let A be the subgroup of diagonal matrices in GL_n. It acts by conjugation (the adjoint action) on $n \times n$ matrices. The space of all matrices is the sum of eigenspaces with respect to this action. The subspace on which A acts trivially is the subspace of diagonal matrices. Other spaces are one-dimensional, and spanned by the elementary matrices $e_{i,j}$, as shown by the equation

$$a \cdot e_{i,j} \cdot a^{-1} = (a_i/a_j) \cdot e_{i,j} \,.$$

These one-dimensional spaces are called the **root spaces** of GL_n, and the characters $\lambda_{i,j}\colon a \mapsto a_i/a_j$ are called roots. In order to simplify notation, the roots are usually written additively: $\lambda_{i,j} = \varepsilon_i - \varepsilon_j$, with $\varepsilon_i\colon a \mapsto a_i$. The **Weyl group** of this root system is the permutation group $\mathfrak{S}_n$, rearranging diagonal entries of A. It also acts dually on the lattice of characters of A, taking the subset of roots into itself.

Let $\mathfrak{n}_\lambda$ be the eigenspace corresponding to the root λ, and N_λ the corresponding subgroup of unipotent matrices $\exp(\mathfrak{n}_\lambda)$. It is made up of all matrices $u_\lambda(x) = I + xe_\lambda$. The vector space $\mathfrak{n}$ is spanned by the $e_{i,j}$ with $i < j$, and these roots are called the **positive roots**. We have the product formula

$$e_{i,j}e_{k,\ell} = \begin{cases} e_{i,\ell} & \text{if } j = k \\ 0 & \text{otherwise.} \end{cases}$$

I write $\lambda \ltimes \mu$ if $\lambda = (i, j)$ and $\mu = (j, k)$. Thus $\lambda = \mu$ is a root if and only if either $\lambda \ltimes \mu$ or $\mu \ltimes \lambda$. From the product formula we deduce the bracket formula

$$[e_\lambda, e_\mu] = \begin{cases} e_{\lambda+\mu} & \text{if } \lambda \ltimes \mu \\ -e_{\lambda+\mu} & \text{if } \mu \ltimes \lambda \\ 0 & \text{otherwise.} \end{cases}$$

An easy calculation tells us that

$$u_\lambda(x)u_\mu(y)u_\lambda^{-1}(x) = 1 + ye_\mu + xy[e_\lambda, e_\mu]$$

and from this we deduce the commutation formula

$$u_\lambda(x)u_\mu(y) = \begin{cases} u_\mu(y)u_\lambda(x)u_{\lambda+\mu}(xy) & \text{if } \lambda \ltimes \mu \\ u_\mu(y)u_\lambda(x)u_{\lambda+\mu}(-xy) & \text{if } \mu \ltimes \lambda \\ u_\mu(y)u_\lambda(x) & \text{otherwise.} \end{cases}$$

How does this discussion relate to filtrations of N? Choose a linear order on the positive roots such that $\lambda + \mu > \lambda, \mu$. Then we can filter N by the normal subgroups $N_{\geq\lambda} = \prod_{\mu \geq \lambda} N_\mu$. Several such orders are possible, but for $N = N_n$ there is one that is especially convenient.

Order the pairs (i, j) with $i < j$ from bottom to top, left to right in each row. The following figure indicates the indexing for $n = 5$.

$$\begin{bmatrix} 1 & 7 & 8 & 9 & 10 \\ & 1 & 4 & 5 & 6 \\ & & 1 & 2 & 3 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}$$

This ordering gives rise to a particularly simple product expression—any unipotent matrix $(x_{i,j})$ is the product

$$u_{n-1,n}(x_{n-1,n})u_{n-2,n-1}(x_{n-2,n-1})u_{n-2,n}(x_{n-2,n}) \dots u_{1,n}(x_{1,n})\,.$$

For example:

$$\begin{bmatrix} 1 & x_2 & x_3 \\ 0 & 1 & x_1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & x_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & x_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & x_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

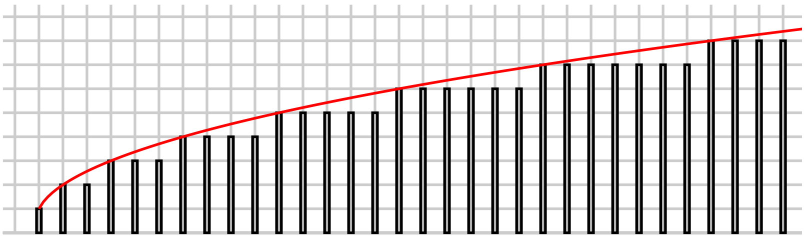
Remark. It is a mildly interesting exercise to find a formula for the location (i, j) of the root labeled by m. Thus:

$$\begin{array}{rcl} m & & (i,j) \\ 1 & \longmapsto & (n-1, n) \\ 2 & \longmapsto & (n-2, n-1) \\ 3 & \longmapsto & (n-2, n) \\ 4 & \longmapsto & (n-3, n-2) \\ 5 & \longmapsto & (n-3, n-1) \\ 6 & \longmapsto & (n-3, n) \\ & \dots & \end{array}$$

The first question is, in which row does item m appear? This will be a certain number of rows above the last, and the offset does not depend on n. Here are a few examples:

m	row offset from last
1	1
2	2
3	2
4	3
5	3
6	3
7	4
8	4
...	

Here is its graph:



After a little thought, you will see that the row for item m is

$$\varphi(m) = \left\lfloor \frac{1+\sqrt{8m-7}}{2} \right\rfloor ,$$

which is the integer-valued approximation to the inverse of the function

$$f(x) = \frac{x^2 - x + 2}{2} .$$

The graph of the function φ is shown in the figure.

The column in which item m can be found is offset from the diagonal by an amount

$$\text{horizontal offset} = m - \varphi(f(m)) + 1 .$$

FLAGS. A **flag** in the vector space V of dimension n is a sequence

$$V_0 = \{0\} \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = V$$

in which V_k/V_{k-1} has dimension one. Let $\mathfrak{B}_n$ be the algebraic variety of flags in $\mathbb{C}^n$. (I shall often ignore the subscript.) It is a projective variety on which GL_n acts transitively. If V_k is the subspace spanned by the e_i for $1 \le i \le k$, the stabilizer of the corresponding flag is the subgroup of upper triangular matrices in GL_n. The stabilizers of other flags are conjugates of this one.

Given a flag in V and a hyperplane U, one can ask, *how does the flag intersect with U?*

6.5. Lemma. *Suppose to be given a flag $V_\bullet$ in V as well a hyperplane U. There exists k such that $U \cap V_i = V_i$ for $i < k$ but $U \cap V_i$ has codimension one in V_i for $i \ge k$.*

That is to say, the dimension of $U \cap V_i$ is i for $i < k$, $i - 1$ for $i \ge k$. In particular, $U \cap V_k = U \cap V_{k-1}$—the index k is where the sequence $U \cap V_i$ **stutters**.

Proof of the Lemma. Let U be defined by the linear equation f_U. Then $f_U|V_i$ either vanishes identically or not. If it vanishes identically then it also vanishes identically on each V_j with $j < i$.

In these circumstances I can therefore define the **intersection flag** $(U_i) = V_\bullet \cap U$ (for $i < n$):

$$U_i = \begin{cases} U \cap V_i = V_i & \text{if } i < k \\ U \cap V_{i+1} & \text{otherwise.} \end{cases}$$

If the $\{v_i\}$ are a set of linearly independent vectors, I write

$$V_\bullet = \langle v_1, v_2, \ldots, v_n \rangle$$

to be the flag with

$$V_k = \text{ the span of } v_1, \ldots, v_k .$$

Sometimes, if the v_k are fixed, I'll just put down the indices. Thus $\langle 3,2,4,1\rangle$ will stand for $\langle v_3, v_2, v_4, v_1\rangle$.

For example, let $V_\bullet = \langle 1,2,3,4,5\rangle$, $U_\bullet = \langle 3,5,2,4,1\rangle$. Then here is how the Lemma works:

$$V_4 \cap U_1 = \langle 3\rangle, V_4 \cap U_2 = \langle 3\rangle, V_4 \cap U_3 = \langle 3,2\rangle, V_4 \cap U_4 = \langle 3,2,4\rangle, V_4 \cap U_5 = \langle 3,2,4,1\rangle$$

6.6. Proposition. (Bruhat decomposition) *If F and F_* are any two flags, there exists a basis (v_i) of V such that F is the flag determined by (v_i) and F_* is that determined by some $(v_{\sigma(i)})$ for some unique permutation σ of $[1,n]$.*

Another way of saying this: any two Borel subgroups contain a common maximal torus (namely, the one that has the common basis as eigenvectors).

Proof. The proof goes by induction on dimension. If $\dim V = 1$, there is no problem. Otherwise, let the two flags be $U_\bullet$ and $V_\bullet$. Apply induction to the intersection $U_\bullet \cap V_{n-1}$.

Suppose w to be in $\mathfrak{S}_n$. The group B acts on $B\backslash BwB$ on the right with isotropy group $B \cap w^{-1}Bw$, $B = TN$, and $w^{-1}Tw = T$. Therefore $BwB = BwN$. The group N acts on the right with isotropy subgroup $N^w = N \cap w^{-1}Nw$. Let

$$N_w = N \cap w^{-1}\overline{N}w = \prod_{\lambda>0, w\lambda<0} N_\lambda\,.$$

It is not necessary to specify order, according to Lemma 6.4. Also according to Lemma 6.4

$$N^w = \prod_{\lambda>0, w\lambda>0} N_\lambda\,,$$

and N is the direct product of N^w and N_w. Therefore

$$B\backslash BwB = BwN_w\,,$$

and every Borel subgroup may be uniquely expressed as gBg^{-1} with $g = wn_w$, $n_w \in N_w$. If $R_w = \{\lambda > 0 \,|\, w\lambda < 0\}$ then

$$R_{xy} = R_y \sqcup y^{-1}R_x$$

if $\ell(xy) = \ell(x) + \ell(y)$. Therefore the product map from $xN_x \times yN_y$ to xyN_{xy} is an isomorphism.

Define $B \overset{w}{\sim} B_*$ to mean that $B_* = wn_wBn_w^{-1}w^{-1}$ for some unique w in W and n_w in N_w. Given $B_0 \overset{xy}{\sim} B_2$, there exists a unique B_1 such that $B_0 \overset{x}{\sim} B_1$ and $B_1 \overset{x}{\sim} B_2$.

SPRINGER FIBERS. Given g, the subset $\mathfrak{B}_g$ of $\mathfrak{B}$ is the set of all Borel subgroups containing g.

6.7. Lemma. *Suppose u to be unipotent. Suppose (B_i) (for $0 \le i \le n$) to be a chain of Borel subgroups with each $B_i \overset{s_i}{\sim} B_{i+1}$. If B_0 and B_n are both in $\mathfrak{B}_u$, then so is each B_i.*

6.8. Lemma. *Suppose that $B \overset{s}{\sim} B_*$. If both lie in $\mathfrak{B}_u$, then so does the line in $\mathfrak{B}$ containing them.*

Remark. So $\mathfrak{B}_u$ is convex, in terms of the spherical building of G. I have the feeling that this is significant, but I don't know how.

6.9. Theorem. *For any unipotent u, the variety $\mathfrak{B}_u$ is connected.*

The next step is to assign to each flag in $\mathfrak{B}_u$ a Young tableau. The construction is basically very simple. Construct the tableau associated to $V_\bullet$ by building up the sequence of Young diagrams associated to the restrictions of ν to the V_k. According to Lemma 6.2 we do this at stage k by adding a box so as to make a corner, and we label this box with k.

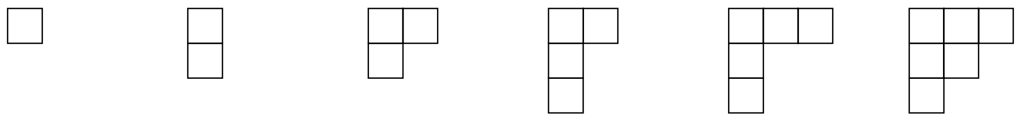
For example, suppose ν to be the Jordan matrix $\nu_{3,2,1}$:

$$\begin{gathered} 0 \leftarrow v_1 \leftarrow v_2 \leftarrow v_3 \\ 0 \leftarrow v_4 \leftarrow v_5 \\ 0 \leftarrow v_6 \end{gathered}$$

whose partition is $(3,2,1)$. Let

$$V_\bullet = \langle 6,1,2,4,3,5 \rangle\,.$$

We get the sequence of Young diagrams



and therefore the tableau

$$\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1} \end{array}$$

In order to explain what $\mathfrak{B}_\nu$ looks like, I'll look at a simple example. Let

$$\nu = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\,.$$

Its partition is $(2,1)$. We can describe $\mathfrak{B}_\nu$ rather explicitly. Suppose

$$\{0\} \subset V_1 \subset V_2 \subset \mathbb{C}^3$$

to be a ν-stable flag. There are two cases: either V_2 is contained in, hence equal to, the kernel of ν, or it is not. In the first case, V_1 may then be any line in $\mathrm{Ker}(\nu)$. In the second, it must contain a vector of the form $v = e_2 + ae_1 + be_3$, and since it is ν-stable, it must also contain $\nu(v) = e_1$. Therefore V_2 is the space spanned by e_1 and $e_2 + be_3$, and V_1 is the line containing v_1. The set $\mathfrak{B}_\nu$ is therefore partitioned into two pieces, with associated Young tableaux

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array} \qquad\qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}$$

The piece corresponding to the first tableau is closed. Both pieces are of dimension one. The first is isomorphic to projective space $\mathbb{P}^1$, and the second is isomorphic to the complement of a single point in $\mathbb{P}^1$.

In general, suppose $V_\bullet$ to be a ν-stable flag. Associate to it the sequence of Young diagrams determined by the restriction of ν to the V_k. In going from $k-1$ to k, one adds a box to the previous diagram, and it is to be numbered by k. In the end one has the Young tableau T defined by the flag $V_\bullet$. This map from flags to Young tableaux partitions $\mathfrak{B}_u$ into disjoint pieces, which are locally closed subvarieties of $\mathfrak{B}_u$.

6.10. Theorem. *Each piece of this partition has dimension*

$$\sum \frac{\overline{\lambda}_i(\overline{\lambda}_i - 1)}{2}\,.$$

The pieces thus all have the same dimension, and coincide with the irreducible components of $\mathfrak{B}_u$.

Proof. To prove the theorem, we need to know more about the structure of each piece. Suppose we are given a flag $V_\bullet$ of dimension k, and we wish to know what box is indexed by k. Let f be the linear function defining V_{k-1} as subspace of V_k, which is well defined up to a non-zero scalar. Since the flag is ν-stable, f vanishes on $\nu\, V_k$. The box labeled k will be at corner (i, j) if and only if $\mathrm{Ker}(f)$ contains $\mathrm{Ker}(\nu^{k-1})$ but does not contain $\mathrm{Ker}(\nu^k$. The space of flags in which k appears in this location therefore determine fibre over the projective space of the dual of $\nu(V) \cap \mathrm{Ker}(\nu^{k-1})$ modulo $\nu(V) \cap \mathrm{Ker}(\nu^k)$. This has dimension $\overline{\lambda}_j$, and an induction argument may then be applied. ∎

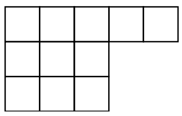
Remark. In the literature, there are two ways to a parametrize the irreducible components of $\mathfrak{B}_u$ by Young tableaux. In Spaltenstein's work, the sequence of diagrams is that of the partitions associated to the action of u on the quotients V/V_k, and in Steinberg's work associated to the restriction of ν to the V_k. These two are dual—Steinberg's sequence is that associated to $\widehat{\nu}$ by Spaltenstein's construction.

CENTRALIZERS. In order to prove Theorem 6.1, it remains to relate $\mathfrak{B}_\nu$ with $\mathcal{C}_\nu \cap \mathfrak{n}$. This will require that we know more about the centralizer $C_G(\nu)$ of u in GL_n.

I'll look at an example in detail, and then state the general result. Suppose

$$\nu = \nu_{5,3,3} = \begin{bmatrix} \nu_5 & & \\ & \nu_3 & \\ & & \nu_3 \end{bmatrix}$$

Its Young diagram is



The centralizer of ν in GL_n is the closure of its centralizer in the matrix algebra, so it suffices to find all matrices commuting with ν. We must solve the equations

$$\begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} \begin{bmatrix} \nu_5 & & \\ & \nu_3 & \\ & & \nu_3 \end{bmatrix} = \begin{bmatrix} \nu_5 & & \\ & \nu_3 & \\ & & \nu_3 \end{bmatrix} \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} .$$

This amounts to the collection of matrix equations

$$A_{i,j} x_i = x_j A_{i,j} \quad (x_1 = \nu_5,\ x_2 = \nu_3,\ x_3 = \nu_3)\,.$$

These equations amount to identifying $A_{i,j}$ with a ν-equvariant map between two modules over $\mathbb{C}[\nu]$. For example, an equation

$$A\nu_5 = \nu_3 A$$

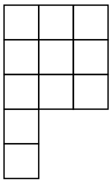
says that A is a homomorphism of $\mathbb{C}[x]/(x^5)$ to $\mathbb{C}[x]/(x^3$. This has dimension 3. So the total dimension of the commuting ring is

$$(5 + 3 + 3) + (3 + 3 + 3) + (3 + 3 + 3) = 29\,.$$

Similar reasoning shows that the commuting ring of ν_λ has dimension $\sum_{i,j} \min(\lambda_i, \lambda_j)$. This formula is fairly simple, but there is a more elegant formulation, which I leave as exercise:

6.11. Theorem. *If $|\lambda| = n$, the dimension of the commuting algebra of ν_λ in M_n is $\sum \overline{\lambda}_i^2$.*

For example:



$$29 = 3^2 + 3^2 + 3^2 + 1^2 + 1^2$$

That $Z_G(\nu)$ is an open set in the commuting ring in M_n implies:

6.12. Proosition. *The group $Z_G(\nu)$ is connected.*

CONCLUDING THE PROOF OF THE THEOREM. Map G to $\mathcal{C}_u$, taking g to gug^{-1}. The image may be identified with $G/Z_G(u)$. But having foxed B teh group G also maps to $\mathfrak{B}$. The inverse images of $\mathfrak{B}_u$ and $\mathcal{C}_u \cap \mathfrak{n}$ coincide, and since $Z_G(u)$ is connected thsi induces a bijection of the components of $\mathfrak{B}_u$ with those of $\mathcal{C}_u \cap \mathfrak{n}$.

7. *Appendix.* Some basic algorithms in linear algebra

In the main text an algorithm for finding the Jordan matrix of an arbitrary nilpotent matrix is given. It calls upon a few more elementary computations that I explain here.

The problems I deal with here are: *Given a linearly independent set of vectors in a vector space, how to extend it to a basis? Given a nilpotent matrix ν, how to find an upper triangular form for it?*

The basic tool for both is the same. I will call a matrix to be in **permuted column echelon form** if

(a) the last non-zero entry in each column, if there are any, is 1;
(b) to the left of any terminal 1 all entries are 0.

If $\overline{B}_r$ is the group of lower triangular $r \times r$ matrices then any $m \times r$ matrix may be reduced to such a form through multiplication on the right by elements of $\overline{B}_r$. This amounts to successively applying certain elementary column operations to a matrix. If $p = q\overline{b}$ is in such form then for any k the k columns of p farthest to the right span the same vector space as the same k columns of q.

If we are given a set of k linearly independent vectors of dimension n, we let Q be the $n \times k$ matrix whose columns are those vectors, adjoin the identity I on its left, and then find its reduced form, at the end the $n-k$ non-zero columns at the left will be a complementary set of linearly independent vectors.

If we start with a nilpotent square matrix n, then by first reducing it and then permuting its columns, we can find a matrix $m = n\overline{b}w$ whose first column is 0. Multiplying m on the left by $\overline{b}^{-1}w^{-1}$ will give us a matrix similar to n whose first column is still 0, since multiplication on the left amounts to row operations. Applying this procedure recursively gives us an upper triangular matrix similar to n.

Part III. The other classical groups

8. Introduction

If H is a non-degenerate $n \times n$ symmetric or skew-symmetric matrix, one can associate to it the group $G(H)$ of all $n \times n$ matrices X such that

(8.1)
$$ {}^t\!XHX = H\,. $$

If H is symmetric this is the orthogonal group $\mathrm{O}(H)$. The special orthogonal group is the subgroup in which $\det = 1$. If H is a skew-symmetric $2n \times 2n$ matrix, this is the symplectic group $\mathrm{Sp}(H)$. Since $\mathrm{Sp}(H)$ also preserves all the bilinear forms $\bigwedge^m H$, all matrices in $\mathrm{Sp}(H)$ have determinant 1.

In all cases, the corresponding Lie algebra $\mathfrak{g}(H)$ is that of all matrices X such that

(8.2) .
$$ {}^t\!XH + HX = 0\,. $$

In this section I shall classify all nilpotent matrices in the Lie algebras of these groups (and, equivalently, all unipotent matrices in the groups themselves). Each of these groups G is defined as a subgroup of SL_m ($m = 2n$ or $2n+1$). Any nilpotent matrix in $\mathfrak{g}$ is also one in $\mathfrak{sl}_m$, hence corresponds to a partition of m. It turns out that, with some few exceptions, two nilpotent matrices in $\mathfrak{g}$ are conjugate under G if and only if they are conjugate under SL_m in $\mathfrak{sl}_m$, so the classification of nilpotents basically reduces to the problem of saying which partitions occur. It also turns out, miraculously, that almost always one nilpotent conjugacy class in $\mathfrak{g}$, with respect to G, lies in the proper closure of another if and only if the same is true in $\mathfrak{sl}_m$, with respect to GL_m. In the first version of this essay, I'll only show explicitly which partitions occur, and otherwise skip proofs.

Define the $n \times n$ matrix

$$\omega_n = \begin{bmatrix} \circ & \circ & \dots & \circ & 1 \\ \circ & \circ & \dots & 1 & \circ \\ & & \dots & & \\ \circ & 1 & \dots & \circ & \circ \\ 1 & \circ & \dots & \circ & \circ \end{bmatrix} .$$

Sometimes I'll forget the subscript. The matrices of the forms that G leaves invariant will be one of

$$\begin{bmatrix} 0 & -\omega_n \\ \omega_n & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \omega_n \\ \omega_n & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & \omega_n \\ 0 & 1 & 0 \\ \omega_n & 0 & 0 \end{bmatrix} .$$

The algebraic variety of nilpotent elements in any semi-simple complex Lie algebra contains a unique open, dense conjugacy class of **principal** nilpotents. These are regular in any one of several senses, most simply because the dimension of the stabilizer of any of them has dimension equal to the minimal possible value, the semi-simple rank of the group.

In the classical groups under consideration here, every nilpotent element of $\mathfrak{g}$ is the sum of principal nilpotents in a direct sum of relatively simple subalgebras (possibly $\mathfrak{g}$ itself). As we shall see, there are two basic types.

The first type depends on an embedding of GL_n into all groups Sp_{2n}, SO_{2n}, or SO_{2n+1}:

$$A \longmapsto \begin{bmatrix} A & 0 \\ 0 & \omega\, {}^t\!A^{-1}\omega \end{bmatrix}$$

or

$$A \longmapsto \begin{bmatrix} A & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \omega\, {}^t\!A^{-1}\omega \end{bmatrix} .$$

It will be convenient to keep in mind that $\omega\, {}^t\!X\omega$ is the reflection of the matrix X along its SW-NE axis.

The second type depends on a decomposition of the given invariant form on $\mathbb{C}^m$ into direct sums of similar forms on smaller subspaces, as we shall see.

9. The symplectic group

All nondegenerate symplectic forms are equivalent, so deciding which one to use is a matter of convenience. I repeat that my choice is

$$H = \begin{bmatrix} 0 & -\omega_n \\ \omega_n & 0 \end{bmatrix} .$$

Before discussing conjugacy classes, I first recall a few basic facts about symplectic groups. A bit more explicitly, the equations defining the Lie algebra become

$$\begin{bmatrix} {}^tA & {}^tC \\ {}^tB & {}^tD \end{bmatrix} \begin{bmatrix} 0 & -\omega_n \\ \omega_n & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\omega_n \\ \omega_n & 0 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = 0 .$$

The twisting by ω can be a minor nuisance in calculation, but the point of this choice of H is that the upper triangular matrices in Sp_{2n} are a Borel subgroup. In particular, as I have already noted, a matrix

$$\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$$

lies in $\mathfrak{sp}_{2n}$ if and only if $D = -\omega\, {}^t\!A \omega^{-1}$. This means that $\mathfrak{gl}_n$ embeds into $\mathfrak{sp}_{2n}$. Also, GL_n embeds into Sp_{2n}, taking

$$A \longmapsto \begin{bmatrix} A & 0 \\ 0 & \omega\, {}^t\!A^{-1}\omega^{-1} \end{bmatrix} .$$

A matrix

$$\begin{bmatrix} 0 & X \\ 0 & 0 \end{bmatrix}$$

lies in $\mathfrak{sp}_{2n}$ if and only if $X = \omega\, {}^t\!X \omega^{-1}$.

9.1. Proposition. *(a) A partition (λ_i) is that of a conjugacy class in $\mathfrak{sp}_{2n}$ if and only if only odd values of λ_i occur an even number of times. (b) Two matrices in $\mathfrak{sp}_{2n}$ are conjugate with respect to Sp_{2n} if and only if they are conjugate with respect to GL_{2n}.*

Proof. SUFFICIENCY. We have seen that GL_n embeds *via* two copies into Sp_{2n}. Hence any pair (k, k) embeds through this embedding, If $n = k + \ell$ we may express

$$\begin{bmatrix} 0 & -\omega_n \\ \omega_n & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -\omega_k \\ 0 & 0 & -\omega_\ell & 0 \\ 0 & \omega_\ell & 0 & 0 \\ \omega_k & 0 & 0 & 0 \end{bmatrix} .$$

Hence $\mathrm{Sp}_{2k} \times \mathrm{Sp}_{2\ell}$ can be embedded into Sp_{2n}:

$$\begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} \times \begin{bmatrix} A_\ell & B_\ell \\ C_\ell & D_\ell \end{bmatrix} \longmapsto \begin{bmatrix} A_k & 0 & 0 & B_k \\ 0 & A_\ell & B_\ell & 0 \\ 0 & C_\ell & D_\ell & 0 \\ C_\ell & 0 & 0 & D_\ell \end{bmatrix} .$$

It remains only to show that any even singleton $\nu_{(2n)}$ can be embedded into Sp_{2n}. For this: the matrix

$$\begin{bmatrix} \Lambda_n & \lambda_{n,n} \\ 0 & -\Lambda_n \end{bmatrix}$$

lies in $\mathfrak{sp}_{2n}$ if

$$\Lambda_n = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} \quad (n \times n)$$

and

$$\lambda_{p,q} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} \quad (p \times q) .$$

Example. Look at Sp_4. If

$$X = \begin{bmatrix} \circ & 1 & a & b \\ \circ & \circ & c & a \\ \circ & \circ & \circ & -1 \\ \circ & \circ & \circ & \circ \end{bmatrix}$$

then

$$
\begin{aligned}
X^2 &= \begin{bmatrix} \circ & \circ & c & \circ \\ \circ & \circ & \circ & -c \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix} \\
X^3 &= \begin{bmatrix} \circ & \circ & \circ & -c \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix} \\
X^4 &= 0\,.
\end{aligned}
$$

There are four allowable partitions, with conjugacy class representatives as follows:

$$
\begin{aligned}
(4) &: \begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & 1 & \circ \\ \circ & \circ & \circ & -1 \\ \circ & \circ & \circ & \circ \end{bmatrix} \\
(2,2) &: \begin{bmatrix} \circ & 1 & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & -1 \\ \circ & \circ & \circ & \circ \end{bmatrix} \\
(2,1,1) &: \begin{bmatrix} \circ & \circ & \circ & \circ \\ \circ & \circ & 1 & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix} \\
(1,1,1,1) &: 0\,.
\end{aligned}
$$

NECESSITY. Skipped, at least in this version.

There is something more to be said. There is a classification of nilpotent elements of all semi-simple Lie algebras due to Bala and Carter in which the notion of a **distinguished nilpotent elements** plays a role. A distinguished element is one that does not meet any proper Levi factor of a parabolic subgroup. The distinguished nilpotent matrices of $\mathfrak{sp}_{2n}$ correspond to partitions with no doubles.

10. The orthogonal groups

All quadrtaic forms of a given dimension over $\mathbb{C}$ are isomorphic, and sometimes it is useful to keep this in mind. Gerally, I take the group SO_{2n} to be the special orthogonal group of

$$
H_{2n} = \begin{bmatrix} 0 & \omega_n \\ \omega_n & 0 \end{bmatrix}\,.
$$

The group SO_{2n+1} is the special orthogonal group of

$$
H_{2n+1} = \begin{bmatrix} 0 & 0 & \omega_n \\ 0 & 1 & 0 \\ \omega_n & 0 & 0 \end{bmatrix}\,.
$$

form a Borel subgroup, and parabolic subgroups are easy to parametrize.

In $\mathfrak{so}_{2n}$ we have matrices

$$
\begin{bmatrix} A & C \\ 0 & -\omega\, {}^tA\omega \end{bmatrix}
$$

and in $\mathfrak{so}_{2n+1}$

$$\begin{bmatrix} A & b & C \\ 0 & d & -\omega\,{}^t b \\ 0 & 0 & -\omega\,{}^t A\omega \end{bmatrix}$$

in which b is a column vector, d a scalar, and $C = -\omega\,{}^t C\omega$.

Since every quadratic form is also equivalent to a sum of squares, $\mathfrak{so}_k \oplus \mathfrak{so}_\ell$ can be embedded into $\mathfrak{so}_{k+\ell}$. It is also true that $\mathfrak{so}_{2k+1}$ is embedded into $\mathfrak{so}_{2(k+1)}$. Exactly how does this relate to my particular choice of form?

Roughly speaking, this is because $x^2 - y^2 = (x-y)(x+y)$, but we'll need an explicit embedding. If ${}^tSQ_1S = Q_2$ and ${}^tXQ_2X = Q_2$ then

$${}^tX\,{}^tSQ_1SX = {}^tSQ_1S\,,$$

so that the map $X \mapsto SXS^{-1}$ takes $O(Q_2)$ to $O(Q_1)$. In our case

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Therefore

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \omega \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \omega & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \omega \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ \omega & 0 & 0 & 0 \end{bmatrix}$$

so that conjugation by

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & I \end{bmatrix}$$

embeds $\mathfrak{so}_{2n+1}$ into $\mathfrak{so}_{2n+2}$.

10.1. Proposition. *(a) A partition (λ_i) is that of a conjugacy class in $\mathfrak{so}_m$ if and only if only even values of λ_i occur an even number of times. (b) Unless a partition is totally even, every conjugacy class in $\mathrm{M}_m \cap \mathfrak{so}_m$ is a single class in $\mathfrak{so}_m$. A class corresponding to a totally even partition is the union of two distinct classes with respect to $\mathrm{SO}(H)$ in $\mathfrak{so}_m$. These become conjugate with respect to $\mathrm{O}(H)$.*

I. e. only odd partitions are allowed to appear as singletons. Two distinct classes are swapped by the involution corresponding to conjugation by reflections in O_m.

Proof. SUFFICIENCY. Again here pairs (k, k) may be embedded through the two copies of GL_n, and $\mathfrak{so}_k \oplus \mathfrak{so}_\ell$ may be embedded into $so_{k+\ell}$, so it suffices to show that $\nu_{(2n+1)}$ may be placed in $\mathrm{SO}(2n+1)$. We can check that

$$\begin{bmatrix} \Lambda_{2n+1} & \lambda_{2n+1,2n} \\ 0 & -\Lambda_{2n} \end{bmatrix}$$

lies in SO_{2n+1}. For example, SO_5 contains the class

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

But for even dimensions something a bit different occurs. All nilpotent matrices corresponding to the same partition are conjugate with respect to the full orthogonal group O_{2n}, but not with respect to the smaller group SO_{2n}. The quotient $\mathrm{O}_{2n}/\mathrm{SO}_{2n}$ acts as a non-trivial outer automorphism on $\mathfrak{sp}_{2n}$, taking

NECESSITY. Also skipped.

The distinguished nilpotent matrices of $\mathfrak{so}_m$, as with $\mathfrak{sp}_{2n}$, correspond to partitions with no doubles.

Part IV. References

1. Thomas Brylawski, 'The lattice of integer partitions', *Discrete Mathematics* **6** (1973), 201–219.

2. Roger Carter, **Finite groups of Lie type: conjugacy classes and complex characters**, Wiley, 1985.

3. Arjeh Cohen, Scott Murray, and Don Taylor, 'Computing in groups of Lie type', *Mathematics of Computation* **73** (2004), 1477–1498.

4. David Collingwood and William McGovern, **Nilpotent orbits in semi-simple Lie algebras**, van Nostrand-Reinhold, 1993.

5. Murray Gerstenhaber, 'On nilalgebras and linear varieties of nilpotent matrices, III', *Annals of Mathematics* **70** (1959), 167–182.

6. ——, 'Dominance over the classical groups', *Annals of Mathematics* **74** (1961), 532–569.

7. G. H. Hardy, J. E. Littlewood, and G. Polya, **Inequalities**, Cambridge University Press, 1952.

8. George Lusztig and Nicolas Spaltenstein, 'Induced unipotent classes', *Journal of the London Mathematical Society* **19** (1979), 41–52.

9. Nicolas Spaltenstein, 'On the fixed point set of a unipotent transformation on the flag manifold', *Indagatione Mathematicae* **38** (1976), 452–456.

10. Robert Steinberg, 'An occurrence of the Robinson-Schensted correspondence', *Journal of Algebra* **113** (1988), 523–28.