Week 1

Conjugacy classes.

Throughout, $\mathbb{C}$ will be the field of complex numbers. Recall that $\mathbb{C}$ is *algebraically closed* and is equipped with a *norm* $\mid a \mid = a\overline{a}$ (here, $\overline{a}$ is the *complex conjugate* of the complex number a) making $\mathbb{C}$ into a topological space.

Let V be a finite dimensional $\mathbb{C}$-vectorspace, say of dimension d, then after choosing a *basis* we can identify V with the space of d-tuples $\mathbb{C}^d$. As such we can equip V with the *analytic topology* induced by the *metric* $d(v,w) = \mid v - w \mid$ coming from the norm

$$\mid v \mid = \max \mid v_i \mid$$

for any $v = (v_1, \ldots, v_d) \in V = \mathbb{C}^d$. In this chapter we will be primarily interested in the analytic topology induced on closed subsets of some V.

With GL_n we denote the group of all invertible $n \times n$ matrices $GL_n(\mathbb{C})$ with coefficients in $\mathbb{C}$. As an $n \times n$ matrix $A \in M_n(\mathbb{C})$ is invertible if and only if its determinant $det(A)$ is non-zero, we see that GL_n is a dense open subset of the n^2-dimensional vectorspace $M_n = M_n(\mathbb{C})$.

1a. Conjugacy classes of matrices.

An $n \times n$ matrix $A \in M_n$ is by left multiplication a linear operator on the n-dimensional vectorspace $V_n = \mathbb{C}^n$ of *column vectors*. If $g \in GL_n$ is the matrix describing the *base change* from the canonical basis of V_n to a new basis, then the linear operator expressed in this new basis is represented by the matrix gAg^{-1}. For a given matrix A we want to find an adapted basis such that the *conjugated matrix* gAg^{-1} has a simple form.

That is, we consider the linear action of GL_n on the n^2-dimensional vectorspace M_n of $n \times n$ matrices determined by

$$GL_n \times M_n \longrightarrow M_n \qquad (g, A) \mapsto g.A = gAg^{-1}.$$

The *orbit* $\mathcal{O}_A$ of A under this action, that is the set of all matrices of the form gAg^{-1} for some $g \in GL_n$, is called the *conjugacy class* of A. We look for a particularly nice *representant* in a given conjugacy class. That is, we want to solve the following orbit space problem.

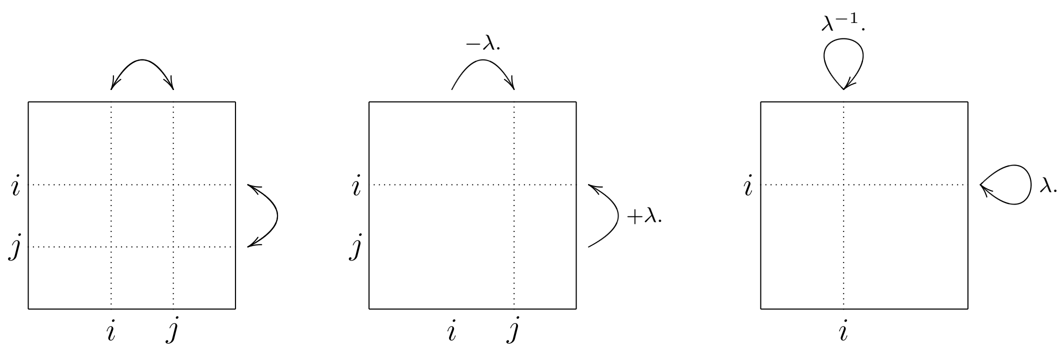
Problem 1.

Classify the conjugacy classes of $n \times n$ matrices.

With $\mathbb{1}_n$ we denote the identity matrix in M_n and with e_{ij} the matrix whose unique non-zero entry is 1 at entry (i, j). Recall that the group GL_n is generated by the following three classes of matrices :

- the *permutation* matrices $p_{ij} = \mathbb{1}_n + e_{ij} + e_{ji} - e_{ii} - e_{jj}$ for all $i \neq j$,
- the *addition* matrices $a_{ij}(\lambda) = \mathbb{1}_n + \lambda e_{ij}$ for all $i \neq j$ and $0 \neq \lambda$, and
- the *multiplication* matrices $m_i(\lambda) = \mathbb{1}_n + (\lambda - 1)e_{ii}$ for all i and $0 \neq \lambda$.

Conjugation by these matrices determine the three types of *Jordan moves* on $n \times n$ matrices, where the altered rows and columns are dashed :



type p type a type m

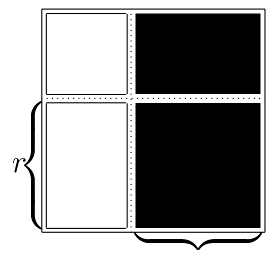
Therefore, it suffices to consider sequences of these moves on a given $n \times n$ matrix $A \in M_n$. The *characteristic polynomial* of A is defined to be the polynomial of degree n in the variable t

$$\chi_A(t) = det(A - t\mathbb{1}_n) \in \mathbb{C}[t].$$

As $\mathbb{C}$ is algebraically closed, $\chi_A(t)$ decomposes as a product of linear terms $\prod_{i=1}^{e}(t - \lambda_i)^{d_i}$ where the $\{\lambda_1, \ldots, \lambda_e\}$ are called the *eigenvalues* of the matrix A. Observe that λ_i is an eigenvalue of A if and only if there is a non-zero *eigenvector* $v \in V_n = \mathbb{C}^n$ with eigenvalue λ_i, that is, $A.v = \lambda_i v$. In particular, the *rank* r_i of the matrix $A_i = \lambda_i \mathbb{1}_n - A$ satisfies $n - d_i \leq r_i < n$.

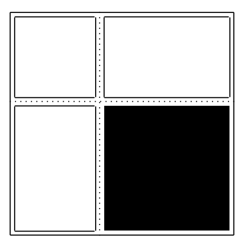
We will apply the following reduction step to the matrices A_i. Let B be a $n \times n$ matrix of rank r. Then, applying type p and type a Jordan moves we can conjugate B to the following block form

$$\begin{bmatrix} 0_{n-r} & \mathrm{top}_1 \\ 0 & \mathrm{bot}_1 \end{bmatrix} \quad =$$

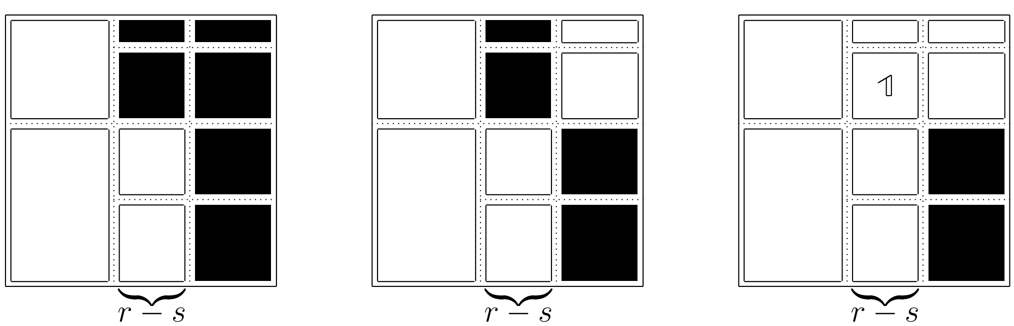


where white blocks denote zero matrices, hence the rank of the (black) right hand side is equal to r. We separate two cases. First, assume that the square $r \times r$ bottom right matrix has rank r (is invertible). Then, the rows of the upper right block are linear combinations of its rows. Then, we apply type a Jordan moves (which do not spoil the zero blocks) and conjugate the matrix to the block form

$$\begin{bmatrix} 0_{n-r} & \\ & \mathrm{bot}_1 \end{bmatrix} \quad =$$



The top left zero block of size $n-r$ splits off and we stop the reduction. In the second case, assume the square $r \times r$ matrix has rank $s < r$ (hence the upper right block has rank $r - s \leq n - r$). Repeating the above reduction to the $r \times r$ block we obtain the situations :



The two bottom right blocks together have rank s, whence all rows of the two upper right blocks are linear combinations of them. Using type a Jordan moves (which preserve the zero blocks) we arrive at the middle block decomposition. Here, the two middle black blocks together have rank $r - s$. Using type p and type m Jordan moves (which preserve the other zero blocks) we can conjugate to obtain the rightmost block decomposition, where the $\mathbb{1}$-block is the identity matrix $\mathbb{1}_{r-s} \in M_{r-s}$. That is, we obtain a matrix of block form

$$\begin{bmatrix} 0_{n+s-2r} & & & \\ & 0_{r-s} & \mathbb{1}_{r-s} & \\ & 0_{r-s} & 0_{r-s} & \text{top}_2 \\ & & & \text{bot}_2 \end{bmatrix}$$

and the top left zero block of size $n+s-2r$ splits off. If the bottom block bot_2 has rank s we can conjugate to make top_2 a zero block in which case also the *Kronecker product* block

$$J_2 \otimes \mathbb{1}_{r-s} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \otimes \mathbb{1}_{r-s} = \begin{bmatrix} 0_{r-s} & \mathbb{1}_{r-s} \\ 0_{r-s} & 0_{r-s} \end{bmatrix}$$

splits off and we stop the reduction. Otherwise, the square $s \times s$ block bot_2 has rank $t < s$ and we repeat the above reduction

$$\begin{bmatrix} 0_{n-r} & \text{top}_1 \\ 0 & \text{bot}_1 \end{bmatrix} \begin{matrix} \scriptstyle \text{r-s} \\ \scriptstyle \text{s} \end{matrix} \xrightarrow{\text{moves}} \begin{bmatrix} 0_{n+s-2r} & & & \\ & 0_{r-s} & \mathbb{1}_{r-s} & \\ & 0_{r-s} & 0_{r-s} & \text{top}_2 \\ & & 0 & \text{bot}_2 \end{bmatrix} \begin{matrix} \\ \\ \scriptstyle \text{s-t} \\ \scriptstyle \text{t} \end{matrix}$$

to the new top/bottom block (the tiny integers give the ranks of the top_i and bot_i blocks). This gives us the following block matrix

$$\begin{bmatrix} 0_{\text{n+s-2r}} & & & & & & \\ & 0_{\text{r+t-2s}} & 0 & \mathbb{1}_{\text{r+t-2s}} & 0 & & \\ & 0 & 0_{\text{s-t}} & 0 & \mathbb{1}_{\text{s-t}} & & \\ & 0_{\text{r+t-2s}} & 0 & 0_{\text{r+t-2s}} & 0 & & \\ & 0 & 0_{\text{s-t}} & 0 & 0_{\text{s-t}} & \mathbb{1}_{\text{s-t}} & \\ & & & & 0_{\text{s-t}} & 0_{\text{s-t}} & \text{top}_3 \\ & & & & & 0 & \text{bot}_3 \end{bmatrix}$$

which after some permutation Jordan moves can be brought into the form

$$\begin{bmatrix} 0_{\text{n+s-2r}} & & & & & & \\ & 0_{\text{r+t-2s}} & \mathbb{1}_{\text{r+t-2s}} & & & & \\ & 0_{\text{r+t-2s}} & 0_{\text{r+t-2s}} & & & & \\ & & & 0_{\text{s-t}} & \mathbb{1}_{\text{s-t}} & 0_{\text{s-t}} & \\ & & & 0_{\text{s-t}} & 0_{\text{s-t}} & \mathbb{1}_{\text{s-t}} & \\ & & & 0_{\text{s-t}} & 0_{\text{s-t}} & 0_{\text{s-t}} & \text{top}_3 \\ & & & & & 0 & \text{bot}_3 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \scriptstyle \text{t-u} \\ \scriptstyle \text{u} \end{matrix}$$

Hence, the block $J_1 \otimes \mathbb{1}_{r+t-2s}$ splits off and if bot_3 is of rank t, so does the newly created Kronecker product

$$J_3 \otimes \mathbb{1}_{s-t} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \otimes \mathbb{1}_{s-t}$$

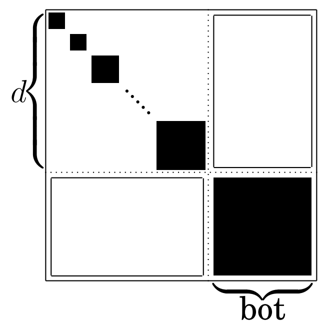
at which stage we stop the reduction. If the rank of bot_3 is $u < t$, then after the next cycle we will be able to split off the Kronecker product $J_3 \otimes \mathbb{1}_{s+u-2t}$ and create a new 'almost'-split Kronecker product $J_4 \otimes \mathbb{1}_{t-u}$.

It is now clear that after a finite number of cycles of this reduction process we will conjugate our $n \times n$ matrix B to a matrix in diagonal block form

$$(J_1 \otimes \mathbb{1}_{\text{n+s-2r}}) \oplus (J_2 \otimes \mathbb{1}_{\text{r+t-2s}}) \oplus (J_3 \otimes \mathbb{1}_{\text{s+u-2t}}) \oplus \ldots \oplus (J_m \otimes \mathbb{1}_{\text{y-z}}) \oplus \text{bot}_{m+1}$$

with bot_{m+1} invertible or of size zero. Here, J_k is the *Jordan block* matrix of size k with zeroes everywhere except ones on the next to main diagonal.

Finally, again by permutation moves we can conjugate B to the block diagonal matrix



Here, the bottom right corner bot is invertible, hence has all its eigenvalues $\{\mu_1, \mu_2, \ldots\}$ nonzero, and all the diagonal blocks in the upper left $d \times d$ corner are Jordan blocks J_k (there are $n+s-2r$ blocks J_1, $r+t-2s$ blocks J_2 etc.). The integer d is determined as the maximal power such that t^d divides the characteristic polynomial $\chi_B(t)$. Hence, the sizes of these Jordan blocks

$$p = (\underbrace{m, \ldots, m}_{\text{y-z}}, \ldots, \underbrace{3, \ldots, 3}_{\text{s+u-2t}}, \underbrace{2, \ldots, 2}_{\text{r+t-2s}}, \underbrace{1, \ldots, 1}_{\text{n+s-2r}})$$

form a *partition* of d.

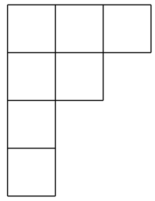
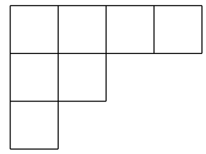
Recall that a partition $p = (a_1, a_2, \ldots, a_k)$ of d is a decompositions in natural numbers

$$d = a_1 + a_2 + \ldots + a_k \quad \text{with} \quad a_1 \geq a_2 \geq \ldots \geq a_k \geq 1$$

It is traditional to assign to a partition $p = (a_1, \ldots, a_k)$ a *Young diagram* with a_i boxes in the i-th row, the rows of boxes lined up to the left.

The *dual partition* $p^* = (a_1^*, \ldots, a_r^*)$ to p is defined by interchanging rows and columns in the Young diagram of p. For example, to the partition $p =$

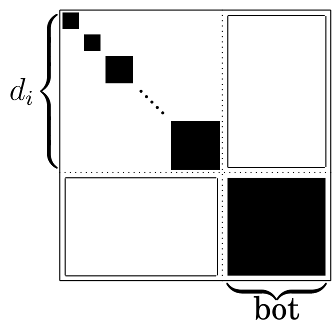
$(3,2,1,1)$ of 7 we assign the Young diagram

$p =$  $p^* =$ 

with dual partition $p^* = (4,2,1)$.

1b. The Jordan-Weierstrass theorem.

Let us return to an arbitrary $n \times n$ matrix A with characteristic polynomial $\chi_A(t) = \prod_{i=1}^e (t-\lambda_i)^{d_i}$. Apply the above reduction to the matrix $B = A_i = \lambda_i \mathbb{1} - A$. Then, A itself is conjugated to a block diagonal matrix of the form



with all the blocks in the $d_i \times d_i$ upper left corner *Jordan blocks with eigenvalue* λ_i, that is of the form

$$J_k(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in M_k(\mathbb{C})$$

and the remaining block bot has eigenvalues $\{\lambda_i, \ldots, \check{\lambda}_i, \ldots, \lambda_e\}$. Repeating this procedure for the other eigenvalues we obtain the *Jordan-Weierstrass theorem*.

Theorem 1.1. *Let $A \in M_n(\mathbb{C})$ with characteristic polynomial $\chi_A(t) = \prod_{i=1}^e (t-\lambda_i)^{d_i}$. Then, A determines unique partitions*

$$p_i = (a_{i1}, a_{i2}, \ldots, a_{im_i}) \quad \textit{of} \quad d_i$$

for $1 \leq i \leq e$ *such that* A *is conjugated to a unique (up to permutation of the blocks) block-diagonal matrix*

$$J_{(p_1,\ldots,p_e)} = \begin{bmatrix} \boxed{B_1} & & & \\ & \boxed{B_2} & & \\ & & \ddots & \\ & & & \boxed{B_m} \end{bmatrix}$$

with $m = m_1 + \ldots + m_e$ *and exactly one block* B_l *of the form* $J_{a_{ij}}(\lambda_i)$ *for all* $1 \leq i \leq e$ *and* $1 \leq j \leq m_i$.

Remains only to prove the unicity of the partitions p_i of d_i corresponding to the eigenvalue λ_i of A. Assume A is conjugated to another Jordan block matrix $J_{(q_1,\ldots,q_e)}$, necessarily with partitions $q_i = (b_{i1}, \ldots, b_{im'_i})$ of d_i. To begin, observe that for a Jordan block of size k we have that

$$rk\ J_k(0)^l = k - l \quad \text{for all } l \leq k \text{ and if } \mu \neq 0 \text{ then} \quad rk\ J_k(\mu)^l = k$$

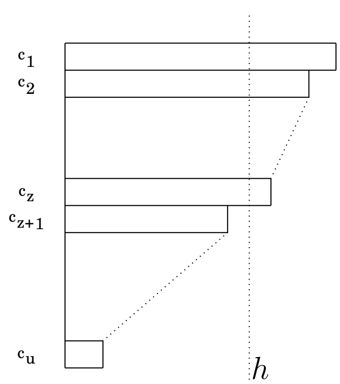
for all l. As $J_{(p_1,\ldots,p_e)}$ is conjugated to $J_{(q_1,\ldots,q_e)}$ we have for all $\lambda \in \mathbb{C}$ and all l

$$rk\ (\lambda \mathbb{1}_n - J_{(p_1,\ldots,p_e)})^l = rk\ (\lambda \mathbb{1}_n - J_{(q_1,\ldots,q_e)})^l$$

Now, take $\lambda = \lambda_i$ then only on the Jordan blocks with eigenvalue λ_i are important in the calculation and one obtains for the ranks eqrank

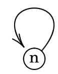
$$n - \sum_{h=1}^{l} \#\{j \mid a_{ij} \geq h\} \quad \text{respectively} \quad n - \sum_{h=1}^{l} \#\{j \mid b_{ij} \geq h\}. \tag{1.1}$$

Now, for any partition $p = (c_1, \ldots, c_u)$ and any natural number h we see that the number $z = \#\{j \mid c_j \geq h\}$



is the number of blocks in the h-th row of the dual partition p^*. Therefore, the above rank equality implies that $p_i^* = q_i^*$ and hence that $p_i = q_i$. As we can repeat this argument for the other eigenvalues we have the required unicity.

This completes the classification of the conjugacy classes of $n \times n$ matrices, or equivalently, the GL_n-orbits in M_n which (for later reference) corresponds to the *pattern* :



We see that the classification consists of two parts : a *discrete* part choosing

- a partition $p = (d_1, d_2, \ldots, d_e)$ of n, and for each d_i,
- a partition $p_i = (a_{i1}, a_{i2}, \ldots, a_{im_i})$ of d_i,

determining the sizes of the Jordan blocks and a *continuous* part choosing

- an e-tuple of distinct complex numbers $(\lambda_1, \lambda_2, \ldots, \lambda_e)$.

fixing the eigenvalues. Moreover, this e-tuple $(\lambda_1, \ldots, \lambda_e)$ is determined only up to permutations of the subgroup G of all permutations S_e on e letters where

$$G = \{\pi \in S_p \mid p_i = p_{\pi(i)} \text{ for all } 1 \leq i \leq e\}$$

Whereas this gives a satisfactory *set-theoretic* description of the orbits, one might ask for a topological *orbit space* C_n the points of which are in one-to-one correspondence with the conjugacy classes, and a continuous surjection

$$M_n \overset{c}{\twoheadrightarrow} C_n$$

which is constant along GL_n-orbits and sends a matrix A to the point of C_n corresponding to the orbit $\mathcal{O}_A$. If we require that this space C_n has at least the separation property that its points should be closed, then continuity of c implies that for any matrix A its conjugacy class $\mathcal{O}_A$ should be a closed subset of M_n.

However, this cannot be the case whenever $n \geq 2$. Consider the matrices

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

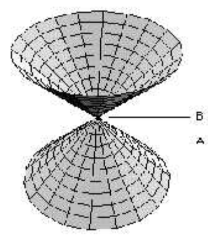
which by theorem 1.1. belong to distinct orbits. For any $\epsilon \neq 0$ we have that

$$\begin{bmatrix} \epsilon & 0 \\ 0 & 1 \end{bmatrix} . \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} . \begin{bmatrix} \epsilon^{-1} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda & \epsilon \\ 0 & \lambda \end{bmatrix}$$

belongs to the orbit of A. Hence if $\epsilon \mapsto 0$, we see that B lies in the closure of $\mathcal{O}_A$ whence $\mathcal{O}_A$ cannot be a closed orbit in M_2. As any matrix in $\mathcal{O}_A$ has trace 2λ, the orbit is contained in the 3-dimensional subspace

$$\begin{bmatrix} \lambda + x & y \\ z & \lambda - x \end{bmatrix} \hookrightarrow M_2$$

In this space, the orbit-closure $\overline{\mathcal{O}_A}$ is the set of points satisfying $x^2 + yz = 0$ (the determinant has to be λ^2), which is a cone having the origin as its top :



The orbit $\mathcal{O}_B$ is the top of the cone and the orbit $\mathcal{O}_A$ is the complement.

Week 2

The quotient space.

Last week we have seen that there is no Hausdorff topological space whose points are in one-to-one correspondence with the conjugacy classes of matrices. Still, we can try to solve :

Problem 2.

Construct the best continuous approximation to the orbit space.

First, we construct a supply of complex valued continuous functions on M_n that are constant along orbits.

2a. Invariant polynomial functions.

If two matrices are conjugated $A \sim B$, then A and B have the same *unordered* n-tuple of eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ (occurring with multiplicities). Hence any symmetric function in the λ_i will have the same values in A as in B. In particular this is the case for the *elementary symmetric functions* σ_l

$$\sigma_l(\lambda_1, \ldots, \lambda_l) = \sum_{i_1 < i_2 < \ldots < i_l} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_l}.$$

Observe that for every $A \in M_n$ with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ we have

$$\prod_{j=1}^{n} (t - \lambda_j) = \chi_A(t) = det(t\mathbb{1}_n - A) = t^n + \sum_{i=1}^{n} (-1)^i \sigma_i(A) t^{n-i}$$

Developing the determinant $det(t\mathbb{1}_n - A)$ we see that each of the coefficients $\sigma_i(A)$ is in fact a *polynomial function* in the entries of A. A fortiori, $\sigma_i(A)$ is a complex valued continuous function on M_n. The above equality also implies that the functions $\sigma_i : M_n \longrightarrow \mathbb{C}$ are constant along orbits. We now construct the continuous map

$$M_n \xrightarrow{\pi} \mathbb{C}^n$$

sending a matrix $A \in M_n$ to the point $(\sigma_1(A), \ldots, \sigma_n(A))$ in $\mathbb{C}^n$. Clearly, if $A \sim B$ then they map to the same point in $\mathbb{C}^n$. We claim that π is surjective. Take any point $(a_1, \ldots, a_n) \in \mathbb{C}^n$ and consider the matrix $A \in M_n$

$$A = \begin{bmatrix} 0 & & & & a_n \\ -1 & 0 & & & a_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & -1 & 0 & a_2 \\ & & & -1 & a_1 \end{bmatrix} \tag{2.1}$$

then we will show that $\pi(A) = (a_1, \ldots, a_n)$, that is,

$$det(t\mathbb{1}_n - A) = t^n - a_1 t^{n-1} + a_2 t^{n-2} - \ldots + (-1)^n a_n.$$

Indeed, developing the determinant of $t\mathbb{1}_n - A$ along the first column we obtain

$$\begin{vmatrix} t & 0 & 0 & \cdots & 0 & -a_n \\ 1 & t & 0 & & 0 & -a_{n-1} \\ 0 & 1 & t & & 0 & -a_{n-2} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & 1 & t & -a_2 \\ 0 & 0 & & & 1 & t-a_1 \end{vmatrix} - \begin{vmatrix} t & 0 & 0 & \cdots & 0 & -a_n \\ 1 & t & 0 & & 0 & -a_{n-1} \\ 0 & 1 & t & & 0 & -a_{n-2} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & 1 & t & -a_2 \\ 0 & 0 & & & 1 & t-a_1 \end{vmatrix}$$

Here, the second determinant is equal to $(-1)^{n-1}a_n$ and by induction on n the first determinant is equal to $t.(t^{n-1} - a_1 t^{n-2} + \ldots + (-1)^{n-1}a_{n-1})$, proving the claim.

Next, we will determine which $n \times n$ matrices can be conjugated to a matrix in the canonical form A as above. We call a matrix $B \in M_n$ *cyclic* if there is a (column) vector $v \in \mathbb{C}^n$ such that $\mathbb{C}^n$ is spanned by the vectors $\{v, B.v, B^2.v, \ldots, B^{n-1}.v\}$. Let $g \in GL_n$ be the basechange transforming the standard basis to the ordered basis

$$(v, -B.v, B^2.v, -B^3.v, \ldots, (-1)^{n-1}B^{n-1}.v).$$

In this new basis, the linear map determined by B (or equivalently, $g.B.g^{-1}$) is equal to the matrix in canonical form

$$\begin{bmatrix} 0 & & & & b_n \\ -1 & 0 & & & b_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & -1 & 0 & b_2 \\ & & & -1 & b_1 \end{bmatrix}$$

where $B^n.v$ has coordinates $(b_n, \ldots, b_2, b_1)$ in the new basis.

Prove that a matrix $B \in M_n$ can be conjugated to one in standard form as above if *and only if* B is a cyclic matrix.

We claim that the set of all cyclic matrices in M_n is a *dense* open subset, that is, its closure is the whole of M_n. To see this take $v = (x_1, \ldots, x_n)^\tau \in \mathbb{C}^n$ and compute the determinant of the $n \times n$ matrix

$$\begin{bmatrix} \boxed{v} & \boxed{Bv} & \cdots & \boxed{B^{n-1}v} \end{bmatrix}$$

This gives a polynomial of total degree n in the x_i with all its coefficients polynomial functions c_j in the entries b_{kl} of B. Now, B is a cyclic matrix if and only if at least one of these coefficients is non-zero. That is, the set of non-cyclic matrices is exactly the intersection of the finitely many *hypersurfaces*

$$V_j = \{B = (b_{kl})_{k,l} \in M_n \mid c_j(b_{11}, b_{12}, \ldots, b_{nn}) = 0\}$$

in the vectorspace M_n. The claim follows because the complement of the hypersurface V_j is a dense open subset (being equal to the inverse image of the dense open set $\mathbb{C} - \{0\}$ in $\mathbb{C}$ under the continuous surjection $c_j : M_n \longrightarrow \mathbb{C}$).

Theorem 2.1. *The best continuous approximation to the orbit space is given by the surjection*

$$M_n \xrightarrow{\pi} \mathbb{C}^n$$

mapping a matrix $A \in M_n(\mathbb{C})$ to the n-tuple $(\sigma_1(A), \ldots, \sigma_n(A))$.

Let $f : M_n \longrightarrow \mathbb{C}$ be a continuous function which is constant along conjugacy classes. We will show that f factors through π, that is, f is really a continuous function in the $\sigma_i(A)$. Consider the diagram

$$\begin{tikzcd} M_n \arrow[r, "f"] \arrow[d, shift left, "\pi"] & \mathbb{C} \\ \mathbb{C}^n \arrow[u, shift left, "s"] \arrow[ur, dotted, "f'=f\circ s"'] & \end{tikzcd}$$

where s is the *section* of π (that is, $\pi \circ s = id_{\mathbb{C}^n}$) determined by sending a point $(a_1, \ldots, a_n)$ to the cyclic matrix in canonical form A as in equation (2.1). Clearly, s is continuous, hence so is $f' = f \circ s$. The approximation property follows if we prove that $f = f' \circ \pi$. By continuity, it suffices to check equality on the dense open set of cyclic matrices in M_n.

There it is a consequence of the following three facts we have proved before : (1) : any cyclic matrix lies in the same orbit as one in standard form, (2) : s is a section of π and (3) : f is constant along orbits.

2b. Some examples.

Miniature 1. Orbits in M_2.

A 2×2 matrix A can be conjugated to an upper triangular matrix with diagonal entries the eigenvalues λ_1, λ_2 of A. As the trace and determinant of both matrices are equal we have

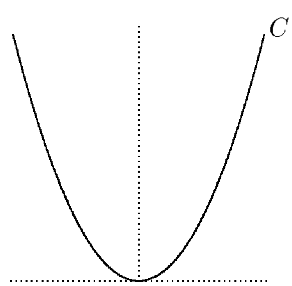
$$\sigma_1(A) = tr(A) \text{ and } \sigma_2(A) = det(A).$$

The best approximation to the orbitspace is therefore given by the surjective map

$$M_2 \xrightarrow{\pi} \mathbb{C}^2 \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto (a+d, ad-bc)$$

The matrix A has two equal eigenvalues if and only if the discriminant of the characteristic polynomial $t^2 - \sigma_1(A)t + \sigma_2(A)$ is zero, that is when $\sigma_1(A)^2 - 4\sigma_2(A) = 0$. This condition determines a closed curve C in $\mathbb{C}^2$ where

$$C = \{(x, y) \in \mathbb{C}^2 \mid x^2 - 4y = 0\}.$$



Observe that C is a smooth 1-dimensional submanifold of $\mathbb{C}^2$. We will describe the *fibers* (that is, the inverse images of points) of the surjective map π .

If $p = (x, y) \in \mathbb{C}^2 - C$, then $\pi^{-1}(p)$ consists of precisely one orbit (which is then necessarily closed in M_2) namely that of the diagonal matrix

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad \text{where} \quad \lambda_{1,2} = \frac{-x \pm \sqrt{x^2 - 4y}}{2}$$

If $p = (x, y) \in C$ then $\pi^{-1}(p)$ consists of two orbits,

$$\mathcal{O}_{\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}} \quad \text{and} \quad \mathcal{O}_{\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}}$$

where $\lambda = \frac{1}{2}x$. We have seen that the second orbit lies in the closure of the first. Observe that the second orbit reduces to one point in M_2 and hence is closed. Hence, also $\pi^{-1}(p)$ contains a unique closed orbit.

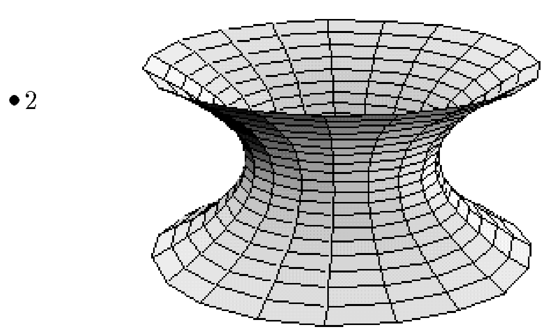
To describe the fibers of π as closed subsets of M_2 it is convenient to write any matrix A as a linear combination

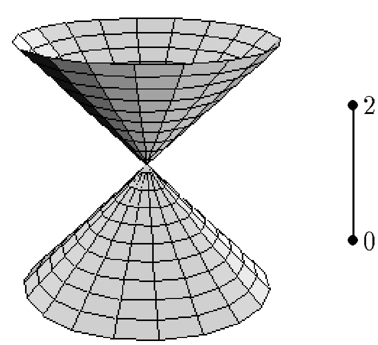
$$A = u(A) \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} + v(A) \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} + w(A) \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + z(A) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Expressed in the coordinate functions u, v, w and z the fibers $\pi^{-1}(p)$ of a point $p = (x, y) \in \mathbb{C}^2$ are the common zeroes of

$$\begin{cases} u & = x \\ v^2 + 4wz & = x^2 - 4y \end{cases}$$

The first equation determines a three dimensional affine subspace of M_2 in which the second equation determines a *quadric*.





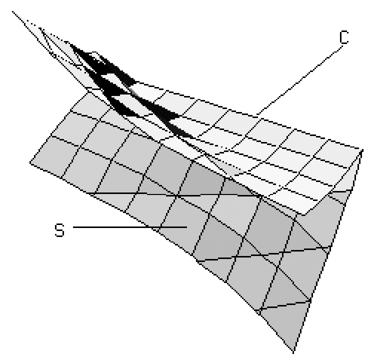
If $p \notin C$ this quadric is non-degenerate and thus $\pi^{-1}(p)$ is a smooth 2-dimensional submanifold of M_2. If $p \in C$, the quadric is a cone with top lying in the point $\frac{x}{2}\mathbb{1}_2$. Under the GL_2-action, the unique singular point of the cone must be clearly fixed giving us the closed orbit of dimension 0 corresponding to the diagonal matrix. The other orbit is the complement of the top and hence is a smooth 2-dimensional (non-closed) submanifold of M_2. The graphs represent the orbit-closures and the dimensions of the orbits.

Miniature 2. Orbits in M_3.

We will describe the fibers of the surjective map $M_3 \xrightarrow{\pi} \mathbb{C}^3$. If a 3×3 matrix has multiple eigenvalues then the *discriminant* $d = (\lambda_1 - \lambda_2)^2(\lambda_2 - \lambda_3)^2(\lambda_3 - \lambda_1)^2$ is zero. Clearly, d is a symmetric polynomial and hence can be expressed in terms of σ_1, σ_2 and σ_3. More precisely,

$$d = 4\sigma_1^3\sigma_3 + 4\sigma_2^3 + 27\sigma_3^2 - \sigma_1^2\sigma_2^2 - 18\sigma_1\sigma_2\sigma_3$$

The set of points in $\mathbb{C}^3$ where d vanishes is a surface S with *singularities*.



These singularities are the common zeroes of the $\frac{\partial d}{\partial \sigma_i}$ for $1 \leq i \leq 3$. One computes that these singularities form a *twisted cubic* curve C in $\mathbb{C}^3$, that is,

$$C = \{(3c, 3c^2, c^3) \mid c \in \mathbb{C}\}.$$

The description of the fibers $\pi^{-1}(p)$ for $p = (x, y, z) \in \mathbb{C}^3$ is as follows. When $p \notin S$, then $\pi^{-1}(p)$ consists of a unique orbit (which is therefore closed in M_3), the conjugacy class of a matrix with pairwise distinct eigenvalues. If $p \in S - C$, then $\pi^{-1}(p)$ consists of the orbits of

$$A_1 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix} \text{ and } A_2 = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}$$

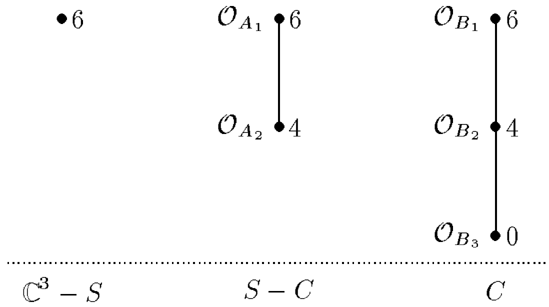
Finally, if $p \in C$, then the matrices in the fiber $\pi^{-1}(p)$ have a single eigenvalue $\lambda = \frac{1}{3}x$ and the fiber consists of the orbits of the matrices

$$B_1 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \quad B_2 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \quad B_3 = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

We observe that the *strata* with distinct fiber behavior (that is, $\mathbb{C}^3 - S$, $S - C$ and C) are all submanifolds of $\mathbb{C}^3$.

The dimension of an orbit $\mathcal{O}_A$ in M_n is computed as follows. Let C_A be the subspace of all matrices in M_n commuting with A. Then, the *stabilizer* subgroup of A is a dense open subset of C_A whence the dimension of $\mathcal{O}_A$ is equal to $n^2 - \dim C_A$.

Performing these calculations for the matrices given above, we obtain the following graphs representing orbit-closures and the dimensions of orbits



Week 3

The orbit closures.

This week we will prove the Gerstenhaber-Hesselink theorem which gives an answer to :

Problem 3.

Describe the orbit-closures for general n.

Consider the *quotient map* $M_n \xrightarrow{\pi} \mathbb{C}^n$ and a point $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$. All matrices A in the *fiber* $\pi^{-1}(x)$ have the same eigenvalues $\{\lambda_1, \ldots, \lambda_e\}$ and multiplicities $\{d_1, \ldots, d_e\}$ because

$$\prod_{i=1}^{e}(t-\lambda_i)^{d_i} = \chi_A(t) = t^n - x_1 t^{n+1} + x_2 t^{n-2} - \ldots + (-1)^n x_n.$$

We have seen in example **??** that the orbit corresponding to the diagonal matrix with these eigenvalues and multiplicities is contained in the closure of $\mathcal{O}_A$ for any $A \in \pi^{-1}(x)$. Moreover, for any $\lambda \in \mathbb{C}$, the linear automorphism

$$M_n \xrightarrow{\phi_\lambda} M_n \qquad \text{defined by} \qquad A \mapsto A - \lambda \mathbb{1}_n$$

commutes with the action of GL_n. Hence, in studying the fibers $\pi^{-1}(x)$ we may assume that one of the eigenvalues λ_i is zero. An important subproblem is therefore to study the orbit-closures in the *nullcone*, $N = \pi^{-1}(\underline{0})$, that is the set of all *nilpotent* $n \times n$ matrices. We recall that a matrix A is said to be nilpotent if $A^k = 0$ for some power k. Clearly, $A \in M_n$ is nilpotent if and only if $A^n = 0$.

A nilpotent matrix A has 0 as its unique eigenvalue (occurring with multiplicity n). Therefore, by theorem **??** the orbits of nilpotent matrices are in one-to-one correspondence with partitions of n. We will first introduce a *dominance* ordering on all partitions of n and consequently show that this ordering determines the orbit closures of nilpotent matrices.

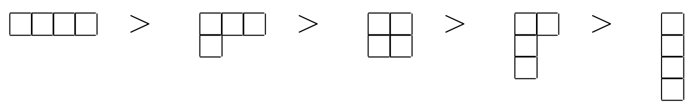
3a. The Gerstenhaber-Hesselink theorem.

It is sometimes convenient to relax our definition of partitions to include zeroes at its tail. That is, a partition p of n is an integral n-tuple $(a_1, a_2, \ldots, a_n)$ with $a_1 \geq a_2 \geq \ldots \geq a_n \geq 0$ with $\sum_{i=1}^n a_i = n$. As before, we represent a partition by a Young diagram by omitting rows corresponding to zeroes.

If $q = (b_1, \ldots, b_n)$ is another partition of n we say that p *dominates* q and write

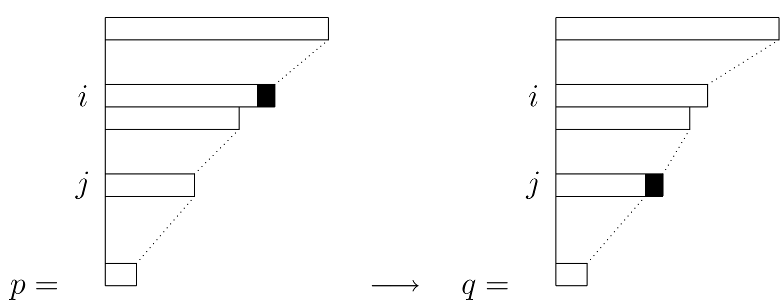
$$p > q \quad \text{if and only if} \quad \sum_{i=1}^r a_i \geq \sum_{i=1}^r b_i \quad \text{for all } 1 \leq r \leq n.$$

For example, the partitions of 4 are ordered as indicated below

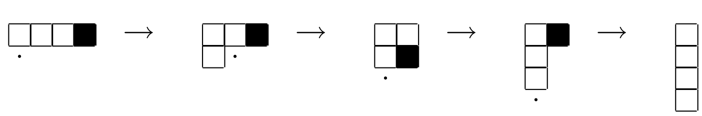


Note however that the dominance relation is *not* a total ordering whenever $n \geq 6$. For example, the following two partition of 6

are not comparable. The dominance order is induced by the *Young move* of throwing a row-ending box down the diagram. Indeed, let p and q be partitions of n such that $p > q$ and assume there is no partition r such that $p > r$ and $r > q$. Let i be the minimal number such that $a_i > b_i$. Then by the assumption $a_i = b_i + 1$. Let $j > i$ be minimal such that $a_j \neq b_j$, then we have $b_j = a_j + 1$ because p dominates q. But then, the remaining rows of p and q must be equal. That is, a Young move can be depicted as



For example, the Young moves between the partitions of 4 given above are as indicated



A *Young p-tableau* is the Young diagram of p with the boxes labeled by integers from $\{1, 2, \ldots, s\}$ for some s such that each label appears at least ones. A Young p-tableau is said to be *of type* q for some partition $q = (b_1, \ldots, b_n)$ of n if the following conditions are met :

- the labels are non-decreasing along rows,
- the labels are strictly increasing along columns, and
- the label i appears exactly b_i times.

For example, if $p = (3, 2, 1, 1)$ and $q = (2, 2, 2, 1)$ then the p-tableau below

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

is of type q (observe that $p > q$ and even $p \to q$). In general, let $p = (a_1, \ldots, a_n)$ and $q = (b_1, \ldots, b_n)$ be partitions of n and assume that $p \to q$. Then, there is a Young p-tableau of type q. For, fill the Young diagram of q by putting 1's in the first row, 2's in the second and so on. Then, upgrade the fallen box together with its label to get a Young p-tableau of type q. In the example above

$$\begin{array}{|c|c|c} \hline 1 & 1 & \bullet \\ \hline 2 & 2 & \\ \hline 3 & ③ & \\ \hline 4 & \multicolumn{1}{c}{} & \\ \cline{1-1} \end{array} \quad \Longrightarrow \quad \begin{array}{|c|c|c|} \hline 1 & 1 & 3 \\ \hline 2 & 2 & \multicolumn{1}{c}{} \\ \cline{1-2} 3 & \multicolumn{2}{c}{} \\ \cline{1-1} 4 & \multicolumn{2}{c}{} \\ \cline{1-1} \end{array}$$

Conversely, assume there is a Young p-tableau of type q. The number of boxes labeled with a number $\leq i$ is equal to $b_1 + \ldots + b_i$. Further, any box with label $\leq i$ must lie in the first i rows (because the labels strictly increase along a column). There are $a_1 + \ldots + a_i$ boxes available in the first i rows whence

$$\sum_{j=1}^{i} b_i \leq \sum_{j=1}^{i} a_i \quad \text{for all} \quad 1 \leq i \leq n$$

and therefore $p > q$. After these preliminaries on partitions, let us return to nilpotent matrices.

Let A be a nilpotent matrix of type $p = (a_1, \ldots, a_n)$, that is, conjugated to a matrix with Jordan blocks (all with eigenvalue zero) of sizes a_i. It follows from equation **??** that the subspace V_l of column vectors $v \in \mathbb{C}^n$ such that $A^l.v = 0$ has dimension

$$\sum_{h=1}^{l} \#\{j \mid a_j \geq h\} = \sum_{h=1}^{l} a_h^*$$

where $p^* = (a_1^*, \ldots, a_n^*)$ is the dual partition of p. Choose a basis $\{v_1, \ldots, v_n\}$ of $\mathbb{C}^n$ such that for all l the first $a_1^* + \ldots + a_l^*$ base vectors span the subspace V_l. For example, if A is in Jordan normal form of type $p = (3, 2, 1, 1)$

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

then $p^* = (4, 2, 1)$ and we can choose the standard base vectors ordered as follows

$$\{\underbrace{e_1, e_4, e_6, e_7}_{4}, \underbrace{e_2, e_5}_{2}, \underbrace{e_3}_{1}\}.$$

Take a partition $q = (b_1, \ldots, b_n)$ with $p \to q$ (in particular, $p > q$), then for the dual partitions we have $q^* \to p^*$ (and thus $q^* > p^*$). But then there is a Young q^*-tableau of type p^*. In the example with $q = (2, 2, 2, 1)$ we have $q^* = (4, 3)$ and $p^* = (4, 2, 1)$ and we can take the Young q^*-tableau of type p^*

1	1	1	1
2	2	3	

Now label the boxes of this tableau by the base vectors $\{v_1, \ldots, v_n\}$ such that the boxes labeled i in the Young q^*-tableau of type p^* are filled with the base vectors from $V_i - V_{i-1}$. Call this tableau T. In the example, we can take

$T =$

e_1	e_4	e_6	e_7
e_2	e_5	e_3	

Define a linear operator F on $\mathbb{C}^n$ by the rule that $F(v_i) = v_j$ if v_j is the label of the box in T just above the box labeled v_i. In case v_i is a label of a box in the first row of T we take $F(v_i) = 0$. Obviously, F is a nilpotent $n \times n$ matrix and by construction we have that

$$rk\ F^l = n - (b_1^* + \ldots + b_l^*)$$

That is, F is nilpotent of type $q = (b_1, \ldots, b_n)$. Moreover, F satisfies $F(V_i) \subset V_{i-1}$ for all i by the way we have labeled the tableau T and defined F.

In the example above, we have $F(e_2) = e_1$, $F(e_5) = e_4$, $F(e_3) = e_6$ and all other $F(e_i) = 0$. That is, F is the matrix

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

which is seen to be of type $(2, 2, 2, 1)$ after performing a few Jordan moves.

Returning to the general case, consider for all $\epsilon \in \mathbb{C}$ the $n \times n$ matrix :

$$F_\epsilon = (1 - \epsilon)F + \epsilon A.$$

We claim that for all but finitely many values of ϵ we have $F_\epsilon \in \mathcal{O}_A$. Indeed, we have seen that $F(V_i) \subset V_{i-1}$ where V_i is defined as the subspace such that $A^i(V_i) = 0$. Hence, $F(V_1) = 0$ and therefore

$$F_\epsilon(V_1) = (1 - \epsilon)F + \epsilon A(V_1) = 0.$$

Assume by induction that $F_\epsilon^i(V_i) = 0$ holds for all $i < l$, then we have that

$$\begin{aligned} F_\epsilon^l(V_l) &= F_\epsilon^{l-1}((1-\epsilon)F + \epsilon A)(V_l) \\ &\subset F_\epsilon^{l-1}(V_{l-1}) = 0 \end{aligned}$$

because $A(V_l) \subset V_{l-1}$ and $F(V_l) \subset V_{l-1}$. But then we have for all l that

$$rk\ F_\epsilon^l \leq dim\ V_l = n - (a_1^* + \ldots + a_l^*) = rk\ A^l \overset{def}{=} r_l.$$

Then for at least one $r_l \times r_l$ submatrix of F_ϵ^l its determinant considered it as a polynomial of degree r_l in ϵ is not identically zero (as it is nonzero for $\epsilon = 1$). But then this determinant is non-zero for all but finitely many ϵ. Hence, $rk\ F_\epsilon^l = rk\ A^l$ for all l for all but finitely many ϵ. As these numbers determine the dual partition p^* of the type of A, F_ϵ is a nilpotent $n \times n$ matrix of type p for all but finitely many values of ϵ, proving the claim. But then, $F_0 = F$ which we have proved to be a nilpotent matrix of type q belongs to the closure of the orbit $\mathcal{O}_A$. That is, we have proved the difficult part of the *Gerstenhaber-Hesselink theorem*.

Theorem 3.1. *Let A be a nilpotent $n \times n$ matrix of type $p = (a_1, \ldots, a_n)$ and B nilpotent of type $q = (b_1, \ldots, b_n)$. Then, B belongs to the closure of the orbit $\mathcal{O}_A$, that is,*

$$B \in \overline{\mathcal{O}_A} \quad \textit{if and only if} \quad p > q$$

in the domination order on partitions of n.

To prove the theorem we only have to observe that if B is contained in the closure of A, then B^l is contained in the closure of A^l and hence $rk\ A^l \geq rk\ B^l$ (because $rk\ A^l < k$ is equivalent to vanishing of all determinants of $k \times k$ minors which is a closed condition). But then,

$$n - \sum_{i=1}^{l} a_i^* \geq n - \sum_{i=1}^{l} b_i^*$$

for all l, that is, $q^* > p^*$ and hence $p > q$. The other implication was proved above if we remember that the domination order was induced by the Young moves and clearly we have that if $B \in \overline{\mathcal{O}_C}$ and $C \in \overline{\mathcal{O}_A}$ then also $B \in \overline{\mathcal{O}_A}$.

3b. Some examples.

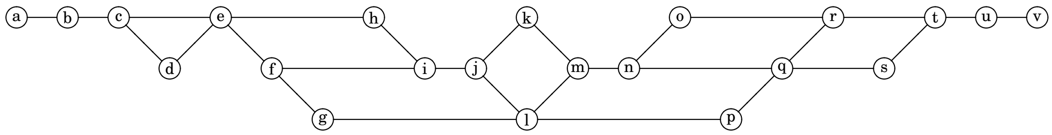
Miniature 3. Nilpotent matrices for small n.

We will apply theorem 1 to describe the orbit-closures of nilpotent matrices of 8×8 matrices. The following table lists all partitions (and their dual in the other column)

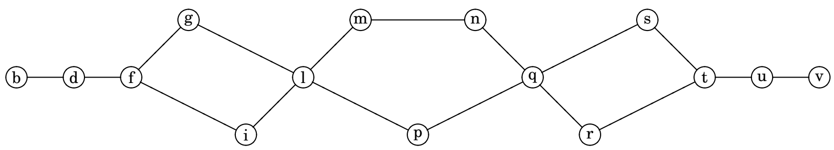
The partitions of 8.

a	(8)	v	(1,1,1,1,1,1,1,1)
b	(7,1)	u	(2,1,1,1,1,1,1)
c	(6,2)	t	(2,2,1,1,1,1)
d	(6,1,1)	s	(3,1,1,1,1,1)
e	(5,3)	r	(2,2,2,1,1)
f	(5,2,1)	q	(3,2,1,1,1)
g	(5,1,1,1)	p	(4,1,1,1,1)
h	(4,4)	o	(2,2,2,2)
i	(4,3,1)	n	(3,2,2,1)
j	(4,2,2)	m	(3,3,1,1)
k	(3,3,2)	k	(3,3,2)
l	(4,2,1,1)	l	(4,2,1,1)

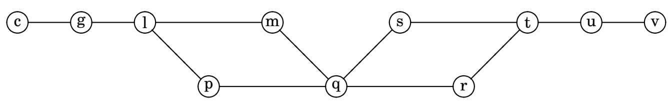
The domination order between these partitions can be depicted as follows where all the Young moves are from left to right



Of course, from this graph we can read off the dominance order graphs for partitions of $n \leq 8$. The trick is to identify a partition of n with that of 8 by throwing in a tail of ones and to look at the relative position of both partitions in the above picture. Using these conventions we get the following graph for partitions of 7



and for partitions of 6 the dominance order is depicted as follows



We have already mentioned that the dominance order on partitions of $n \leq 5$ is a total ordering.

We will prove later that knowledge of the orbit closures of nilpotent matrices for $m \times m$ matrices for all $m \leq n$ is enough to understand the orbit closures in all the fibers $\pi^{-1}(x)$ for the quotient map $M_n \xrightarrow{\pi} \mathbb{C}^n$.

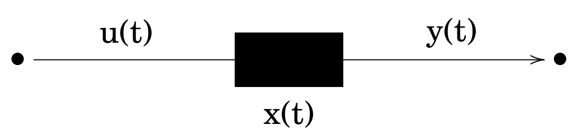
Week 4

Dynamical systems.

In this section we will consider *linear time-invariant dynamical systems*. Whereas this is a gross simplification of actual processes, often one can reduce to such a situation (as a first approximation), for example near an equilibrium state of the system. A linear time invariant dynamical system Σ is governed by the following system of differential equations diffeq

$$\begin{cases} \dfrac{dx}{dt} &= Bx + Au \\ y &= Cx. \end{cases} \tag{4.1}$$

Here, $u(t) \in \mathbb{C}^m$ is the *input* or *control* of the system at tome t, $x(t) \in \mathbb{C}^n$ the *state* of the system and $y(t) \in \mathbb{C}^p$ the *output* of the system Σ. *Time invariance* of Σ means that the matrices $A \in M_{n\times m}(\mathbb{C})$, $B \in M_n(\mathbb{C})$ and $C \in M_{p\times n}(\mathbb{C})$ are constant. The system Σ can be represented as a *black box*



which is in a certain state $x(t)$ that we can try to change by using the input controls $u(t)$. By reading the output signals $y(t)$ we can try to determine the state of the system. We briefly recall how one solves a linear dynamical system.

4a. Solving linear systems.

Recall that the *matrix exponential* e^B of any $n \times n$ matrix B is defined by the infinite series

$$e^B = \mathbb{1}_n + B + \frac{B^2}{2!} + \ldots + \frac{B^m}{m!} + \ldots$$

Observe that this series converges to a matrix in $M_n(\mathbb{C})$ as the norm $\mid B^n \mid$ is bounded by $\mid B \mid^m$ for any m. The importance of this construction is clear from the fact that e^{Bt} is the *fundamental matrix* for the homogeneous differential equation $\frac{dx}{dt} = Bx$. That is, the columns of e^{Bt} are a basis for the n-dimensional space of solutions of the equation $\frac{dx}{dt} = Bx$.

Motivated by this, let us look for a solution to equation (4.1) as the form $x(t) = e^{Bt}g(t)$ for some function $g(t)$. Substitution gives the condition

$$\frac{dg}{dt} = e^{-Bt}Au \quad \text{whence} \quad g(\tau) = g(\tau_0) + \int_{\tau_0}^{\tau} e^{-Bt}Au(t)dt.$$

Observe that $x(\tau_0) = e^{B\tau_0}g(\tau_0)$ and we obtain the solution of the linear dynamical system $\Sigma = (A, B, C)$:

$$\begin{cases} x(\tau) & = e^{(\tau-\tau_0)B}x(\tau_0) + \int_{\tau_0}^{\tau} e^{(\tau-t)B}Au(t)dt \\ y(\tau) & = Ce^{B(\tau-\tau_0)}x(\tau_0) + \int_{\tau_0}^{\tau} Ce^{(\tau-t)B}Au(t)dt. \end{cases}$$

Differentiating we see that this is indeed a solution and it is the unique one having a prescribed starting state $x(\tau_0)$. Indeed, given another solution $x_1(\tau)$ we have that $x_1(\tau) - x(\tau)$ is a solution to the homogeneous system $\frac{dx}{dt} = Bt$, but then

$$x_1(\tau) = x(\tau) + e^{\tau B}e^{-\tau_0 B}(x_1(\tau_0) - x(\tau_0)).$$

4b. Observable and controllable systems.

We will recall some important system-theoretic notions describing the level of *control* or *observation* a given system allows.

We call the system Σ *completely controllable* if we can steer any starting state $x(\tau_0)$ to the zero state by some control function $u(t)$ in a finite time span $[\tau_0, \tau]$. That is, the equation

$$0 = x(\tau_0) + \int_{\tau_0}^{\tau} e^{(\tau_0-t)B}Au(t)dt$$

has a solution in τ and $u(t)$. As the system is time-invariant we may always
eqcontrol assume that $\tau_0 = 0$ and have to satisfy the equation

$$0 = x_0 + \int_0^{\tau} e^{tB}Au(t)dt \quad \text{for every} \quad x_0 \in \mathbb{C}^n \tag{4.2}$$

Consider the *control matrix* $c(\Sigma)$ which is the $n \times mn$ matrix

$$c(\Sigma) = \begin{array}{|c|c|c|c|c|} \hline A & BA & B^2A & \cdots & B^{n-1}A \\ \hline \end{array}$$

Assume that $rk\ c(\Sigma) < n$ then there is a non-zero state $s \in \mathbb{C}^n$ such that $s^{tr} c(\Sigma) = 0$, where s^{tr} denotes the transpose (row column) of s. Because B satisfies the characteristic polynomial $\chi_B(t)$, B^n and all higher powers B^m are linear combinations of $\{\mathbb{1}_n, B, B^2, \dots, B^{n-1}\}$. Hence, $s^\tau B^m A = 0$ for all m. Writing out the power series expansion of e^{tB} in equation (4.2) this leads to the contradiction that $0 = s^\tau x_0$ for all $x_0 \in \mathbb{C}^n$. Hence, if $rk\ c(\Sigma) < n$, then Σ is not completely controllable.

Conversely, let $rk\ c(\Sigma) = n$ and assume that Σ is not completely controllable. That is, the space of all states

$$s(\tau, u) = \int_0^\tau e^{-tB} A u(t) dt$$

is a proper subspace of $\mathbb{C}^n$. But then, there is a non-zero state $s \in \mathbb{C}^n$ such that $s^{tr} s(\tau, u) = 0$ for all τ and all functions $u(t)$. Differentiating this with respect to τ we obtain eq11

$$s^{tr} e^{-\tau B} A u(\tau) = 0 \quad \text{whence} \quad s^{tr} e^{-\tau B} A = 0 \tag{4.3}$$

for any τ as $u(\tau)$ can take on any vector. For $\tau = 0$ this gives $s^{tr} A = 0$. If we differentiate (4.3) with respect to τ we get $s^{tr} B e^{-\tau B} A = 0$ for all τ and for $\tau = 0$ this gives $s^{tr} BA = 0$. Iterating this process we show that $s^{tr} B^m A = 0$ for any m, whence

$$s^{tr} \begin{bmatrix} A & BA & B^2 A & \dots & B^{n-1} A \end{bmatrix} = 0$$

contradicting the assumption that $rk\ c(\Sigma) = n$. That is, we have proved :

Proposition 4.1. *A linear time-invariant dynamical system Σ determined by the matrices (A, B, C) is completely controllable if and only if $rk\ c(\Sigma)$ is maximal.*

$$\begin{bmatrix} C \\ CB \\ CB^2 \\ \dots \\ CB^{n-1} \end{bmatrix}$$

Next, we turn to the problem to what degree we can obtain information about the system by reading off its output displays. We say that a state $x(\tau)$ at time τ is *unobservable* if $Ce^{(\tau-t)B}x(\tau) = 0$ for all t. Intuitively this means that the state $x(\tau)$ cannot be detected uniquely from the output of the system Σ. Again, if we differentiate this condition a number of times and evaluate at $t = \tau$ we obtain the conditions

$$Cx(\tau) = CBx(\tau) = \ldots = CB^{n-1}x(\tau) = 0.$$

We say that Σ is *completely observable* if the zero state is the only unobservable state at any time τ. Consider the *observation matrix* $o(\Sigma)$ of the system Σ which is the $pn \times n$ matrix

$$o(\Sigma) = \begin{bmatrix} C^{tr} & (CB)^{tr} & \ldots & (CB^{n-1})^{tr} \end{bmatrix}^{tr}$$

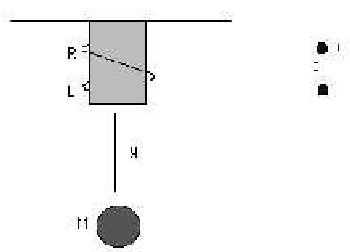
An analogous argument as in the proof of proposition 4.1. gives us :

Proposition 4.2. *A linear time-invariant dynamical system Σ determined by the matrices (A, B, C) is completely observable if and only if $rk\ o(\Sigma)$ is maximal.*

4c. Some examples.

Miniature 4. A physical dynamical system.

We consider a simple magnetic-ball suspension system. The objective of the system is to control the position of the steel ball by adjusting the current in the electromagnet through the input voltage $e(t)$.



The differential equations determining this system are :

$$\begin{cases} M\dfrac{d^2y(t)}{dt^2} & = Mg - \dfrac{i^2(t)}{y(t)} \\ e(t) & = Ri(t) + L\dfrac{di(t)}{dt} \end{cases}$$

Here, $y(t)$ is the distance of teh ball from the magnet, M the mass of the ball and g the gravitational constant. The electromagnet has winding inductance L, winding resistance R and $i(t)$ is the winding current. If we want to approximate this system by a linear system we have to solve two problems : (1) to replace the higher order differential term in the left hand side by first order terms, and (2) to linearize the non-linear terms in teh right hand side.

The first problem is solved by teh standard trick that a single n-th order differential equation

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_n y = u(t)$$

is equivalent to a linear system

$$\frac{dx}{dt} = Fx + u$$

where we take $x = (x_1, \ldots, x_n)$ with $x_1 = y$, $x_i = \frac{d^i y}{dt^i}$ for $i \geq 2$ and where F is then$n \times n$ matrix

$$F = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_n \\ 1 & 0 & \ldots & 0 & -a_{n-1} \\ 0 & 1 & & 0 & -a_{n-2} \\ & & \ddots & & \vdots \\ 0 & 0 & & 1 & -a_1 \end{bmatrix}$$

Hence, in the above suspension system we take as state vaiables $x_1(t) = y(t)$, $x_2(t) = \frac{dy}{dt}$ and $x_3(t) = i(t)$. Then, the defining equations of teh system become

$$\begin{cases} \frac{dx_1}{dt} & = x_2 \\ \frac{dx_2}{dt} & = g - \frac{1}{M}\frac{x_3^2}{x_1} \\ \frac{dx_3}{dt} & = -\frac{R}{L}x_3 + \frac{1}{L}e \end{cases}$$

To remove the non-linear terms in teh right hand side we consider an equilibrium state with $y_0(t) = a$ is constant. Then, $\frac{d^i y_0(t)}{dt^i} = 0$ and usbstitution in the differential equations gives that the equilibrium state determined by a is

$$x_0(t) = (a, 0, \sqrt{Mga}).$$

Expanding the nonlinear terms into Taylor series about $x_0(t)$ and neglecting higher order terms we can approximate the system by the linear time-invariant system $\Sigma : \frac{dx'}{dt} = Bx' + A$ where $x'_i(t) = x_i(t) - x_{0i}(t)$ and

$$B = \begin{bmatrix} 0 & 1 & 0 \\ \frac{g}{a} & 0 & -2\sqrt{\frac{g}{Ma}} \\ 0 & 0 & -\frac{R}{L} \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

The control matrix $c(\Sigma)$ is teh 3×3 matrix

$$[A, BA, B^2A] = \begin{bmatrix} 0 & 0 & -\frac{2}{L} \\ 0 & -\frac{2}{L}\sqrt{\frac{g}{Ma}} & \frac{2R}{L^2}\sqrt{\frac{g}{Ma}} \\ \frac{1}{L} & -\frac{R}{L^2} & \frac{R^2}{L^3} \end{bmatrix}$$

which is of rank 3 hence Σ is completely controllable. Observability of the system depends on which variable we define as the output. First, assume that the output signal is the distance x_1 from teh ball to the magnet, that is, $C = \begin{bmatrix}1,0,0\end{bmatrix}$. Then, teh observation matrix $o(\Sigma)$ is equal to

$$\begin{bmatrix} C \\ CB \\ C^2B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{g}{a} & 0 & -2\sqrt{\frac{g}{Ma}} \end{bmatrix}$$

and the system is completely controllable. Similarly, if the output is teh speed x_2 of the ball then one verifies complete controllability. However, if our output signal is the current i, that is, $C = \begin{bmatrix}0,0,1\end{bmatrix}$ then the controllability matrix is

$$\begin{bmatrix} C \\ CB \\ C^2B \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -\frac{R}{L} \\ 0 & 0 & 2\frac{R}{L}\sqrt{\frac{g}{Ma}} \end{bmatrix}$$

Thus, the system is not completely controllable. That is, observing only the current $i(t)$ we are not always able to reconstruct the state of teh system. If however, we would observe the distance $y(t)$ (or the speed $\frac{dy}{dt}$) we are able to reconstruct the state of the system.

Week 5

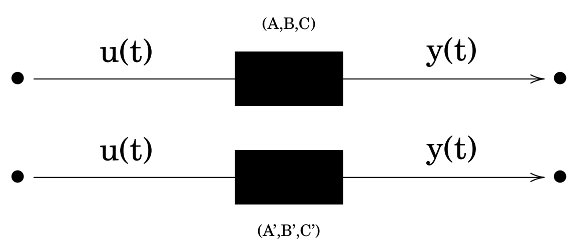
The orbit space.

Usually, a system is a black box, that is, its inner workings are unknown to us and we can only detect its input/output behavior. Let us restrict to linear time-invariant dynamical systems which are both completely controllable and completely observable and call such systems *Schurian*. Of fundamental importance in system theory is the solution to :

Problem 4.

Classify Schurian dynamical systems with the same input/output behavior.

First, we will reduce this problem to the study of GL_n-orbits in an open subset of a certain vectorspace. Assume we have two systems Σ and Σ', determined by matrix triples from $Sys = M_{n\times m}(\mathbb{C}) \times M_n(\mathbb{C}) \times M_{p\times n}(\mathbb{C})$



producing the same output $y(t)$ when given the same input $u(t)$, for all possible input functions $u(t)$. We recall that the output function y for a system $\Sigma = (A, B, C)$ is determined by

$$y(\tau) = Ce^{B(\tau-\tau_0)}x(\tau_0) + \int_{\tau_0}^{\tau} Ce^{(\tau-t)B}Au(t)dt.$$

Differentiating this a number of times and evaluating at $\tau = \tau_0$ as in the proof of proposition **??** equality of input/output for Σ and Σ' gives the conditions

$$CB^iA = C'B'^iA' \quad \text{for all} \quad i.$$

5a. The Kalman code.

As a consequence the systems Σ and Σ' have the same *Hankel matrix* which by definition is the product of the observation matrix with the control matrix of the system :

$$H(\Sigma) = \begin{bmatrix} C \\ CB \\ \vdots \\ CB^{n-1} \end{bmatrix} \begin{bmatrix} A & BA & \ldots & B^{n-1}A \end{bmatrix} = \begin{bmatrix} \boxed{CA} & \boxed{CBA} & \cdots & \\ \boxed{CBA} & \boxed{CB^2A} & \cdots & \\ \vdots & \vdots & \ddots & \\ & & & \boxed{CB^{2n-2}A} \end{bmatrix}$$

Alternatively, we can express this condition in terms of linear maps. Consider the two compositions

$$\begin{cases} \mathbb{C}^{mn} \xrightarrow{c(\Sigma)} \mathbb{C}^n \xrightarrow{o(\Sigma)} \mathbb{C}^{pn} \\ \mathbb{C}^{mn} \xrightarrow{c(\Sigma')} \mathbb{C}^n \xrightarrow{o(\Sigma')} \mathbb{C}^{pn} \end{cases}$$

Here, the control maps are onto by complete controllability and the observation maps are into by complete observability. Equality of input/output implies equality of the Hankel matrices and so the composed linear maps $\mathbb{C}^{mn} \longrightarrow \mathbb{C}^{pn}$ are equal.

But then, we have for any $v \in \mathbb{C}^{mn}$ that $c(\Sigma)(v) = 0 \Leftrightarrow c(\Sigma')(v)$ and we can decompose $\mathbb{C}^{pn} = V \oplus W$ such that the restriction of $c(\Sigma)$ and $c(\Sigma')$ to V are the zero map and the restrictions to W give isomorphisms with $\mathbb{C}^n$. Hence, there is an invertible matrix $g \in GL_n$ such that $c(\Sigma') = gc(\Sigma)$ and from the commutative diagram

$$\begin{tikzcd} \mathbb{C}^{mn} \arrow[r, two heads, "c(\Sigma)"] \arrow[d, equal] & \mathbb{C}^n \arrow[r, hook, "o(\Sigma)"] \arrow[d, "g"] & \mathbb{C}^{pn} \arrow[d, equal] \\ \mathbb{C}^{mn} \arrow[r, two heads, "c(\Sigma')"] & \mathbb{C}^n \arrow[r, hook, "o(\Sigma')"] & \mathbb{C}^{pn} \end{tikzcd}$$

we obtain that also $o(\Sigma') = o(\Sigma)g^{-1}$.

Consider the system $\Sigma_1 = (A_1, B_1, C_1)$ *equivalent* with Σ under the basechange matrix g. That is, $\Sigma_1 = g.\Sigma = (gA, gBg^{-1}, Cg^{-1})$. Then,

$$\left[A_1, B_1A_1, \ldots, B_1^{n-1}A_1\right] = gc(\Sigma) = c(\Sigma') = \left[A', B'A', \ldots, B'^{n-1}A'\right]$$

and so $A_1 = A'$. Further, as $B_1^{i+1}A_1 = B'^{i+1}A'$ we have by induction on i that the restriction of B_1 on the subspace of $B'^iIm(A')$ is equal to the restriction of B' on this space. Moreover, as $\sum_{i=0}^{n-1} B'^i Im(A') = \mathbb{C}^n$ it follows that $B_1 = B'$. Because $o(\Sigma') = o(\Sigma)g^{-1}$ we also have $C_1 = C'$. This finishes the proof of :

Proposition 5.1. *Let $\Sigma = (A, B, C)$ and $\Sigma' = (A', B', C')$ be two Schurian dynamical systems. The following are equivalent*

1. *The input/output behavior of Σ and Σ' are equal.*
2. *The systems Σ and Σ' are equivalent, that is, there exists an invertible matrix $g \in GL_n$ such that*

$$A' = gA, \quad B' = gBg^{-1} \quad \textit{and} \quad C' = Cg^{-1}.$$

This proposition reduces the system theoretic problem to a *fabric* setting. We consider the linear action of GL_n on the vectorspace of matrix triples defining linear time-invariant dynamical systems

$$Sys = M_{n\times m}(\mathbb{C}) \times M_n(\mathbb{C}) \times M_{p\times n}(\mathbb{C})$$

defined for all $g \in GL_n$ by

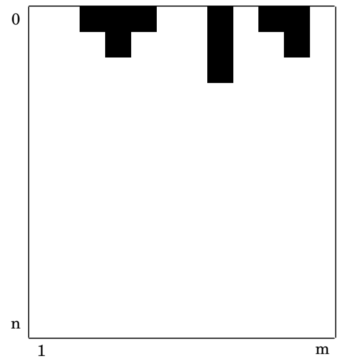
$$g.(A, B, C) = (gA, gBg^{-1}, Cg^{-1}).$$

For later reference we depict this action by the *pattern*

$$\begin{tikzcd} \circled{m} \arrow[r] & \circled{n} \arrow[loop above] \arrow[r] & \circled{p} \end{tikzcd}$$

By definition, a dynamical system $\Sigma = (A, B, C)$ is Schurian if (and only if) the determinant of at least one $n \times n$ minor of $c(\Sigma)$ and $o(\Sigma)$ is non-zero. That is, the subset Sys^s of Schurian dynamical systems is open in Sys and is stable under the GL_n-action. Our next job is to classify the orbits under this action.

We introduce a combinatorial gadget : the *Kalman code*. It is an array consisting of $(n + 1) \times m$ boxes each having a position label (i, j) where $0 \leq i \leq n$ and $1 \leq j \leq m$. These boxes are ordered *lexicographically* that is $(i', j') < (i, j)$ if and only if either $i' < i$ or $i' = i$ and $j' < j$. Exactly n of these boxes are painted black subject to the rule that if box (i, j) is black, then so is box (i', j) for all $i' < i$. That is, a Kalman code looks like



We assign to a completely controllable system $\Sigma = (A, B, C)$ its Kalman code $K(\Sigma)$ as follows : let $A = \begin{bmatrix} A_1 & A_2 & \dots & A_m \end{bmatrix}$, that is A_i is the i-th column of A. Paint the box (i, j) black if and only if the column vector $B^i A_j$ is linearly independent of the column vectors $B^k A_l$ for all $(k, l) < (i, j)$. The painted array $K(\Sigma)$ is indeed a Kalman code. Assume that box (i, j) is black but box (i', j) white for $i' < i$, then

$$B^{i'} A_j = \sum_{(k,l)<(i',j)} \alpha_{kl} B^k A_l \quad \text{but then,} \quad B^i A_j = \sum_{(k,l)<(i',j)} \alpha_{kl} B^{k+i-i'} A_l$$

and all $(k + i - i', l) < (i, l)$, a contradiction. Moreover, $K(\Sigma)$ has exactly n black boxes as there are n linearly independent columns of the control matrix $c(\Sigma)$ when Σ is completely controllable.

The Kalman code is a discrete invariant of the orbit $\mathcal{O}_\Sigma$ under the action of GL_n. This follows from the fact that $B^i A_j$ is linearly independent of the $B^k A_l$ for all $(k, l) < (i, j)$ if and only if $g B^i A_j$ is linearly independent of the $g B^k A_l$ for any $g \in GL_n$ and the observation that $g B^k A_l = (g B g^{-1})^k (g A)_l$. Next, we will clarify the geometric significance of the Kalman code.

As the Kalman code depends only on the input part (A, B) of the system $\Sigma = (A, B, C)$ we consider the linear action of GL_n on the vectorspace of matrixpairs $V = M_{n \times m}(\mathbb{C}) \times M_n(\mathbb{C})$ defined by $g.(A, B) = (gA, gBg^{-1})$.

With V_c we will denote the open subset of all completely controllable pairs (A, B) that is, those for which the rank of the $n \times nm$ matrix $\begin{bmatrix} A & BA & B^2A & \dots & B^{n-1}A \end{bmatrix}$ is maximal. We consider the map

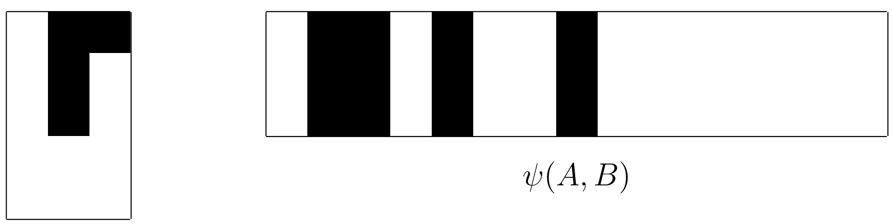
$$\begin{array}{ccc} V = M_{n\times m}(\mathbb{C}) \times M_n(\mathbb{C}) & \xrightarrow{\psi} & M_{n\times(n+1)m}(\mathbb{C}) \\ (A, B) & \mapsto & \begin{bmatrix} A & BA & B^2A & \dots & B^{n-1}A & B^nA \end{bmatrix} \end{array}$$

The matrix $\psi(A, B)$ determines a linear map $\psi_{(A,B)} : \mathbb{C}^{(n+1)m} \longrightarrow \mathbb{C}^n$ and (A, B) is a completely controllable pair if and only if the corresponding linear map $\psi_{(A,B)}$ is surjective. Moreover, there is a linear action of GL_n on $M_{n\times(n+1)m}(\mathbb{C})$ by left multiplication and the map ψ is GL_n-*equivariant* meaning that $\psi(g.(A,B)) = g\psi(A,B)$.

The Kalman code allows us to find a canonical pair in the orbit $\mathcal{O}_{(A,B)}$ when (A,B) is a completely controllable pair. There is a natural one-to-one correspondence between the boxes in the Kalman code array and the columns of $\psi(A,B)$ by sending box (i,j) to the j-th column of the submatrix B^iA.

5b. Grassman manifolds.

The Kalman code induces a *barcode* on $\psi(A,B)$, that is the $n\times n$ minor of $\psi(A,B)$ determined by the columns corresponding to black boxes in the Kalman code.



$\psi(A,B)$

By construction this minor is an invertible matrix $g^{-1}\in GL_n$. The canonical element in the orbit $\mathcal{O}_{(A,B)}$ we have in mind is the pair $g.(A,B)$. It has the characteristic property that the $n\times n$ minor of its image under ψ, determined by the Kalman code is the identity matrix $\mathbb{1}_n$. The matrix $\psi(g.(A,B))$ will be denoted by $b(A,B)$ and is called barcode of the pair (A,B). We now claim that the barcode determines the orbit uniquely.

In fact, the map ψ is injective on the open set V_c of completely controllable pairs. Indeed, if

$$\begin{bmatrix} A & BA & \dots & B^nA\end{bmatrix} = \begin{bmatrix} A' & B'A' & \dots & B'^nA'\end{bmatrix}$$

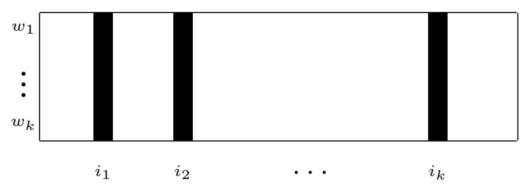
then $A = A'$, $B \mid Im(A) = B' \mid Im(A)$ and hence by induction also

$$B \mid B^iIm(A) = B' \mid B'^iIm(A') \quad \text{for all } i\leq n-1.$$

But then, $B = B'$ as both pairs (A,B) and (A',B') are completely controllable, that is, $\sum_{i=0}^{n-1} B^iIm(A) = \mathbb{C}^n = \sum_{i=0}^{n-1} B'^iIm(A')$. Hence, the barcode $b(A,B)$ determines the orbit $\mathcal{O}_{(A,B)}$ and is a point in the *Grassman manifold* $Gras_n(\mathbb{C}^{m(n+1)})$.

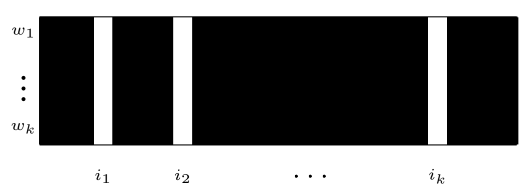
We recall briefly the definition of the Grassman manifolds. Let $k\leq l$ be integers, then the points of Grassman manifold $Gras_k(\mathbb{C}^l)$ are in one-to-one correspondence with k-dimensional subspaces of $\mathbb{C}^l$. For example, if $k = 1$ then $Gras_1(\mathbb{C}^l)$ is the projective $l-1$-space $\mathbb{P}^{l-1}$. We know that projective space can be covered by affine spaces defining a manifold structure on it. Also Grassman manifold admit a cover by affine spaces.

Let W be a k-dimensional subspace of $\mathbb{C}^l$ then fixing a basis $\{w_1, \ldots, w_k\}$ of W determines an $k \times l$ matrix M having as i-th row the coordinates of w_i with respect to the standard basis of $\mathbb{C}^l$. Linear independence of the vectors w_i means that there is a barcode design I on M



where $I = 1 \leq i_1 < i_2 < \ldots < i_k \leq l$ such that the corresponding $k \times k$ minor M_I of M is invertible. Observe that M can have several such designs.

Conversely, given a $k \times l$ matrix M of rank k determines a k-dimensional subspace of l spanned by the transposed rows. Two $k \times l$ M and M' matrices of rank k determine the same subspace provided there is a basechange matrix $g \in GL_k$ such that $gM = M'$. That is, we can identify $Grass_k(\mathbb{C}^l)$ with the orbit space of the linear action of GL_k by left multiplication on the open set $M^{max}_{k \times l}(\mathbb{C})$ of $M_{k \times l}(\mathbb{C})$ of matrices of maximal rank. Let I be a barcode design and consider the subset of $Grass_k(\mathbb{C}^l)(I)$ of subspaces having a matrix representation M having I as barcode design. Multiplying on the left with M_I^{-1} the GL_k-orbit $\mathcal{O}_M$ has a unique representant N with $N_I = \mathbb{1}_k$. Conversely, any matrix N with $N_I = \mathbb{1}_k$ determines a point in $Grass_k(\mathbb{C}^l)(I)$. Thus, $Grass_k(\mathbb{C}^l)(I)$ depends on $k(l-k)$ free parameters (the entries of the negative of the barcode)



and we have an identification $Grass_k(\mathbb{C}^l) \xrightarrow{\pi_I} \mathbb{C}^{k(l-k)}$. For a different barcode design I' the image $\pi_I(Grass_k(\mathbb{C}^l)(I) \cap Grass_k(\mathbb{C}^l)(I'))$ is an open subset of $\mathbb{C}^{k(l-k)}$ (one extra nonsingular minor condition) and $\pi_{I'} \circ \pi_I^{-1}$ is a diffeomorphism on this set. That is, the maps π_I provide us with an atlas and determine a manifold structure on $Grass_k(\mathbb{C}^l)$.

Applying the foregoing construction, the barcode $b(A, B)$ determined by the Kalman code determines a unique point in $Grass_n(\mathbb{C}^{m(n+1)})$. We have the following diagram

$$\begin{tikzcd} V_c \arrow[r, hook, "\psi"] \arrow[dr, "b(.)"'] & M^{max}_{n \times m(n+1)}(\mathbb{C}) \arrow[d, "\chi"] \\ & Grass_n(\mathbb{C}^{m(n+1)}) \end{tikzcd}$$

where ψ is a GL_n-equivariant embedding and χ the orbit map. Observe that both ψ and $'chi$ are clearly continuous maps, hence so is the orbit map b. Observe that the barcode matrix $b(A,B)$ shows that the stabilizer of (A,B) is trivial. Indeed, the minor of $g.b(A,B)$ determined by the Kalman code is equal to g. Moreover, continuity of b implies that the orbit $\mathcal{O}_{(A,B)}$ is closed in V_c.

Our final aim is to prove that ψ is a diffeomorphism to a locally closed submanifold of $M_{n \times m(n+1)}(\mathbb{C})$. To prove this we have to consider the *differential* of ψ. We recall briefly the definition of a differential. Consider a map

$$f = (f_1, \ldots, f_l) : \mathbb{C}^k \longrightarrow \mathbb{C}^l$$

with all the f_i differentiable complex valued maps in the coordinate functions x_i of $\mathbb{C}^k$. For a point $p \in \mathbb{C}^k$ the *differential of f at p* is a linear map

$$df_p : T_p\, \mathbb{C}^k \simeq \mathbb{C}^k \longrightarrow \mathbb{C}^l = T_{f(p)}\, \mathbb{C}^l$$

between the tangent space to $\mathbb{C}^k$ in p and the tangent space to $\mathbb{C}^l$ in the image $f(p)$. This linear map is determined by the matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(p) & \ldots & \frac{\partial f_1}{\partial x_k}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_l}{\partial x_1}(p) & \ldots & \frac{\partial f_l}{\partial x_k}(p) \end{bmatrix} \in M_{l \times k}(\mathbb{C}).$$

When all the f_i are polynomials in the variables x_i we can compute the differential map by the *ϵ-method* : compute $f_i(p + \epsilon v)$ bearing in mind that $\epsilon^2 = 0$, then one has

$$f(p + \epsilon v) = f(p) + \epsilon\; df_p(v) \qquad \text{for all } v \in T_p\, \mathbb{C}^k.$$

If df_p is injective, then the *implicit function theorem* implies that the image of f is locally around $f(p)$ a *closed submanifold* of dimension n in $\mathbb{C}^l$.

All the coordinate functions of ψ are polynomials in the coordinates of W. For all $(A,B) \in W$ and $(X,Y) \in T_{(A,B)}(W) \simeq W$ we have

$$(B + \epsilon Y)^j (A + \epsilon X) = B^n A + \epsilon\; (B^n X + \sum_{i=0}^{j-1} B^i Y B^{n-1-i} A).$$

Therefore the differential of ψ in $(A,B) \in W$, $d\psi_{(A,B)}(X,Y)$ is equal to

$$\begin{bmatrix} X & BX+YA & B^2X+BYA+YBA & \dots & B^nX+\sum_{i=0}^{n-1} B^iYB^{n-1-i}A \end{bmatrix}.$$

Assume $d\psi_{(A,B)}(X,Y)$ is the zero matrix, then $X=0$ and substituting in the next term also $YA=0$. Substituting in the third gives $YBA=0$, then in the fourth $YB^2A=0$ and so on until $YB^{n-1}A=0$. But then,

$$Y\begin{bmatrix} A & BA & B^2A & \dots & B^{n-1}A \end{bmatrix} = 0.$$

If (A,B) is a completely controllable pair, this implies that $Y=0$ and hence shows that $d\psi_{(A,B)}$ is injective for all $(A,B) \in V_c$.

By the implicit function theorem, ψ induces a GL_n-equivariant diffeomorphism between the open subset V_c of completely controllable pairs and a locally closed submanifold of $M_{n\times(n+1)m}(\mathbb{C})^{max}$. The image of this submanifold under the orbit map χ is again a manifold as all fibers are equal to GL_n. This concludes the difficult part of the *Kalman theorem* :

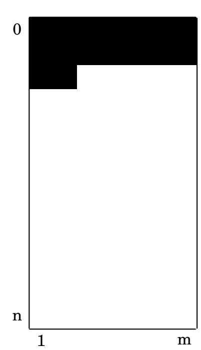
Theorem 5.2. *The orbit space $O_c = V_c/GL_n$ of equivalence classes of completely controllable pairs is a locally closed submanifold of dimension $m.n$ of the Grassman manifold $Grass_n(\mathbb{C}^{m(n+1)})$. In fact $V_c \xrightarrow{b} O_c$ is a principal GL_n-bundle.*

To prove the dimension statement, consider $V_c(K)$ the set of completely controllable pairs (A,B) having Kalman code K and let $O_c(K)$ be the image under the orbit map. After identifying $V_c(K)$ with its image under ψ, the barcode matrix $b(A,B)$ gives a section $O_c(K) \xhookrightarrow{s} V_c(K)$. In fact,

$$GL_n \times O_c(K) \longrightarrow V_c(K)$$

$$(g,x) \mapsto g.s(x)$$

is a GL_n-equivariant diffeomorphism because the $n \times n$ minor determined by K of $g.b(A,B)$ is g. Apply the local product decomposition to the *generic* Kalman code K^g



obtained by painting the top boxes black from left to right until one has n black boxes. Clearly $V_c(K^g)$ is open in V_c and one deduces

$$dim\ O_c = dim\ O_c(K^g) = dim\ V_c(K^g) - dim\ GL_n = mn + n^2 - n^2 = mn.$$

The Kalman theorem 5.2. implies immediately the existence of an orbit space for completely controllable and Schurian systems. Indeed, let $\Sigma = (A,B,C)$ completely controllable and let $g = g_{(A,B)} \in GL_n$ be the uniquely determined basechange such that $g.(A,B) = b(A,B)$, then we have a canonical representant (gA, gBg^{-1}, Cg^{-1}) in the orbit $\mathcal{O}_\Sigma$. As the stabilizer $Stab(A,B)$ is trivial the orbits of (A,B,C) and (A,B,C') are distinct if $C = C'$. That is the natural projection pr_3

$$\begin{tikzcd}
Sys_c \arrow[r, twoheadrightarrow, "pr_3"] \arrow[d, dotted] & V_c \arrow[d, "b"] \\
Sys_c/GL_n \arrow[r, twoheadrightarrow] & O_c
\end{tikzcd}$$

descends to define an orbit space which is an $M_{p\times n}(\mathbb{C})$- bundle over O_c and hence is a manifold. The Schurian systems Sys_s form a GL_n-stable open subset of Sys_c and hence their orbit space is an open submanifold of Sys_c/GL_n. This concludes the solution to problem 5 :

Theorem 5.3. *Let Sys_c (resp. Sys_s) the the open subset of $Sys = M_{n\times m}(\mathbb{C}) \times M_n(\mathbb{C}) \times M_{p\times n}(\mathbb{C})$ determined by the completely controllable (resp. Schurian) linear dynamical systems.*

1. *The orbit space for the GL_n action on Sys_c exists and is a vectorbundle of rank pn over O_c.*

2. *The orbit space for the GL_n-action on Sys_s exists and is a manifold of dimension mn^2p.*

Week 6

The slice space.

Another important problem in system theory is to determine how a system can change under small perturbations. That is, given a completely controllable pair (A, B) or system $\Sigma = (A, B, C)$, we want to construct a *slice* giving unique representants in nearby orbits.

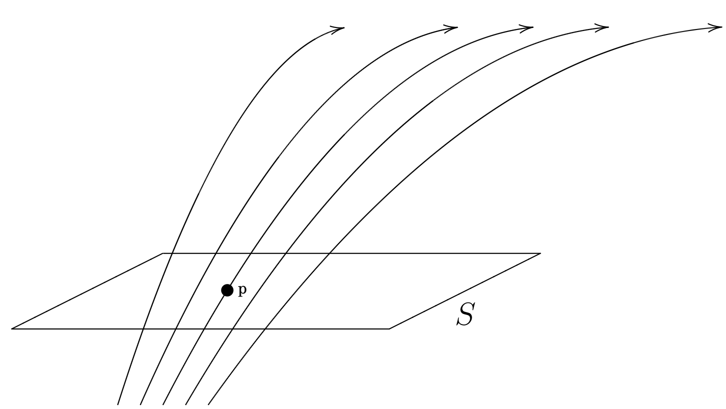
Problem 5.

Construct a slice for completely controllable systems.

Consider first the case of a controllable pair $p = (A, B)$. By a *slice* we mean a submanifold S of $V = M_{n\times m}(\mathbb{C}) \times M_n(\mathbb{C})$ passing through (A, B) and such that we have a GL_n-equivariant diffeomorphism

$$GL_n \times S \longrightarrow V \quad \text{given by} \quad (g, s) \mapsto g.s$$

in a neighborhood of (A, B). Hence, near (A, B) an orbit intersects S uniquely. We have the following situation



6a. Slice representation.

In order to determine a good candidate for S let us compute the differential of the *action map*

$$GL_n \times V \xrightarrow{a} V \quad \text{defined by} \quad (g,(A',B')) \mapsto (gA', gBg^{-1})$$

in the point $(\mathbb{1}_n, p)$ where $p = (A,B)$. Because GL_n is an open submanifold of $M_n(\mathbb{C})$ we know that $T_{\mathbb{1}_n}\, GL_n = M_n(\mathbb{C}) = \{\mathbb{1}_n + \epsilon\, m \mid m \in M_n(\mathbb{C})\}$. Similarly, the tangent space $T_p\, V = V$ and can be identified with $\{(A + \epsilon\, X, B + \epsilon\, Y) \mid X \in M_{n\times m}(\mathbb{C}), Y \in M_n(\mathbb{C})\}$. By the ϵ-method to compute the differential we have to expand

$$(\ (\mathbb{1}_n + \epsilon\, m)(A + \epsilon\, X)\ ,\ (\mathbb{1}_n + \epsilon\, m)(B + \epsilon\, Y)(\mathbb{1}_n - \epsilon\, m)\ )$$

which is equal to

$$(A,B) + \epsilon\,(X + mA, Y + [m,B]) \quad \text{where} \quad [m,B] = mB - Bm.$$

That is, the differential of the action map in the point $(\mathbb{1}_n, p)$ is given by

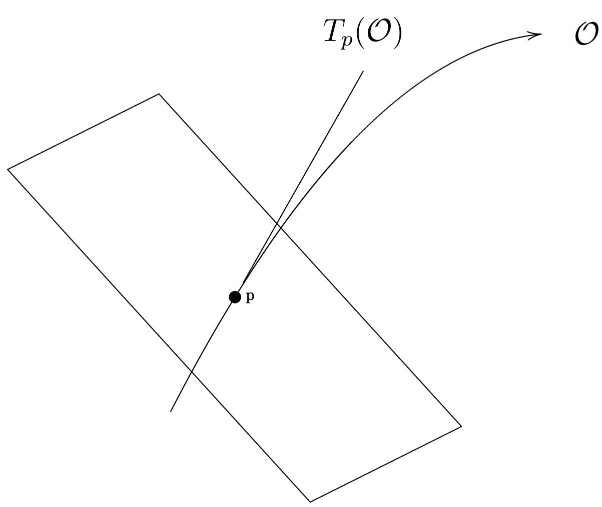
$$da_{(\mathbb{1}_n,p)}(m,(X,Y)) = (X + mA, Y + [m,B]).$$

Likewise, we can compute that the differential of the *orbit map* $GL_n \xrightarrow{o} V$ of $\mathcal{O}_{(A,B)}$ defined by $g \mapsto g.(A,B)$ in the point $\mathbb{1}_n$ is given by the linear map

$$do_{\mathbb{1}_n}(m) = (mA, [m,B]).$$

Observe that $do_{\mathbb{1}_n}$ is injective. For if $mA = 0$ and $mB = Bm$ then for all i we have that $mB^iA = B^imA = 0$ whence $m\begin{bmatrix} A & BA & \dots & B^{n-1}A \end{bmatrix} = 0$. But then $m = 0$ by complete controllability of (A,B).

The tangent space $T_p\, \mathcal{O}_{(A,B)}$ in $p = (A,B)$ to the orbit is the subset of pairs $(A,B) + Im\ do_{\mathbb{1}_n} = \{(A + mA, B + [m,B]) \mid m \in M_n(\mathbb{C})\}$. Take S to be the normal space in p to the orbit $\mathcal{O}_{(A,B)}$. That is, $S = (A,B) + (Im\ do_{\mathbb{1}_n})^{\perp}$ where $(Im\ do_{\mathbb{1}_n})^{\perp}$ is subspace of V orthogonal to $Im\ do_{\mathbb{1}_n}$.



Here, orthogonality is meant with respect to the *Hermitian inner product* on $V = M_{n\times m}(\mathbb{C}) \times M_n(\mathbb{C})$ defined by

$$\langle (A_1,B_1),(A_2,B_2)\rangle = Tr(A_1\overline{A_2}^{tr}) + Tr(B_1\overline{B_2}^{tr})$$

where Tr is the trace on $n\times n$ matrices and $\overline{X}^{tr}$ is the Hermitian transpose, that is if $U = (u_{ij})_{i,j}$, then $\overline{U}^{tr} = (\overline{u}_{ij})_{j,i}$. Now, $(Im\ do_{\mathbb{1}_n})^\perp$ is the subspace of pairs (X,Y) such that for all $m \in M_n(\mathbb{C})$ we have

$$\begin{aligned} 0 = \langle (mA,[m,B]),(X,Y)\rangle &= Tr(mA\overline{X}^{tr}) + Tr([m,B]\overline{Y}^{tr}) \\ &= Tr(mA\overline{X}^{tr}) + Tr(m[B,\overline{Y}^{tr}]) \\ &= Tr(m(A\overline{X}^{tr} + [B,\overline{Y}^{tr}])) \end{aligned}$$

where we used that traces of products are the same for cyclic permutations of the factors. By the non-degeneracy of the trace on $M_n(\mathbb{C})$ (that is, $n = 0$ if and only if $Tr(mn) = 0$ for all $m \in M_n(\mathbb{C})$) we conclude from this that

$$(Im\ do_{\mathbb{1}_n})^\perp = \{(X,Y) \mid A\overline{X}^{tr} + [B,\overline{Y}^{tr}] = 0\}.$$

Theorem 6.1. *Let $p = (A,B)$ be a completely controllable pair. Then the manifold*

$$S = \{(A+X,B+Y) \mid A\overline{X}^{tr} + [B,\overline{Y}^{tr}] = 0\}$$

is a slice for the action of GL_n on V in $p = (A,B)$.

By the implicit function theorem the action map $GL_n \times S \xrightarrow{a} V$ is a diffeomorphism in $p = (A,B)$ if and only if the differential

$$da_{(\mathbb{1}_n,p)} : T_{\mathbb{1}_n}\ GL_n \oplus T_p\ S \longrightarrow T_p\ V = V$$

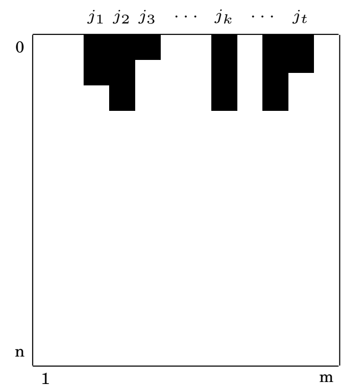
is a linear isomorphism. We have seen that $T_{\mathbb{1}_n}\ GL_n = M_n(\mathbb{C}) = Im\ do_{\mathbb{1}_n}$ and that $T_p\ S = (Im\ do_{\mathbb{1}_n})^\perp$ so the dimensions of both left factors add up to $dim\ V$. It therefore suffices to check injectivity of the differential. We have calculated that

$$da_{(\mathbb{1}_n,p)}(m,(X,Y)) = (mA+X,[m,B]+Y)$$

Assume that $X = -mA$ and $Y = -[m,B]$ then because $(X,Y) \in (Im\ do_{\mathbb{1}_n})^\perp$ we have $\langle (X,Y),(X,Y)\rangle = 0$ whence $X = 0$ and $Y = 0$. But then as $mA = 0$ and $mB = Bm$ and (A,B) is completely controllable we also have $m = 0$ finishing the proof of the slice theorem.

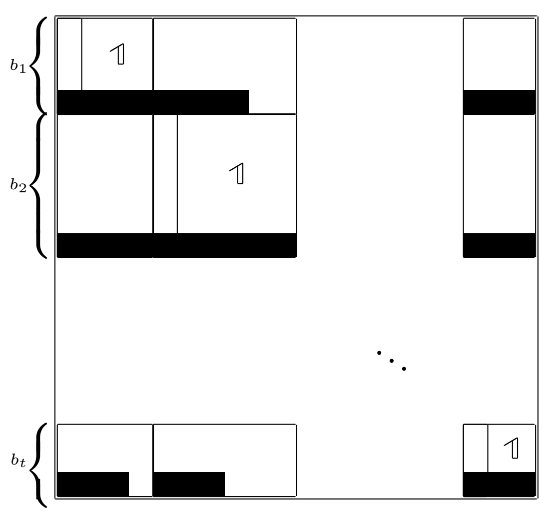
6b. Control canonical form.

The determination of the slice S is particularly simple in the point (A', B') of the orbit $\mathcal{O}_{(A,B)}$ where (A', B') is in *control canonical form*. This canonical form is similar to the Jordan normal form and is determined by the Kalman code K associated to the completely controllable pair (A, B). Assume K has the following shape :

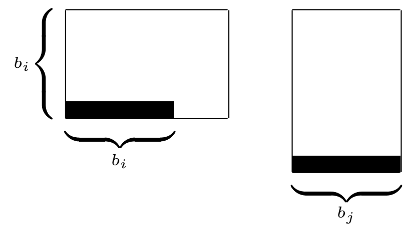


Here, $1 \leq j_1 < \ldots < j_t \leq m$ form the set J of integers j such that the box $(0, j)$ is painted black, hence $t = \# J = rank\ A$. For each $j_k \in J$ define b_k to be the number of painted boxes in the j-th column of K.

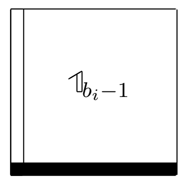
With notations as above, prove that there is a base-change matrix $g \in GL_n$ such that $g.(A, B) = (A', B')$ where the $n \times n$ matrix B' has the following block form



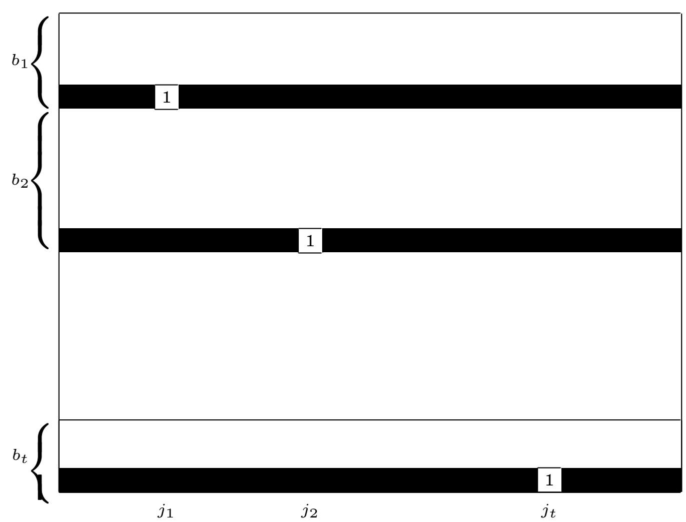
where the nonzero entries are in one of the identity matrix components $\mathbb{1}_{b_k - 1}$ or in the painted stripes which represent one row. The off-diagonal block B_{ij} is a $b_i \times b_j$ matrix of one of the following types



depending on whether $b_i \leq b_j$ or $b_j \leq b_i$. The diagonal block B_{ii} is a $b_i \times b_i$ matrix of the form



The $n \times m$ matrix A has the following block structure :



6c. Another slice result.

Finally, we will generalize the slice construction to dynamical systems $\Sigma = (A, B, C)$ which are completely controllable. We define an Hermitian inner product on $Sys = M_{n \times m}(\mathbb{C}) \times M_n(\mathbb{C}) \times M_{p \times n}(\mathbb{C})$ by the rule

$$\langle (A_1, B_1, C_1), (A_2, B_2, C_2) \rangle = Tr(A_1 \overline{A}_2^{tr}) + Tr(B_1 \overline{B}_2^{tr}) + Tr(\overline{C}_2^{tr} C_1).$$

This time, the differential of the orbit map $GL_n \xrightarrow{o} Sys$ defined by $g \mapsto g.(A, B, C)$ is computed by expanding the expression

$$(\ (\mathbb{1}_n + \epsilon\ m)(A + \epsilon\ X)\ ,\ (\mathbb{1}_n + \epsilon\ m)(B + \epsilon\ Y)(\mathbb{1}_n - \epsilon\ m)\ ,\ (C + \epsilon\ Z)(\mathbb{1}_n - \epsilon\ m)\ )$$

and we obtain that the image of the orbit map $do_{\mathbb{1}_n}$ consists of the triples

$$Im\ do_{\mathbb{1}_n} = \{(mA, [m, B], -Cm) \mid m \in M_n(\mathbb{C})\}.$$

Again, using non-degeneracy of the trace map on $n \times n$ matrices one can identify the orthogonal complement of this space as the subspace

$$(Im\ do_{\mathbb{1}_n})^{\perp} = \{(X, Y, Z) \mid A\overline{X}^{tr} + [B, \overline{Y}^{tr}] - \overline{Z}^{tr}C = 0\}$$

From this we again construct a slice as in the proof of theorem 6.1. .

Theorem 6.2. *Let $\Sigma = (A, B, C)$ be a completely controllable system. Then the manifold*

$$S = \{(A + X, B + Y, C + Z) \mid A\overline{X}^{tr} + [B, \overline{Y}^{tr}] - \overline{Z}^{tr}C = 0\}$$

is a slice for the GL_n-action on Sys in Σ.

Week 7

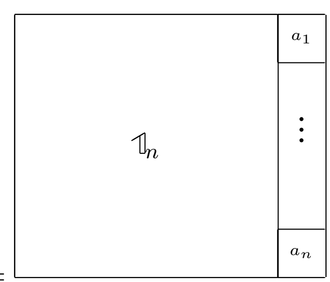
Hilbert schemes.

Consider completely controllable pairs of matrices (A, B) with $m = 1$. That is, $A = v$ is a column vector and we must have that

$$B(v) = \begin{bmatrix} v & Bv & B^2v & \dots & B^{n-1}v \end{bmatrix}$$

has rank n, or equivalently, v is a cyclic vector of B. There is only one Kalman code $K(v, B)$ possible, the generic one. The barcode $b(v, B)$ of the pair (v, B) is the matrix $B(v)^{-1} \begin{bmatrix} B(v) & B^n v \end{bmatrix}$. That is, we have :

$K(v, B) =$ 

and $b(v, B) =$ 

Hence, the orbit space $O_c(1, n) \simeq \mathbb{C}^n$ and the point $(a_1, \dots, a_n) \in \mathbb{C}^n$ corresponding to the orbit of (v, B) is determined by the characteristic polynomial of B :

$$\chi_B(t) = det(t\mathbb{1}_n - B) = t^n - a_n t^{n-1} - \dots - a_2 t - a_1.$$

Hence, we can identify the orbit space $O_c(1, n)$ with the space of all monic polynomials in $\mathbb{C}[t]$ of degree n. Such polynomials $f(t)$ are uniquely determined by their unordered n-tuple of roots $\{\lambda_1, \dots, \lambda_n\}$ as

$$f(t) = \prod_{i=1}^{n} (t - \lambda_i) = t^n - a_n t^{n-1} - \dots - a_2 t - a_1$$

where $a_{i+1} = (-1)^{n-1}\sigma_{n-i}(\lambda_1,\ldots,\lambda_n)$ with σ_j the j-th elementary symmetric function in the λ_k.

That is, we can also identify $O_c(1,n)$ with $Hilb_n(\mathbb{C}^1)$, the ***Hilbert scheme of** n **points in*** $\mathbb{C}^1$. The points of $Hilb_n(\mathbb{C}^1)$ parametrize the ideals $I \triangleleft \mathbb{C}[t]$ of codimension n and as $\mathbb{C}[t]$ is a principal ideal domain such ideals have a unique monic generator of degree n.

The *symmetric product* $S^n\,\mathbb{C}^1$ of n copies of $\mathbb{C}^1$ is defined to be

$$S^n\,\mathbb{C}^1 = \underbrace{\mathbb{C}^1 \times \ldots \times \mathbb{C}^1}_{n}/S_n$$

where the symmetric group on n-letters S_n acts on the entries. A point of $S^n\,\mathbb{C}^1$ is represented as a formal sum $\sum_i n_i[\mu_i]$ where $n_i \in \mathbb{N}$ with $\sum_i n_i = n$ and $\mu_i \in \mathbb{C}^1$. In this case, the *Hilbert-Chow map*

$$Hilb_n\,\mathbb{C}^1 \xrightarrow{\pi} S^n\,\mathbb{C}^1$$

defined by sending a monic polynomial $f(t)$ of degree n to $\sum_i n_i[\mu_i]$ where the μ_i are the roots of $f(t)$ occurring with multiplicity n_i, is a one-to-one correspondence. Observe that the numbers n_i determine (when ordered) a partition of n and that the projection

$$\underbrace{\mathbb{C}^1 \times \ldots \times \mathbb{C}^1}_{n} \twoheadrightarrow S^n\,\mathbb{C}^1$$

is defined by sending an n-tuple $(c_1,\ldots,c_n)$ to $(s_1,\ldots,s_n)$ where $s_i = \sigma_i(c_1,\ldots,c_n)$. So, we still have another description of the orbit space $O_c(1,n)$ as $S^n\,\mathbb{C}^1$. However, all these identifications are particular to dimension 1. Let us consider :

Problem 6.

Describe $Hilb_n\,\mathbb{C}^2$, the Hilbert scheme of n points in the plane $\mathbb{C}^2$. That is, parametrize all ideals $I \triangleleft \mathbb{C}[x,y]$ of codimension n.

Let $I \triangleleft \mathbb{C}[x,y]$ be such that $V = \frac{\mathbb{C}[x,y]}{I}$ is an n-dimensional vectorspace and fix a basis $\{v_1,\ldots,v_n\}$ of V. Multiplication by x (resp. y) on $\mathbb{C}[x,y]$ induces a linear operator on V and hence determines a matrix $X \in M_n(\mathbb{C})$ (resp. $Y \in M_n(\mathbb{C})$). Clearly, $[X,Y] = 0$ and they generate an n-dimensional subalgebra $\mathbb{C}[X,Y] \simeq \frac{\mathbb{C}[x,y]}{I}$ of $M_n(\mathbb{C})$. Further, the image of the unit element $1 \in \mathbb{C}[x,y]$ determines a column vector $v \in V = \mathbb{C}^n$ with the property that

$$\mathbb{C}[X,Y]v = \mathbb{C}^n.$$

Note however that the triple $(v, X, Y) \in \mathbb{C}^n \oplus M_n \oplus M_n$ is not uniquely determined by the ideal I as it depends on the choice of basis of V. If we choose a different basis $\{v'_1, \ldots, v'_n\}$ with basechange matrix $g \in GL_n$, then the corresponding triple is

$$(v', X', Y') = (gv, gXg^{-1}, gYg^{-1}).$$

Consider the vectorspace of all triples

$$H_n = \mathbb{C}^n \oplus M_n \oplus M_n \quad \text{with action} \quad g.(v, X, Y) = (gv, gXg^{-1}, gYg^{-1})$$

for all $g \in GL_n$. For later reference, we depict this action by the pattern

The above discussion shows that the ideal $I \triangleleft \mathbb{C}[x, y]$ of codimension n determines an orbit $\mathcal{O}_I$ in H_n. Conversely, let C^c_n be the subset of triples $(v, X, Y) \in H_n$ satisfying the additional conditions :

1. The matrix pair commutes : $[X, Y] = 0$, and
2. v is a cyclic vector for this pair : $\mathbb{C}[X, Y]v = \mathbb{C}^n$.

For $(v, X, Y) \in C^c_n$ we can define a map $\mathbb{C}[x, y] \xrightarrow{\phi} \mathbb{C}^n$ by sending a polynomial $f = f(x, y)$ to the vector $\phi(f) = f(X, Y)v$. By the second condition, ϕ is surjective and therefore, its kernel $I = \{f \in \mathbb{C}[x, y] \mid \phi(f) = 0\}$ is an ideal of codimension n. That is, the Hilbert scheme $Hilb_n\ \mathbb{C}^2$ of n points in the plane $\mathbb{C}^2$ is the orbit space for the GL_n-action on the subset C^c_n.

7a. An example.

Miniature 5. The Hilbert scheme $Hilb_2\ \mathbb{C}^2$.

Let us first consider the Hilbert scheme $Hilb_1\ C^2$ of one point in $\mathbb{C}^2$ which we expect to be $\mathbb{C}^2$. Indeed, $H_1 = \{(v, X, Y) \mid v, X, Y \in \mathbb{C}\}$ and any pair (X, Y) is commuting. Moreover, v is cyclic for (X, Y) if and only if $v \neq 0$. That is, $C^c_1 = \mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$. The group $GL_1 = \mathbb{C}^*$ acts via $c.(v, X, Y) = (cv, X, Y)$ and hence the triples $\{(1, X, Y)\} = \mathbb{C}^2$ parametrize the orbits of C^c_1, that is, $Hilb_1\ \mathbb{C}^2 = \mathbb{C}^2$ and the ideal I of codimension one corresponding to the point $p = (X, Y) \in \mathbb{C}^2$ is the ideal of polynomials $f \in \mathbb{C}[x, y]$ vanishing in p, $f(X, Y) = 0$.

Next, we consider the Hilbert scheme $Hilb_2\ \mathbb{C}^2$ of two points in $\mathbb{C}^2$. Let $(v, X, Y) \in C^c_2$ and assume that either X or Y has distinct eigenvalues (type a). As

$$[\begin{bmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{bmatrix}, \begin{bmatrix} a & b \\ c & d \end{bmatrix}] = \begin{bmatrix} 0 & (\nu_1 - \nu_2)b \\ (\nu_2 - \nu_1)a & 0 \end{bmatrix}$$

we have a representant in the orbit of the form

$$(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix})$$

where cyclicity of the column vector implies that $v_1 v_2 \neq 0$. The stabilizer subgroup of the matrix-pair is the group of diagonal matrices $\mathbb{C}^* \times \mathbb{C}^* \hookrightarrow GL_2$, hence the orbit has a unique representant with $v_1 = v_2 = 1$. The corresponding ideal $I \triangleleft \mathbb{C}[x,y]$ is then

$$I = \{f(x,y) \in \mathbb{C}[x,y] \mid f(\lambda_1,\mu_1) = 0 = f(\lambda_2,\mu_2)\}$$

hence these orbits in C_2^c correspond to sets of two distinct points in $\mathbb{C}^2$.

The situation is slightly more complicated when X and Y have only one eigenvalue (type b). If $(v,X,Y) \in C_2^c$ then either X or Y is not diagonalizable. But then, as

$$[\begin{bmatrix} \nu & 1 \\ 0 & \nu \end{bmatrix}, \begin{bmatrix} a & b \\ c & d \end{bmatrix}] = \begin{bmatrix} c & d-a \\ 0 & c \end{bmatrix}$$

we have a representant in the orbit of the form

$$(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} \lambda & \alpha \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} \mu & \beta \\ 0 & \mu \end{bmatrix})$$

with $[\alpha : \beta] \in \mathbb{P}^1$ and $v_2 \neq 0$. The stabilizer of the matrixpair is the subgroup

$$\{\begin{bmatrix} c & d \\ 0 & c \end{bmatrix} \mid c \neq 0\} \hookrightarrow GL_2$$

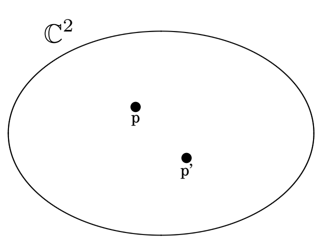
and hence we have a unique representant with $v_1 = 0$ and $v_2 = 1$. The corresponding ideal $I \triangleleft \mathbb{C}[x,y]$ is

$$I = \{f(x,y) \in \mathbb{C}[x,y] \mid f(\lambda,\mu) = 0 \text{ and } \alpha\frac{\partial f}{\partial x}(\lambda,\mu) + \beta\frac{\partial f}{\partial y}(\lambda,\mu) = 0\}$$

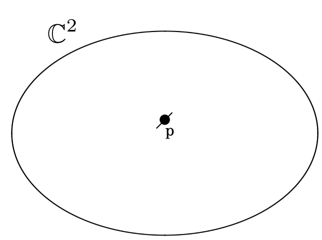
as one proves by verification on monomials because

$$\begin{bmatrix} \lambda & \alpha \\ 0 & \lambda \end{bmatrix}^k \begin{bmatrix} \mu & \beta \\ 0 & \mu \end{bmatrix}^l \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} k\alpha\lambda^{k-1}\mu^l + l\beta\lambda^k\mu^{l-1} \\ \lambda^k\mu^l \end{bmatrix}$$

Therefore, I corresponds to the set of two points at $(\lambda,\mu) \in \mathbb{C}^2$ infinitesimally attached to each other in the direction $\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}$. For each point in $\mathbb{C}^2$ there is a $\mathbb{P}^1$ family of such *fat points*. Thus, points of $Hilb_2\ \mathbb{C}^2$ correspond to either of the following two situations :



type a

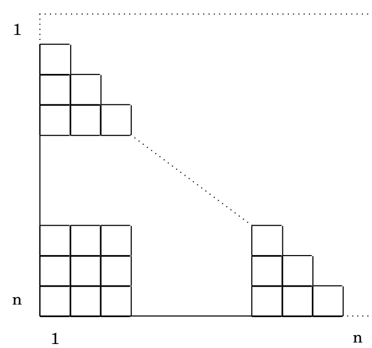


type b

The Hilbert-Chow map $Hilb_2\ \mathbb{C}^2 \xrightarrow{\pi} S^2\ \mathbb{C}^2$ sends a point of type a to the formal sum $[p] + [p']$ and a point of type b to $2[p]$. Over the complement of (the image of) the diagonal, this map is a one-to-one correspondence. However, over points on the diagonal the fibers are $\mathbb{P}^1$, soπ is not a one-to-one correspondence as in the case of $Hilb_n\ \mathbb{C}^1$. The situation is nicer for $\mathbb{C}^1$ because there points can only collide along one direction, whereas in $\mathbb{C}^2$ they can approach each other along a $\mathbb{P}^1$ family of lines leading to different ideals. In fact, the symmetric power $S^2\ \mathbb{C}^2$ has singularities and the Hilbert-Chow map $Hilb_2\ \mathbb{C}^2 \xrightarrow{\pi} S^2\ \mathbb{C}^2$ is a *resolution of singularities.*

7b. Hilbert stairs.

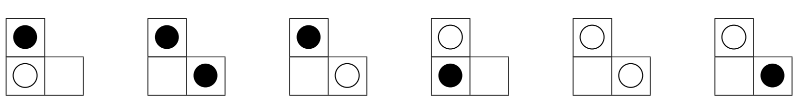
For the investigation of the GL_n-action on H_n and on the subset C_n^c we introduce a combinatorial gadget : the ***Hilbert*** *n-stair*. This is the lower triangular part of a square $n \times n$ array of boxes



filled with *go-stones* according to the following two rules :

- each row contains exactly one stone, and
- each column contains at most one stone of each color.

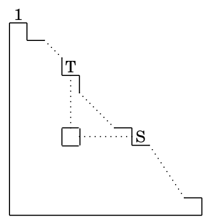
For example, the set of all possible Hilbert 3-stairs is given below.



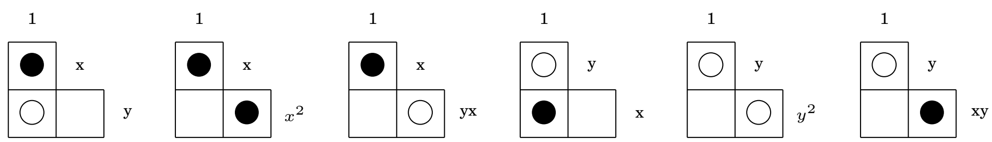
Let $\mathbb{C}\langle x, y\rangle$ be the *free associative algebra* on the non-commuting variables x and y. That is, $\mathbb{C}\langle x, y\rangle$ is the vectorspace with basis all *words* in x and y and with multiplication induced by *concatenation* of words. To every Hilbert stair we will now associate a sequence of words in x and y.

At the top of the stairs we place the identity element 1. Then, we descend the stairs according to the following rule. Every go-stone has a *top word T* which we may assume we have constructed before and a *side word*

S and they are related as indicated below



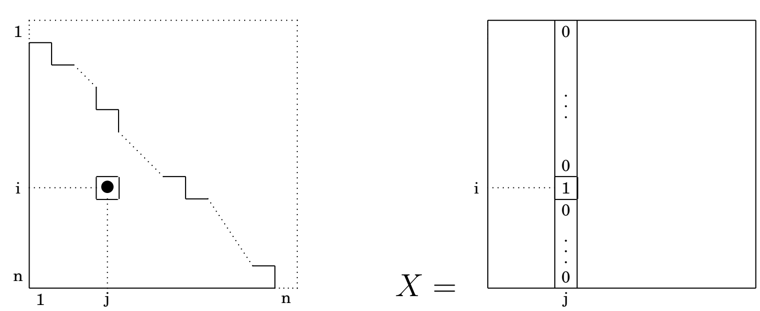
For example, for the Hilbert 3-stairs we have the following sequences of non-commutative words



Let σ be a Hilbert n-stair with associated sequence of non-commutative words $W(\sigma) = \{1, w_2(x,y), \ldots, w_n(x,y)\}$. Let $(v, X, Y) \in H_n$ then replacing each occurrence of x in the word $w_i(x,y)$ by X and each occurrence of y by Y we obtain an $n \times n$ matrix $w_i(X,Y) \in M_n(\mathbb{C})$ and by left multiplication a column vector $w_i(X,Y)v$. We call the *evaluation of* σ *in* (v, X, Y) the determinant of the $n \times n$ matrix

$$\sigma(v, X, Y) = det \begin{bmatrix} v & w_1(X,Y)v & w_2(X,Y)v & \ldots & w_n(X,Y)v \end{bmatrix}.$$

For a fixed Hilbert n-stair σ we denote with $H_n(\sigma)$ the subset of triples $(v, X, Y) \in H_n$ with non-zero evaluation $\sigma(v, X, Y) \neq 0$. We claim that none of the $H_n(\sigma)$ is empty. Indeed, let v be the basic column vector $e_1 = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}^{tr}$ and let every black stone in the Hilbert stair σ fix a column of X by the rule that if it lies in box (i, j) the j-th column of X is the basic column vector $e_i = \begin{bmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix}^{tr}$ (a 1 at place i)



and the same rule applies to white stones determining columns of Y. That is, one replaces each stone in σ by 1 at the same spot in X or Y and fills the remaining spots in the same column by zeroes. We say that such a triple

(v, X, Y) is in σ-standard form. With these conventions one easily verifies by induction that

$$w_i(X,Y)e_1 = e_i \quad \text{for all } 2 \leq i \leq n.$$

Hence, filling up the remaining spots in X and Y arbitrarily one has that $\sigma(v, X, Y) \neq 0$ proving the claim. Hence, $H_n(\sigma)$ is an open subset of H_n for every Hilbert n-stair σ. Further, for every word (monomial) $w(x, y)$ and every $g \in GL_n$ we have that

$$w(gXg^{-1}, gYg^{-1})gv = gw(X,Y)v$$

and therefore the open sets $H_n(\sigma)$ are stable under the GL_n-action on H_n. We will give representants of the orbits in $H_n(\sigma)$.

Let $W_n = \{1, x, y, xy, \ldots, y^n\}$ be the set of all words in the non-commuting variables x and y of length $\leq n$, ordered lexicographically. For every triple $(v, X, Y) \in H_n$ consider the $n \times m$ matrix

$$\psi(v, X, Y) = \begin{bmatrix} v & Xv & Yv & XYv & \ldots & Y^n v \end{bmatrix}$$

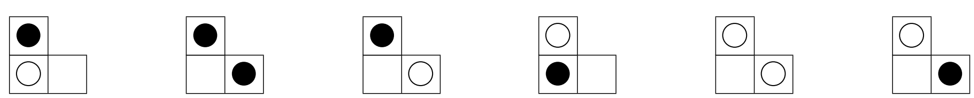
where $m = 2^{n+1} - 1$ and the j-th column is the column vector $w(X, Y)v$ with $w(x, y)$ the j-th word in W_n. Hence, $(v, X, Y) \in H_n(\sigma)$ if and only if the $n \times n$ minor of $\psi(v, X, Y)$ determined by the word-sequence $\{1, w_2, \ldots, w_n\}$ of σ is invertible. Moreover, as

$$\psi(gv, gXg^{-1}, gYg^{-1}) = g\psi(v, X, Y)$$

we deduce that the GL_n-orbit of $(v, X, Y) \in H_n(\sigma)$ contains a *unique* triple (v', X', Y') such that the corresponding minor of $\psi(v', X', Y') = \mathbb{1}_n$. Hence, each GL_n-orbit in $H_n(\sigma)$ contains a unique representant in σ-standard form. Therefore, the orbit space $O_n(\sigma)$ of $H_n(\sigma)$ is an affine space of dimension the number of non-forced entries in X and Y. As we fixed $n-1$ columns in X or Y this dimension is equal to

$$O_c(\sigma) = \mathbb{C}^k \quad \text{with} \quad k = 2n^2 - (n-1)n = n^2 + n.$$

For example, representants for the orbits in $H_3(\sigma)$ are given by (v, X, Y) with $v = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{tr}$ and



$$\begin{array}{c|cccccc}
X & \begin{bmatrix} 0 & a & b \\ 1 & c & d \\ 0 & e & f \end{bmatrix} & \begin{bmatrix} 0 & 0 & a \\ 1 & 0 & b \\ 0 & 1 & c \end{bmatrix} & \begin{bmatrix} 0 & a & b \\ 1 & c & d \\ 0 & e & f \end{bmatrix} & \begin{bmatrix} 0 & a & b \\ 0 & c & d \\ 1 & e & f \end{bmatrix} & \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} & \begin{bmatrix} a & 0 & b \\ c & 0 & d \\ e & 1 & f \end{bmatrix} \\
 & & & & & & \\
Y & \begin{bmatrix} 0 & g & h \\ 0 & i & j \\ 1 & k & l \end{bmatrix} & \begin{bmatrix} d & e & f \\ g & h & i \\ j & k & l \end{bmatrix} & \begin{bmatrix} g & 0 & h \\ i & 0 & j \\ k & 1 & l \end{bmatrix} & \begin{bmatrix} 0 & g & h \\ 1 & i & j \\ 0 & k & l \end{bmatrix} & \begin{bmatrix} 0 & 0 & j \\ 1 & 0 & k \\ 0 & 1 & l \end{bmatrix} & \begin{bmatrix} 0 & g & h \\ 1 & i & j \\ 0 & k & l \end{bmatrix}
\end{array}$$

We define the subset H_n^c of *cyclic triples*, that is those $(v, X, Y) \in H_n$ such that there exists *no* proper subspace $W \subset \mathbb{C}^n$ containing v and stable under the action of X and Y, that is, $XW \subset W$ and $YW \subset W$. If we denote with $\mathbb{C}\langle X, Y\rangle$ the (not necessarily commutative) $\mathbb{C}$-subalgebra of $M_n(\mathbb{C})$ generated by the matrix-pair (X, Y), then (v, X, Y) is a cyclic triple if and only if

$$\mathbb{C}\langle X, Y\rangle v = \mathbb{C}^n.$$

Hence, clearly $H_n(\sigma) \subset H_n^c$ for any Hilbert n-stair σ. Conversely, we claim that a cyclic triple $(v, X, Y) \in H_n^c$ belongs to at least one of the open subsets $H_n(\sigma)$. Indeed, either $Xv \notin \mathbb{C}v$ or $Yv \notin \mathbb{C}v$ as otherwise the subspace $W = \mathbb{C}v$ would contradict the cyclicity assumption. Fill the top box of the stairs with the corresponding stone and define the 2-dimensional subspace $V_2 = \mathbb{C}v_1 + \mathbb{C}v_2$ where $v_1 = v$ and $v_2 = w_2(X, Y)v$ with w_2 the corresponding word (either x or y).

Assume by induction we have been able to fill the first i rows of the stairs with stones leading to the sequence of words $\{1, w_2(x, y), \ldots, w_i(x, y)\}$ such that the subspace $V_i = \mathbb{C}v_1 + \ldots + \mathbb{C}v_i$ with $v_i = w_i(X, Y)v$ has dimension i. Then, either $Xv_j \notin V_i$ for some j or $Yv_j \notin V_i$ (if not V_i would contradict cyclicity). Then fill the j-th box in the $i+1$-th row of the stairs with the corresponding stone. Then, the top $i+1$ rows of the stairs form a Hilbert $i+1$-stair as there can be no stone of the same color lying in the same column. Define $w_{i+1}(x, y) = xw_i(x, y)$ (or $yw_i(x, y)$) and $v_{i+1} = w_{i+1}(X, Y)v$. Then, $V_{i+1} = \mathbb{C}v_1 + \ldots + \mathbb{C}v_{i+1}$ has dimension $i+1$. Continuing we end up with a Hilbert n-stair σ such that $(v, X, Y) \in H_n(\sigma)$. This concludes the proof of the following result.

Theorem 7.1. *The orbit space O_n^c for the GL_n-action on the open submanifold H_n^c of cyclic triples is a manifold of dimension $n^2 + n$.*

Week 8

Hilbert manifolds.

Recall that C_n^c is the subset of the affine space $H_n = \mathbb{C}^n \oplus M_n \oplus M_n$ consisting of those triples (v, X, Y) such that $[X, Y] = 0$ and v is a cyclic vector for the pair (X, Y). Clearly, C_n^c is a subset of the open submanifold of cyclic triples H_n^c.

We aim to show that C_n^c is a manifold of dimension $n^2 + 2n$. This is slightly surprising as the closely related *commuting variety*

$$C = \{(X, Y) \in M_n \oplus M_n \mid [X, Y] = 0\}$$

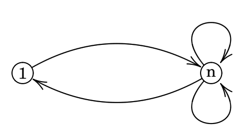
is known to have singularities. In fact, not much is known about this commuting variety except that it is *connected*.

8a. A slice argument.

Consider the vectorspace $Q_n = \mathbb{C}^n \oplus C^{n*} \oplus M_n \oplus M_n$ where $\mathbb{C}^{n*}$ denotes the n-dimensional vectorspace of row vectors. We define a linear action of GL_n on Q_n by the rule

$$g.(v, w, X, Y) = (gv, wg^{-1}, gXg^{-1}, gYg^{-1})$$

for all $g \in GL_n$. For later reference, we depict this action by the pattern



Observe that modding out the row vectors gives a surjection $Q_n \overset{\gamma}{\twoheadrightarrow} H_n$ which is GL_n-equivariant. With Q_n^c we will denote the open submanifold

$\gamma^{-1}(Q_n^c)$ of *cyclic quartets*. Consider the map

$$\begin{array}{ccc} Q_n & \xrightarrow{\quad\chi\quad} & M_n \\ \\ (v,w,X,Y) & \mapsto & v.w + [X,Y] \end{array}$$

The differential $d\chi$ of this map is computed by expanding $(v+\epsilon a)(w+\epsilon b)$ $+[(X+\epsilon C),(Y+\epsilon D)]$ which shows

$$d\chi_{(v,w,X,Y)}(a,b,C,D) = v.b + a.w + [X,D] + [C,Y]$$

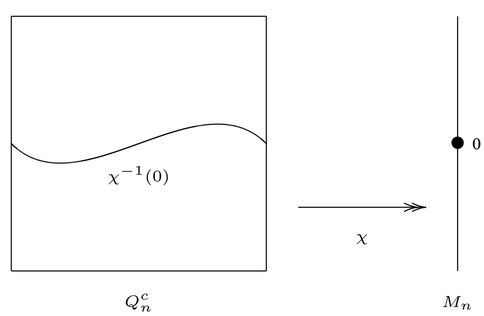
We claim that the differential is surjective whenever (v,w,X,Y) is a cyclic quartet. Consider the Hermitian inner product $\langle M,N\rangle = Tr(M\overline{N}^{tr})$ on $M_n(\mathbb{C})$, then the space orthogonal to the image of $d\chi_{(v,w,X,Y)}$ is equal to

$$\{M \in M_n(\mathbb{C}) \mid Tr(vb\overline{M}^{tr} + aw\overline{M}^{tr} + [X,D]\overline{M}^{tr} + [C,Y]\overline{M}^{tr}) = 0, \forall (a,b,C,D)\}$$

Because Tr does not change under cyclic permutation and is non degenerate on $M_n(\mathbb{C})$, we see that this orthogonal space is equal to

$$\{M \in M_n(\mathbb{C}) \mid \overline{M}^{tr}v = 0 \quad w\overline{M}^{tr} = 0 \quad [\overline{M}^{tr}, X] = 0 \quad \text{and} \quad [Y, \overline{M}^{tr}] = 0\}.$$

If (v,w,X,Y) is a cyclic quartet, for such a matrix M we have that the nullspace $Ker\ \overline{M}^{tr}$ is a proper subspace of $\mathbb{C}^n$ containing v and stable under X and Y. By the cyclicity condition this implies that $Ker\ \overline{M}^{tr} = \mathbb{C}^n$ or equivalently that $\overline{M}^{tr} = 0$, proving the claim. By the implicit function theorem this implies that the fiber of any point in the image of χ



is a submanifold of Q_n^c of dimension $dim\ Q_n - dim\ M_n = n^2+2n$. In particular, $\chi^{-1}(0)$ is a submanifold of dimension n^2+2n. We will now identify the fiber $\chi^{-1}(0)$ with the subset C_n^c and hence prove :

Proposition 8.1. *The subset C_n^c of cyclic triples (v,X,Y) with $[X,Y]=0$ is a submanifold of dimension n^2+2n of H_n.*

Indeed, let (v,w,X,Y) be a cyclic quartet satisfying $v.w+[X,Y]=0$ and let $m(x,y)$ be any word in the non-commuting variables x and y. We claim that $wm(X,Y)v=0$. We prove this by induction on the length $l(m)$ of the word m. If $l(m)=0$ then $m=1$ and we have

$$wm(X,Y)v = wv = Tr(v.w) = -Tr([X,Y]) = 0.$$

Assume we proved the claim for all words of length $< l$ and take a word of the form $m(x,y)=m_1(x,y)yxm_2(x,y)$ with $l(m_1)+l(m_2)+2=l$. Then, we have

$$\begin{aligned} wm(X,Y) &= wm_1(X,Y)YXm_2(X,Y) \\ &= wm_1(X,Y)([Y,X]+XY)m_2(X,Y) \\ &= (wm_1(X,Y)v).wm_2(X,Y)+wm_1(X,Y)XYm_2(X,Y) \\ &= wm_1(X,Y)XYm_2(X,Y) \end{aligned}$$

where we used the induction hypotheses in the last equality (the bracketed term vanishes). Hence we can reorder the terms in $m(x,y)$ if necessary and have that $wm(X,Y)=wX^{l_1}Y^{l_2}$ with $l_1+l_2=l$ and l_1 the number of occurrences of x in $m(x,y)$. Hence, we have to prove the claim for $X^{l_1}Y^{l_2}$.

$$\begin{aligned} wX^{l_1}Y^{l_2}v &= Tr(X^{l_1}Y^{l_2}vw) = -Tr(X^{l_1}Y^{l_2}[X,Y]) \\ &= -Tr([X^{l_1}Y^{l_2},X]Y) = -Tr(X^{l_1}[Y^{l_2},X]Y) \\ &= -\sum_{i=0}^{l_2-1} Tr(X^{l_1}Y^i[Y,X]Y^{l_2-i}) = -\sum_{i=0}^{l_2-1} Tr(Y^{l_2-i}X^{l_1}Y^i[Y,X] \\ &= -\sum_{i=0}^{l_2-1} Tr(Y^{l_2-i}X^{l_1}Y^iv.w = -\sum_{i=0}^{l_2-1} wY^{m_2-i}X^{l_1}Y^iv \end{aligned}$$

But we have seen that $wY^{l_2-i}X^{l_1}Y^i = wX^{l_1}Y^{l_2}$ hence the above implies that $wX^{l_1}Y^{l_2}v = -l_2wX^{l_1}Y^{l_2}v$. But then $wX^{l_1}Y^{l_2}v=0$, proving the claim.

Consequently, $w\mathbb{C}\langle X,Y\rangle v = 0$ and by the cyclicity condition we have $w\mathbb{C}^n = 0$ hence $w=0$. Finally, as $v.w+[X,Y]=0$ this implies that $[X,Y]=0$ and we can identify the fiber $\chi^{-1}(0)$ with C_n^c (identifying H_n with the closed submanifold of Q_n where $w=0$), finishing the proof of proposition 8.1. . Recalling that the Hilbert scheme $Hilb_n\ \mathbb{C}^2$ is the orbit space of the GL_n-action on C_n^c we have the situation

$$\begin{tikzcd} C_n^c \arrow[r, hook] \arrow[d] & H_n^c \arrow[d] \\ Hilb_n\ \mathbb{C}^2 \arrow[r, hook] & O_n^c \end{tikzcd}$$

We will construct a slice for the GL_n-action on C_n^c. First, consider the orbit map $GL_n \xrightarrow{o} H_n$ for a point $(v,X,Y)\in H_n^c$ defined by $g \mapsto g.(v,X,Y)$. The differential of this map in the point $\mathbb{1}_n$ is given by the linear map

$$do_{\mathbb{1}_n} : M_n(\mathbb{C}) \longrightarrow H_n \quad \text{where} \quad do_{\mathbb{1}_n}(m) = (mv,[m,X],[m,Y])$$

as one verifies as in the previous section. Observe that this differential is injective whenever $(v,X,Y) \in H_n^c$. Indeed, if m satisfies $mv = 0$ and $[m,X] = 0 = [m,Y]$ then the nullspace $Ker\ m$ is a subspace of $\mathbb{C}^n$ containing v and stable under X and Y so must be $\mathbb{C}^n$ whence $m = 0$. Define an Hermitian inner product on the vectorspace H_n by the rule

$$\langle (v_1,X_1,Y_1),(v_2,X_2,Y_2)\rangle = Tr(v_1\overline{v}_2^{tr} + X_1\overline{X}_2^{tr} + Y_1\overline{Y}_2^{tr}).$$

The subspace of the tangentspace H_n in (v,X,Y) orthogonal to $Im\ do_{\mathbb{1}_n}$ is then the subspace of triples (a,C,D) such that

$$\begin{aligned} 0 &= Tr(mv\overline{a}^{tr}) + Tr([m,X]\overline{C}^{tr}) + Tr([m,Y]\overline{D}^{tr}) \\ &= Tr(m(v\overline{a}^{tr} + [X,\overline{C}^{tr}] + [Y,\overline{D}^{tr}]) \end{aligned}$$

Again using the non-degeneracy of the trace map on $n \times n$ matrices, we have

$$(Im\ do_{\mathbb{1}_n})^{\perp} = \{(a,C,D) \in H_n \mid v\overline{a}^{tr} + [X,\overline{C}^{tr}] + [Y,\overline{D}^{tr}] = 0\}$$

which is of dimension n^2+n (using injectivity of $do_{\mathbb{1}_n}$). Reasoning as in the foregoing section we obtain the slice.

Proposition 8.2. *Let $(v,X,Y) \in H_n^c$, then the n^2+n-dimensional manifold*

$$S = \{(v+a, X+C, Y+D) \mid v\overline{a}^{tr} + [X,\overline{C}^{tr}] + [Y,\overline{D}^{tr}] = 0\}$$

is a slice for the action of GL_n on H_n in (v,X,Y).

Now let $(v,X,Y) \in C_n^c$, that is, assume that $[X,Y] = 0$. We have seen that C_n^c is a submanifold of dimension n^2+2n. Thus, the tangentspace to C_n^c in $p = (v,X,Y)$ is of dimension n^2+2n. This is the subspace $(a,C,D) \in H_n$ such that $[X+\epsilon C, Y+\epsilon D] = 0$, that is

$$T_p\ C_n^c = \{(a,C,D) \mid [X,D]+[C,Y] = 0\}$$

Observe that as C_n^c is GL_n-stable the image of $do_{\mathbb{1}_n}$ is contained in the tangentspace (can be checked immediately using the Jacobi identity and $[X,Y] = 0$). Therefore, the orthogonal complement of $Im\ do_{\mathbb{1}_n}$ in the tangentspace is of dimension $n^2+2n-n^2 = 2n$.

Proposition 8.3. *Let $(v,X,Y) \in C_n^c$, then the $2n$-dimensional manifold S'*

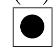
$$\{(v+a, X+C, Y+D) \mid v\overline{a}^{tr} + [X,\overline{C}^{tr}] + [Y,\overline{D}^{tr}] = 0 \text{ and } [X,D]+[C,Y] = 0\}$$

is a slice for the GL_n-action on C_n^c.

Because the dimension of the slice is independent of the point in C_n^c this concludes the proof of the following result.

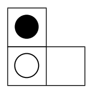
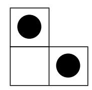
Theorem 8.4. *The Hilbert scheme $Hilb_n\ \mathbb{C}^2$ of n points in $\mathbb{C}^2$ is a manifold of dimension $2n$.*

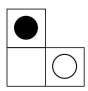
8b. The Hilbert game.

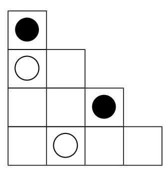
The determination of the slice is easier in a triple (v, X, Y) in standard σ- form for a Hilbert n-stair σ. In fact the description of $O_n(\sigma)$ is a slice for the action of GL_n on $H_n(\sigma)$. It is an interesting exercise to determine the covering of $Hilb_n\ \mathbb{C}^2$ by the $2n$-dimensional submanifolds $Hilb_n\ \mathbb{C}^2(\sigma) = Hilb_n\ \mathbb{C}^2 \cap O_n(\sigma)$ of $O_n(\sigma) = \mathbb{C}^{n^2+n}$. For example, consider $Hilb_2\ \mathbb{C}^2($  $)$. Because

$$[\begin{bmatrix} 0 & a \\ 1 & b \end{bmatrix}, \begin{bmatrix} c & d \\ e & f \end{bmatrix}] = \begin{bmatrix} ae - d & af - ac - bd \\ c + be - f & d - ae \end{bmatrix}$$

this subset can be identified with $\mathbb{C}^4$ using the equalities $d = ar$ and $f = c + be$. Similarly, one has identifications

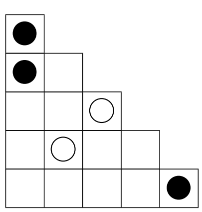
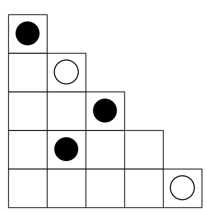
$Hilb_3\ \mathbb{C}^2($  $) = \mathbb{C}^6 = Hilb_3\ \mathbb{C}^2($  $)$

However, for $Hilb_3\ \mathbb{C}^2($  $)$ the description is more complicated. Observe that some of these intersections may be empty. For example, consider the Hilbert 5-stair

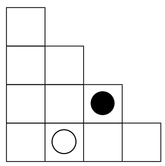
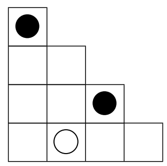
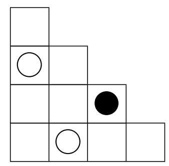
$\sigma_5 =$ 

then the associated series of words is $\{1, x, y, xy, yx\}$ whence $\sigma(v, X, Y) = 0$ whenever $[X, Y] = 0$. Hence all Hilbert stairs σ containing σ_5 (that is, if we recover σ_5 after removing certain rows and columns) satisfy $Hilb_n\ \mathbb{C}^2(\sigma) = \emptyset$.

Call a Hilbert n-stair σ a *forbidden position* if $Hilb_n\ \mathbb{C}^2(\sigma)$ is empty. A forbidden position of minimal size is called a *blocking position*. Examples of blocking positions are σ_5 above and the Hilbert 6-stairs

 and 

Determine all blocking positions for small n (up to color changes). Consider the following two person *game* on an n-stair. Left and right take turns where left places bLack stones and right white stones according to the Hilbert n-stair rules. The person unable to move or forced to move to a forbidden position is declared the looser. Determine the values of positions for small n following the rules of *combinatorial game theory* as explained in J.H. Conway's "On Numbers and Games". For example,

 $= \langle$  $\mid$  $\rangle = \langle 1 \mid 0 \rangle$

8c. Connectedness.

We have shown that $Hilb_n\ \mathbb{C}^2$ is a manifold of dimension $2n$. A priori it may have many connected components (all of dimension $2n$). We will now show that $Hilb_n\ \mathbb{C}^2$ is *connected*. As there is clearly a component of $Hilb_n\ \mathbb{C}^2$ of which points in general position correspond to n distinct points in $\mathbb{C}^2$, this result implies that any fat n-point of $\mathbb{C}^2$ can be deformed into n distinct simple points. In the next section we will show that a similar result does *not* hold for $Hilb_n\ \mathbb{C}^m$ with m, n sufficiently large.

Recall that the symmetric power $S^n\ \mathbb{C}^1$ parametrizes *sets* of n-points on the line $\mathbb{C}^1$ and can be identified with $\mathbb{C}^n$. Consider the map

$$Hilb_n\ \mathbb{C}^2 \xrightarrow{\pi} S^n\ \mathbb{C}^1$$

defined by mapping a cyclic triple $(v, X, Y) \in C_n^c$ with $[X, Y] = 0$ in the orbit corresponding to the point of $Hilb_n\ \mathbb{C}^2$ to the set $\{\lambda_1, \ldots, \lambda_n\}$ of eigenvalues of X. Observe that this map does not depend on the point chosen in the orbit. Let Δ be the *big diagonal* in $S^n\ \mathbb{C}^1$, that is, $S^n\ \mathbb{C}^1 - \Delta$ is the space of all sets of n distinct points from $\mathbb{C}^1$. Clearly, $S^n\ \mathbb{C}^1 - \Delta$ is a connected n-dimensional manifold. We claim that

$$\pi^{-1}(S^n\ \mathbb{C}^1 - \Delta) \simeq (S^n\ \mathbb{C}^1 - \Delta) \times \mathbb{C}^n$$

and hence is connected. Indeed, take a matrix X with n distinct eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. We may diagonalize X. But then, as

$$[\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \begin{bmatrix} y_{11} & \ldots & y_{1n} \\ \vdots & & \vdots \\ y_{n1} & \ldots & y_{nn} \end{bmatrix}] = \begin{bmatrix} (\lambda_1 - \lambda_1)y_{11} & \ldots & (\lambda_1 - \lambda_n)y_{1n} \\ \vdots & & \vdots \\ (\lambda_n - \lambda_1)y_{n1} & \ldots & (\lambda_n - \lambda_n)y_{nn} \end{bmatrix}$$

we see that also Y must be a diagonal matrix with entries $(\mu_1, \ldots, \mu_n) \in \mathbb{C}^n$ where $\mu_i = y_{ii}$. But then the cyclicity condition implies that all coordinates of v must be non-zero. Now, the stabilizer subgroup of the commuting (diagonal) matrix-pair (X,Y) is the *maximal torus* $T_n = \mathbb{C}^* \times \ldots \times \mathbb{C}^*$ of diagonal invertible $n \times n$ matrices. Using its action we may assume that all coordinates of v are equal to 1. That is, the points in $\pi^{-1}(\{\lambda_1,\ldots,\lambda_n\})$ with $\lambda_i \neq \lambda_j$ have unique representants of the form

$$(\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}, \begin{bmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n \end{bmatrix})$$

that is $\pi^{-1}(\{\lambda_1,\ldots,\lambda_n\}$ can be identified with $\mathbb{C}^n$, proving the claim. Next, we claim that *all* the fibers of π have dimension at most n.

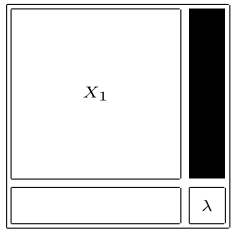
Let $\{\lambda_1,\ldots,\lambda_n\} \in S^n\ \mathbb{C}^1$ then there are only finitely many X in Jordan normalform with eigenvalues $\{\lambda_1,\ldots,\lambda_n\}$. Fix such an X, then the subset $T(X)$ of cyclic triples $(X,Y.v) \in C_n^c$ has dimension at most $n + dim\ C(X)$ where $C(X)$ is the centralizer of X in $M_n(\mathbb{C})$, taht is, $C(X) = \{Y \in M_n(\mathbb{C}) \mid XY = YX\}$. The stabilizer subgroup $Stab(X) = \{g \in GL_n \mid gXg^{-1} = X\}$ is an open subset of the vectorspace $C(X)$ and acts freely on the subset $T(X)$ because the action of GL_n on C_n^c has trivial stabilizers. But then, the orbitspace for the $Stab(X)$-action on $T(X)$ has dimension at most $n + dim\ C(X) - dim\ Stab(X) = n$. As we only have to consider finitely many X this proves the claim.

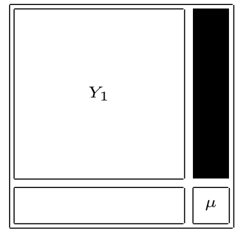
The diagonal Δ has dimension $n-1$ in $S^n\ \mathbb{C}^1$ and hence by the foregoing we know that the dimension of $\pi^{-1}(\Delta)$ is at most $2n-1$. Let H be the connected component of $Hilb_n\ \mathbb{C}^2$ containing teh connected subset $\pi^{-1}(S^n\ \mathbb{C}^1 - \Delta)$. If $\pi^{-1}(\Delta)$ were not entirely contained in H, then $Hilb_n\ \mathbb{C}^2$ would have a component of dimension less than $2n$, which we proved not to be the case. Thus $H = Hilb_n\ \mathbb{C}^2$ and we have proved :

Theorem 8.5. *The Hilbert scheme $Hilb_n\ \mathbb{C}^2$ of n points in $\mathbb{C}^2$ is a connected manifold of dimension $2n$.*

Let (v,X,Y) be a cyclic triple representing a point in $h \in Hilb_n\ \mathbb{C}^2$, that is $[X,Y] = 0$. We claim that X and Y are simultaneously upper triangularizable. Let λ be an eigenvalue of X and consider the eigenspace $V_\lambda = \{w \in \mathbb{C}^n \mid Xw = \lambda w\}$. Then, $YV_\lambda \subset V_\lambda$ as $XYw = Y(Xw) = Y(\lambda w) = \lambda Yw$. Let Y_λ be the matrix representing the action of Y on V_λ then up to basechange in V_λ we may assume that Y_λ is in Jordan normal form, but then X and Y have at least one common eigenvector $w \in V_\lambda$ such that $Yw = \mu w$. Consider a new basis $\{f_1,\ldots,f_n\}$ with $f_n = w$, then in this basis the matrixpair (X,Y)

has blockform

$X =$ 

$Y =$ 

But then (X_1, Y_1) is a commuting matrixpair in $M_{n-1}(\mathbb{C})$ and by induction we may assume that they are simultaneously upper triangularizable, proving the claim. Hence, we have a cyclic triple in the orbit of (v, X, Y) of the form

$$\left(\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}, \begin{bmatrix} \lambda_1 & \dots & * \\ & \ddots & \vdots \\ & & \lambda_n \end{bmatrix}, \begin{bmatrix} \mu_1 & \dots & * \\ & \ddots & \vdots \\ & & \mu_n \end{bmatrix} \right)$$

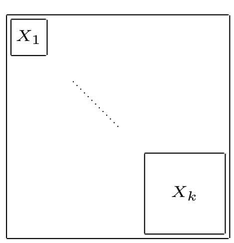
The *Hilbert-Chow map* $Hilb_n\,\mathbb{C}^2 \xrightarrow{H} S^n\,\mathbb{C}^2$ is defined by sending the point h representing the orbit of (v, X, Y) to the set $\{(\lambda_1, \mu_1), \dots, (\lambda_n, \nu_n)\}$ of n points in $\mathbb{C}^2$. If $(\lambda_i, \mu_i) \neq (\lambda_j, \mu_j)$ for $i < j$ and consider the 2×2 minors of X and Y

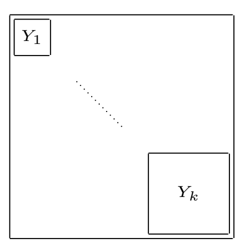
$$X_{ij} = \begin{bmatrix} \lambda_i & x_{ij} \\ 0 & \lambda_j \end{bmatrix} \qquad Y_{ij} = \begin{bmatrix} \mu_i & y_{ij} \\ 0 & \mu_j \end{bmatrix}$$

Assume that $\lambda_i \neq \lambda_j$, conjugating with the matrix $g = \begin{bmatrix} 1 & \frac{x_{ij}}{\lambda_i - \lambda_j} \\ 0 & 1 \end{bmatrix}$ gives

$$\begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_j \end{bmatrix} \qquad \begin{bmatrix} \mu_i & y'_{ij} \\ 0 & \mu_j \end{bmatrix}$$

but then the commutation relation forces $y'_{ij} = 0$. Repeating this argument and possibly permuting the base vectors we may assume that the commuting matrix pair (X, Y) can be brought into block form

$X =$ 

and $Y =$ 

Here, X_i (resp. Y_i) is an upper triangular $m_i \times m_i$ matrix with single eigenvalue λ_i (resp. μ_i) where p_i is the multiplicity with which $p_i = (\lambda_i, \mu_i)$ appears in the n-set of points. The image under the Hilbert-Chow map is then the formal sum

$$H(v, X, Y) = m_1[p_1] + \dots m_k[p_k] \in S^n\,\mathbb{C}^2.$$

Another way to phrase the above block decomposition is in terms of the n-dimensional algebra $\mathbb{C}[X,Y] = \frac{\mathbb{C}[x,y]}{I}$ determined by $h \in Hilb_n\ \mathbb{C}^2$. As $\mathbb{C}[X,Y]$ is the $\mathbb{C}$-subalgebra of $M_n(\mathbb{C})$ generated by the commuting matrix-pair (X,Y) we have a decomposition

$$\mathbb{C}[X,Y] = \mathbb{C}[X_1,Y_1] \oplus \ldots \oplus \mathbb{C}[X_k, Y-k] = \frac{\mathbb{C}[x,y]}{I_1} \oplus \ldots \oplus \frac{\mathbb{C}[x,y]}{I_k}$$

where I_j is an ideal of $\mathbb{C}[x,y]$ of codimension m_j concentrated in the point $p_j = (\lambda_j, \mu_j)$. This means that p_j is the only point of $\mathbb{C}^2$ where every polynomial $f \in I_j$ vanishes. Let us draw some consequences from this decomposition. If all multiplicities m_j are one, that is if $H(v,X,Y)$ does not lie on the diagonal Δ in $S^n\ \mathbb{C}^2$, then X and Y are simultaneously *diagonalizable*. But then, as the stabilizer subgroup of the commuting matrix-pair (X,Y) is equal to the *maximal torus* $T_n = \mathbb{C}^* \times \ldots \times \mathbb{C}^*$ of diagonal elements in GL_n we see that there is precisely one orbit in C_n^c lying over this point represented by the cyclic triple

$$(\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{bmatrix})$$

Hence, over the open subset $S^n\ \mathbb{C}^2 - \Delta$ the Hilbert-Chow map H is a one-to-one correspondence. As we have seen that $Hilb_n\ \mathbb{C}^2$ is a manifold, the Hilbert-Chow map is a resolution of singularities of $S^n\ \mathbb{C}^2$.

When $\delta = m_1[p_1] + \ldots + m_k[p_k] \in \Delta$, then any point in the fiber of the Hilbert-Chow map is determined by a k-tuple $(h_1, \ldots, h_k)$ where $h_i \in Hilb_{m_i}\ \mathbb{C}^2$ and *concentrated* in p_i, that is, the image of h_i under the Hilbert-Chow map $Hilb_{m_i}\ \mathbb{C}^2 \longrightarrow S^{m_i}\ \mathbb{C}^2$ is $m_i[p_i]$. The parallel translations in $\mathbb{C}^2$ give a natural one-to-one correspondence between $H^{-1}(m_i[(0,0)])$ and $H^{-1}(m_i[p_i])$. Hence, it is important to study the subset $H^{-1}(n[(0,0)])$ of $Hilb_n\ \mathbb{C}^2$. Its points correspond to codimension n ideals I of $\mathbb{C}[x,y]$ concentrated in $(0,0)$. The n-dimensional algebra $\mathbb{C}[x,y]/I$ has a unique underlying point $(0,0)$, that is, it is a *local* algebra of dimension n. Hence, the algebras $\mathbb{C}[x,y]/I$ are examples of *fat points* of multiplicity n, that is, commutative local algebras of dimension n.

Week 9

Reductivity of GL_n.

In this section we will put the examples considered before in a general framework. To begin, in each of the examples we have a *linear action* of the basechange group GL_n on a finite dimensional $\mathbb{C}$-vectorspace V, that is $g(v+w) = g.v + g,w$, $\mathbb{1}_n.v = v$ and $(gg').v = g.(g'.v)$ for all $g, g' \in GL_n$ and $v, w \in V$

Jordan forms	$V = M_n(\mathbb{C})$ $g.m = gmg^{-1}$
Dynamical systems	$V = M_{n\times m}(\mathbb{C}) \oplus M_n(\mathbb{C}) \oplus M_{p\times n}(\mathbb{C})$ $g.(A,B,C) = (gA, gBg^{-1}, Cg^{-1})$
Hilbert schemes	$V = \mathbb{C}^n \oplus M_n(\mathbb{C}) \oplus M_n(\mathbb{C})$ $g.(v,X,Y) = (gv, gXg^{-1}, gYg^{-1})$

Our first objective will be to control all linear actions possible. We will call actions on V and W *isomorphic* if there is a linear isomorphism $V \xrightarrow{\phi} W$ which is GL_n-equivariant, that is $g.\phi(v) = \phi(g.v)$ for all $g \in GL_n$ and $v \in V$.

Problem 7.

Describe all linear actions of GL_n on finite dimensional vectorspaces V up to isomorphism.

Let G be a group and V a finite dimensional $\mathbb{C}$-vectorspace on which G acts linearly. We say that V is a *G-representation*. If V and W are G-representations, then so are $V \oplus W$ and $V \otimes W$ where the actions are given by

$$g.(v,w) = (g.v, g.w) \quad \text{and} \quad g.(v \otimes w) = g.v \otimes g.w.$$

for all $g \in G$ and all $v \in V, w \in W$. A G-subrepresentation of V is a subspace W which is left stable under the action of G. A G-representation V is said to be *simple* if V contains no proper G-subrepresentations and is said to be *completely reducible* if V is the direct sum of simple G-subrepresentations.

9a. Haar measures.

Our first aim is to prove that GL_n is a *reductive group*, that is, all GL_n-representations are completely reducible. The method of proof is based on the 'averaging over the group'-idea used to prove that *finite* groups are reductive. We will briefly sketch it. Let G be a finite group and W a G-subrepresentation of a G-representation V. Let $\phi : V \twoheadrightarrow W$ be a linear projection obtained from extending a basis of W to one of V. Then, we consider the averaged linear map

$$\pi : V \twoheadrightarrow W \quad \text{where} \quad v \mapsto \sum_{g \in G} g(\phi(g^{-1}.v)).$$

This map is G-equivariant and restricted to W it is multiplication by $\#(G)$. The kernel K of π is a subspace of V, stable under the action of G and complementary to W. That is, $V = W \oplus K$ is a decomposition of G-representations. Continuing gives a complete decomposition of V.

We will replace the sum by an integral and the finite group by the *compact subgroup* of *unitary matrices*

$$U_n = \{A \in GL_n \mid A.\overline{A}^{tr} = I_n\}$$

Clearly, U_n is a subgroup of GL_n and we claim that it is a compact *Lie group*, that is a real compact differentiable manifold (a C^∞-manifold) with a differentiable groupstructure. Because U_n is a group it suffices to verify the manifold property in a neighborhood of the unit element $e = \mathbb{1}_n$. Let $Herm_n$ be the $\mathbb{R}$-vectorspace of Hermitian $n \times n$ matrices $Herm_n = \{H \in M_n(\mathbb{C}) \mid H = \overline{H}^{tr}\}$ and consider the map

$$f : GL_n \longrightarrow Herm_n \quad \text{defined by} \quad A \longrightarrow A\overline{A}^{tr}$$

Calculating the differential with the ϵ-method gives that $Im\ df_{\mathbb{1}_n}(X) = X + \overline{X}^{tr}$ and as any Hermitian matrix H can be written as $\frac{1}{2}(H + \overline{H}^{tr})$ this differential is surjective. By the implicit function theorem (over $\mathbb{R}$) we deduce that the fiber $f^{-1}(\mathbb{1}_n) = U_n$ is a real manifold of $\mathbb{R}$-dimension $2n^2 - dim\ Herm_n = n^2$. Finally, U_n is *compact* as it is closed in the $\mathbb{R}$-topology on M_n and *bounded* as the norms of all entries in a fixed row (or column) add up to one.

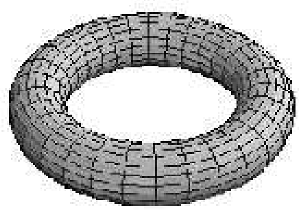
Compact Lie groups have a *Haar measure* allowing to integrate complex valued continuous functions in an invariant way. For example $U_1 = \{c \in C \mid c\bar{c} = 1\}$ is the unit circle $S^1 = \{e^{ix} \mid 0 \leq x < 2\pi\}$. We define a complex valued linear map $\int_{U_1} .(g)\ dg$ on the space of all continuous functions $f : U_1 \longrightarrow \mathbb{C}$ by

$$\int_{U_1} f(g)\ dg = \frac{1}{2\pi} \int_0^{2\pi} f(e^{ix})\ dx.$$

We see that this map is normalized, that is $\int_{U_1}\ dg = 1$ and is left and right invariant, that is for any $h = e^{iy} \in U_1$

$$\int_{U_1} f(gh)\ dg = \frac{1}{2\pi} \int_0^{2\pi} f(e^{x+y})\ dx = \frac{1}{2\pi} \int_0^{2\pi} f(e^{x'})\ dx' = \int_{U_1} f(g)\ dg$$

and similarly for multiplication by h on the left. For the compact group $G = U_1 \times U_1$, that is the *torus group* $S^1 \times S^1$



we can take as the normalized invariant integral $\int_{U_1 \times U_1} f(g)\ dg = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} f(e^{ix}, e^{iy})\ dx\ dy$. In general, a Haar measure on a compact Lie group is a linear functional $\int_G f(g)\ dg$ which is normalized $\int_G\ dg = 1$ and is left and right invariant $\int_G f(gh)\ dg = \int_G f(g)\ dg = \int_G f(hg)\ dg$. Assuming its existence we can prove :

Proposition 9.1. *If G is a compact Lie group, then every G-representation is completely reducible.*

Let W be a subspace of a finite dimensional G-representation V which is invariant under the G-action. Extending a basis of W to one of V gives a linear projection $V \overset{\pi}{\twoheadrightarrow} W$ which is the identity on W. For $v \in V$ we have complex valued coordinate maps

$$G \longrightarrow W = \mathbb{C}^k \quad \text{defined by} \quad g \mapsto g.(\pi(g^{-1}.v)$$

Integrating these coordinate maps defines a map $\phi : V \longrightarrow W$

$$\phi(v) = \int_G g.\pi(g^{-1}.v)\ dg$$

which is linear and the identity on W. Moreover, ϕ commutes with the G-action as for every $h \in G$

$$\begin{aligned} \phi(h.v) = & \int_G g.\pi(g^{-1}.h.v)\,dg = h.\int_G h^{-1}g.\pi(g^{-1}h.v)\,dg \\ \overset{*}{=} & \qquad h.\int_G g.\pi(g^{-1}.v)\,dg = h.\phi(v) \end{aligned}$$

where the starred equality uses invariance of the Haar measure on G. But then, $V = W \oplus Ker\ \phi$ is a decomposition as G-representations. Continuing whenever one of the components has a non-trivial G-subrepresentation, we arrive after a finite number of steps at a decomposition of V into simple G-representations proving proposition 9.1. .

9b. Cartan decomposition.

Next, we want to move from the (real) compact Lie group to the associated (complex) algebraic group. In particular, from the (real) torus $T_n(\mathbb{R}) = U_1 \times \ldots \times U_1$ to the (complex) torus $T_n = \mathbb{C}^* \times \ldots \times \mathbb{C}^*$ and from the unitary group U_n to the basechange group GL_n. Let us first consider the torus case. A polynomial $f(t) \in \mathbb{C}[t]$ has only finitely many zeroes so if $f(U_1) = 0$, f must be the zero polynomial whence $f(\mathbb{C}^*) = 0$. A similar result holds for n-dimensional tori. Let $f(x_1,\ldots,x_n)$ be a polynomial in $\mathbb{C}[x_1,\ldots,x_n]$ and write it as a polynomial in x_n with coefficients in $\mathbb{C}[x_1,\ldots,x_{n-1}]$

$$f(x_1,\ldots,x_n) = f_0(x_1,\ldots,x_{n-1})x_n^k + \ldots + f_k(x_1,\ldots,x_{n-1})x_n + f_{k+1}(x_1,\ldots,x_{n-1})$$

Assume that $f(T_n(\mathbb{R})) = 0$ and fix $(u_1,\ldots,u_{n-1}) \in T_{n-1}(\mathbb{R})$, then $f(u_1,\ldots,u_{n-1},x_n)$ is a polynomial vanishing on U_1, whence all the coefficients $f_i(u_1,\ldots,u_{n-1})$ must be zero. Hence, the f_i are polynomials such that $f_i(T_{n-1}(\mathbb{R})) = 0$ and by induction on n we may assume that then $f_i(T_{n-1}) = 0$, whence $f(T_n) = 0$.

Assume now that V is a T_n-representation. Assume that V has a decomposition $V = W \oplus W'$ as a $T_n(\mathbb{R})$-representation. Consider the *normalizer subgroup*

$$N = N_{T_n}(W) = \{c = (c_1,\ldots,c_n) \in T_n \mid c.W \subset W\}$$

Extending a basis $\{w_1,\ldots,w_l\}$ of W to V we see that this condition can be expressed by the fact that certain certain coordinates of $c.w_i$ (which are polynomial functions in the c_j) must be zero. Hence, N can be identified as the subset of points of T_n which are simultaneous zeroes of a set of polynomials $\{f_a\}$. Because W is a $T_n(\mathbb{R})$-representation we have $T_n(\mathbb{R}) \subset N$ and hence $f_a(T_n(\mathbb{R})) = 0$ for all f_a, whence $f_a(T_n) = 0$. That is, $N = T_n$ and so W is also a T_n-representation. That is, a decomposition of V in simple

$T_n(\mathbb{R})$-representations is also a decomposition of simple T_n-representations. Hence, the complex torus $T_n = \mathbb{C}^* \times \ldots \times \mathbb{C}^*$ is a reductive group.

To prove a similar result for GL_n we need a polynomial density result for U_n with respect to GL_n. This follows from the *Cartan decomposition* asserting that

$$GL_n = U_n \, T_n \, U_n$$

where T_n is the *maximal torus* of diagonal matrices in GL_n. Indeed, for $g \in GL_n$, $g\overline{g}^{tr}$ is an Hermitian matrix and hence diagonalizable by unitary matrices (this follows from the Gramm-Schmidt procedure of orthonormal bases in linear algebra). So, there is a $u \in U_n$

$$u^{-1} g \overline{g}^{tr} u = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} = \underbrace{u^{-1} g u}_{p} \,.\, \underbrace{u^{-1} \overline{g}^{tr} u}_{\overline{p}^{tr}}$$

Because $\alpha_i = \sum_{j=1}^n \parallel p_{ij} \parallel^2$, $\alpha_i > 0 \in \mathbb{R}$. Let $\beta_i = \sqrt{\alpha_i}$ and consider the diagonal matrix

$$d = \begin{bmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_n \end{bmatrix}$$

Clearly, $g = ud(d^{-1}u^{-1}g)$ and $v = d^{-1}u^{-1}g$ is a unitary matrix as

$$\begin{aligned} v\overline{v}^{tr} =& (d^{-1}u^{-1}g).(\overline{g}^{tr} u d^{-1}) = d^{-1}(u^{-1} g \overline{g}^{tr} u) d^{-1} \\ =& d^{-1} d^2 d^{-1} = \mathbb{1}_n \end{aligned}$$

proving the Cartan decomposition. Let $f = f(x_{11}, x_{12}, \ldots, x_{nn})$ be a polynomial function in the matrix entries of $M_n(\mathbb{C})$ such that $f(U_n) = 0$. Then $f' = f \mid T_n$ is a polynomial such that $f'(T_n(\mathbb{R})) = 0$ whence $f(T_n) = 0$. By continuity of the matrix-multiplication in GL_n and the Cartan decomposition, it follows that $f(GL_n) = 0$. Using this polynomial denseness of U_n in GL_n we can repeat the above T_n argument verbatim and conclude :

Theorem 9.2. *The n-dimensional torus T_n and the basechange group GL_n are reductive groups. That is, every representation V admits a decomposition as a direct sum of simple subrepresentations S_i*

$$V = S_1 \oplus \ldots \oplus S_k.$$

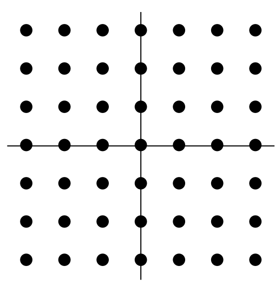
Moreover, such a decomposition is unique up to permutation of the factors.

The last statement follows because any G-equivariant map $S \xrightarrow{\phi} S'$ between two simple G-representations is either the zero map or an isomorphism (the kernel $Ker\ \phi$ is a G-subrepresentation of S).

Week 10

GL_n-representations.

Reductivity of GL_n reduces our problem to that of classifying all *simple* GL_n-representations V. Again, we will first consider this problem for the n-dimensional complex torus T_n. We claim that the simple T_n-representations are all one dimensional and are classified by the lattice $\mathbb{Z}^n$. For example, for T_2 we get the discrete set



10a. Characters.

Let V be a simple T_n-representation of dimension m and $h = (r_1, \ldots, r_n) \in T_n(\mathbb{R})$ with all r_i roots of unity. The subgroup $< h >$ of $T_n(\mathbb{R})$ is a finite Abelian group so as an $< h >$- representation $V = \oplus V_{\lambda_i}$ where V_{λ_i} is the eigenspace $\{v \in V \mid h.v = \lambda_i v\}$. T_n being Abelian, each of the V_{λ_i} is a T_n-subrepresentation. By simplicity, there is only one non-zero eigenspace, that is $V = V_\lambda$. Varying h we see that the subgroup $\mu \times \ldots \times \mu$ of $T_n(\mathbb{R})$ consisting of elements having all its entries roots of unities acts diagonally on V. But then so does its closure, which is $T_n(\mathbb{R})$ so V is a direct sum of one-dimensional $T_n(\mathbb{R})$-subrepresentations, a contradiction unless $m = 1$. Hence, all simple T_n-representations are one dimensional. For V a one-dimensional simple T_n-representation, the action determines (and is determined by) a groupmorphism $T_n \xrightarrow{\chi} \mathbb{C}^*$. The only groupmorphisms $\mathbb{C}^* \longrightarrow \mathbb{C}^*$ are easily seen to be the maps $x \mapsto x^\nu$ for some $\nu \in \mathbb{Z}$. Hence V

is determined by the *character*

$$\chi : T_n \longrightarrow \mathbb{C}^* \quad \text{given by} \quad (x_1, \ldots, x_n) \mapsto x_1^{a_1} \ldots x_n^{a_n}$$

for some $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$, proving the claim. Traditionally, one writes the *character group* $X(T_n) = \mathbb{Z}^n$ additively and denotes the character corresponding to the standard basis vector e_i by $\epsilon_i : T_n \longrightarrow \mathbb{C}^*$ defined by $(x_1, \ldots, x_n) \mapsto x_i$.

Proposition 10.1. *Every finite dimensional T_n-representation V is the direct sum of its eigenspaces*

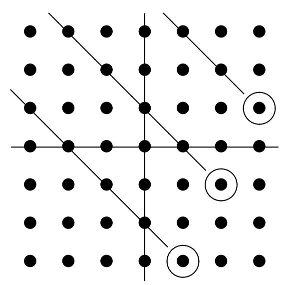
$$V_\chi = \{v \in V \mid x.v = \chi(x)v \ \textit{for all} \ x = (x_1, \ldots, x_n) \in T_n \ \}$$

where $\chi = \sum_i a_i \epsilon_i$ is a character in $X(T_n) = \mathbb{Z}^n$.

On $X(T_n) = \mathbb{Z}^n$ we put an ordering defined as follows. Let $\lambda = \sum a_i \epsilon_i$ and $\mu = \sum b_i \epsilon_i$, then

$$\lambda \leq \mu \Leftrightarrow \sum_{i=1}^n a_i = \sum_{i=1}^n b_i \quad \text{and} \sum_{i=1}^k a_i \leq \sum_{i=1}^k b_i \text{ for all } 1 \leq k \leq n$$

For example, for T_2 the characters smaller than the circled characters all lie on the indicated halflines



Let V be a finite dimensional GL_n-representation. As a T_n-representation (using the diagonal embedding $T_n \hookrightarrow GL_n$) V decomposes as a direct sum of *weight spaces*

$$V = \oplus_{\lambda \in X(T_n)} V_\lambda$$

We will now investigate the restriction imposed on the V_λ by the GL_n-action. To begin, consider the *permutation matrixes* P_σ for $\sigma \in S_n$, having a 1 at all positions $(i, \sigma(i))$ and zeroes elsewhere. There is a natural action of the symmetric group S_n on T_n and on $X(T_n)$ defined by $\sigma.x = \sigma.(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ and $\sigma.\lambda = \sigma.(\sum a_i \epsilon_i) = \sum a_i \epsilon_{\sigma(i)}$. If $v \in V_\lambda$ and $x = diag(x_1, \ldots, x_n) \in T_n \hookrightarrow GL_n$, we have

$$\begin{aligned} x.(P_\sigma.v) = & \quad P_\sigma.(P_\sigma^{-1} x P_\sigma).v = P_\sigma.(\sigma.x).v \\ = & \quad P_\sigma.(\sigma\lambda(x)v) = \sigma\lambda(x) P_\sigma.v \end{aligned}$$

that is, $P_\sigma.v \in V_{\sigma.\lambda}$. So we have that $dim\ V_\lambda = dim\ V_{\sigma\lambda}$ for all $\lambda \in X(T_n)$ and all $\sigma \in S_n$. Further, consider the *elementary matrices*

$$x_{ij}(t) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & t & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

having 1's on the diagonal, $t \in \mathbb{C}$ at place (i,j) and zeroes elsewhere. One calculates that for every $x = diag(x_1, \ldots, x_n) \in T_n \hookrightarrow GL_n$ we have

$$x x_{ij}(t) x^{-1} = x_{ij}((x_i - x_j)t) = x_{ij}((\epsilon_i - \epsilon_j)(x).t)$$

If $v \in V_\lambda$, then as the map $t \mapsto x_{ij}(t).v$ is a polynomial map in t we can write with respect to a basis $\{w_1, \ldots, w_l\}$ of V that

$$x_{ij}(t).v = \sum_{a=1}^{l} (\sum_{b \geq 0} c_{ab} t^b) w_a = \sum_{b \geq 0} t^b v_b \quad \text{with} \quad v_b = \sum_{a=1}^{l} c_{ab} w_a$$

As $x_{ij}(0) = \mathbb{1}_n$ we have $v_0 = v$ and we claim that $v_b \in V_{\lambda + b(\epsilon_i - \epsilon_j)}$. Indeed, for all $x = diag(x_1, \ldots, x_n) \in T_n \hookrightarrow GL_n$ and $t \in \mathbb{C}$ we have

$$\begin{aligned} x.x_{ij}(t).v = & \quad (x x_{ij}(t) x^{-1}).x.v = x_{ij}((\epsilon_i - \epsilon_j)(x).t).(\lambda(x)v) \\ = & \quad \lambda(x) \textstyle\sum_{b \geq 0} ((\epsilon_i - \epsilon_j)(x).t)^b v_b = \lambda(x) \sum_{b \geq 0} t^b ((\epsilon_i - \epsilon_j)(x))^b v_b \\ = & \quad \textstyle\sum_{b \geq 0} t^b (\lambda + b(\epsilon_i - \epsilon_j))(x) v_b \end{aligned}$$

and on the other hand we have the equalities

$$x.x_{ij}(t).v = x.(\sum_{b \geq 0} t^b v_b) = \sum_{b \geq 0} t^b (x.v_b)$$

So, comparing terms we establish our claim as

$$x.v_b = (\lambda + b(\epsilon_i - \epsilon_j))(x) v_b \quad \text{that is,} \quad v_b \in V_{\lambda + b(\epsilon_i - \epsilon_j)}.$$

10b. Borel subgroups.

Let $\blacktriangle_n$ be the subgroup of GL_n consisting of upper triangular matrices with 1's on the diagonal and let $\blacktriangledown_n$ be the corresponding subgroup of lower triangular matrices in GL_n. Consider the product map

$$\blacktriangledown_n \times T_n \times \blacktriangle_n \xrightarrow{prod} GL_n \qquad (l, d, u) \mapsto ldu$$

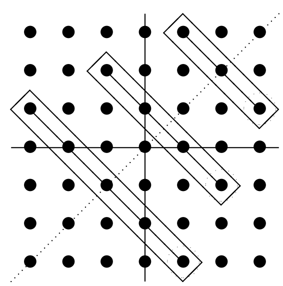
The differential of $prod$ in the unit element $(\mathbb{1}_n, \mathbb{1}_n, \mathbb{1}_n)$ can be computed by the ϵ-method expanding

$$(\mathbb{1}_n + \epsilon L)(\mathbb{1}_n + \epsilon D)(\mathbb{1}_n + \epsilon U) = \mathbb{1}_n + \epsilon(L + D + U)$$

with L (resp. U) a strictly lower (resp. upper) triangular matrix and D a diagonal matrix. Hence, $d\ prod_{(\mathbb{1}_n, \mathbb{1}_n, \mathbb{1}_n)}$ is an isomorphism and so the product map has a dense image in GL_n. Therefore,

$$V = V_{l_1} \oplus \ldots \oplus V_{l_s} \qquad V_l = \oplus\{V_\lambda \mid \lambda = \sum a_i \epsilon_i \text{ with } \sum_i a_i = l\}$$

as $\blacktriangle_n$ and $\blacktriangledown_n$ are generated by the respective $x_{ij}(t)$ and by the above equality we have that for any $v \in V_l$ that $x_{ij}(t)v \in V_l$. Hence, the non-zero *weight spaces* V_λ of a GL_n-representation V lie in a finite number of hyperplanes where they form a configuration stable under the action of S_n. For T_2 the configuration of non-zero weight spaces is symmetric with respect to the main diagonal and constrained to the indicated boxes



The dashed circled weights are the *highest weights* with respect to the ordering on $X(T_2)$. For general n we similarly have for each component V_l a *highest weight* V_λ with respect to the ordering on $X(T_n)$. An element $v \in V_\lambda$ is then called a *highest weight vector* of V (or of the component V_l). Let v_b be the vectors in $V_{\lambda + b(\epsilon_i - \epsilon_j)}$ introduced before, then again using the dense map $\blacktriangledown_n \times T_n \times \blacktriangle_n \longrightarrow GL_n$ we see that the subspace of V spanned by $v = v_0$ and the v_b is a simple GL_n-subrepresentation of V. We now have all the necessary ingredients to finish the proof of :

Theorem 10.2. *There is a one-to-one correspondence between*

1. *isomorphism classes of simple GL_n-representations, and*
2. $\lambda = \sum_i a_i \epsilon_i \in X(T_n)$ *which are highest weights, that is,*

$$a_1 \geq a_2 \geq \ldots \geq a_n.$$

We have already seen that a highest weight vector generates a simple component. Observe that a highest weight vector spans a one-dimensional subspace which is stable under the action of the *Borel subgroup* B_n of upper triangular elements of GL_n. Further, if V is a simple GL_n-representation then using the S_n-action we know that V has a highest weight vector. Remains to prove that two GL_n-representations spanned by a highest weight vector of the same weight are isomorphic and that every $\lambda = \sum_i a_i \epsilon_i$ with $a_1 \geq \ldots \geq a_n$ occurs as a highest weight for a GL_n-representation. First, let V and V' be simple GL_n-representations with highest weight vectors v and v' of weight λ. Then, $W = V \oplus V'$ is a GL_n-representation and $w = (v, v')$ a vector of highest weight λ so generates a simple component W' of W with $W'_\lambda = \mathbb{C}w$. $V \cap W'$ is a GL_n-subrepresentation of V so is either 0 or V. As $W'_\lambda = \mathbb{C}w \not\subset V_\lambda = \mathbb{C}(v, 0)$ this intersection must be zero. Then, the composition

$$W' \hookrightarrow V \oplus V' \xrightarrow{pr_1} V$$

is into and gives therefore an isomorphism $W \simeq V$. Repeating the argument with V' instead of V we deduce that $V \simeq V'$.

As for the existence, let $\lambda = \sum_i a_i \epsilon_i$ with $a_1 \geq \ldots \geq a_n$ and consider $\omega_i = \epsilon_1 + \ldots + \epsilon_i$, then we have

$$\lambda = b_1 \omega_1 + \ldots b_n \omega_n$$

with $b_n = a_n \in \mathbb{Z}$ and all $b_i = a_i - a_{i+1} \in \mathbb{N}$ for $1 \leq i \leq n-1$. Recall the construction of *exterior product* $\wedge^i \mathbb{C}^n$ for $1 \leq i \leq n$ that is, the subspace of $\mathbb{C}^n \otimes \ldots \otimes \mathbb{C}^n$ (i terms) spanned by the anti-symmetric tensors which is clearly a GL_n- subrepresentation of $\mathbb{C}^n \otimes \ldots \otimes \mathbb{C}^n$ with the diagonal action on the tensorproduct, $g.(c_1 \otimes \ldots \otimes c_i) = (g.c_1) \otimes \ldots \otimes (g.c_i)$. Consider the GL_n-representation

$$V = (\wedge^1 \mathbb{C}^n)^{\otimes b_1} \otimes (\wedge^2 \mathbb{C}^2)^{\otimes b_2} \otimes \ldots \otimes (\wedge^n \mathbb{C}^n)^{\otimes b_n}$$

where $(\wedge^i \mathbb{C}^n)^{\otimes k}$ is the k-fold tensorproduct of $\wedge^i \mathbb{C}^n$ for $k \in \mathbb{N}_+$, $(\wedge^i \mathbb{C}^n)^0 = \mathbb{C}_{triv}$ the trivial one-dimensional GLn-representation and $(\wedge^n \mathbb{C}^n)^{\otimes m}$ is the one-dimensional GL_n-representation defined by $g \mapsto (det\ g)^m$ (note that $m < 0$ is possible). If $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{C}^n$, then $e_1 \wedge e_2 \wedge \ldots \wedge e_i$ is stable under the action of the Borel subgroup B_n and has weight ω_i as clearly

$$diag(x_1, \ldots, x_n).(e_1 \wedge e_2 \wedge \ldots \wedge e_i) = \prod_{j=1}^{i} x_j (e_1 \wedge e_2 \wedge \ldots \wedge e_i)$$

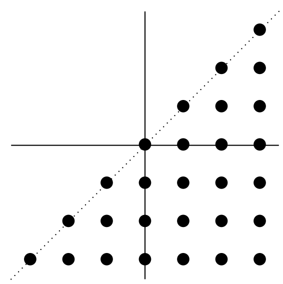
But then, the following vector v is stable under the Borel B_n and has weight λ

$$v = \begin{cases} e_1^{\otimes b_1} \otimes (e_1 \wedge e_2)^{\otimes b_2} \otimes \ldots \otimes (e_1 \wedge \ldots \wedge e_n)^{\otimes b_n} & \text{when } b_n \in \mathbb{N}, \text{ and} \\ e_1^{\otimes b_1} \otimes (e_1 \wedge e_2)^{\otimes b_2} \otimes \ldots \otimes (f_1 \wedge \ldots \wedge f_n)^{\otimes -b_n} & \text{when } b_n < 0 \end{cases}$$

where $\{f_1, \ldots, f_n\}$ is the dual basis, that is the standard basis of $\mathbb{C}^{n*}$. Hence, v generates a simple GL_n-representation of highest weight λ. This concludes the proof of theorem 10.2. . That is, we can identify the set Ω_{GL_n} of isomorphism classes of simple GL_n-representations with the subset of $X(T_n)$

$$\Omega_{GL_n} = \mathbb{N}\omega_1 \oplus \mathbb{N}\omega_2 \oplus \ldots \oplus \mathbb{N}\omega_{n-1} \oplus \mathbb{Z}\omega_n$$

For example, the simple GL_2-representations are in one-to-one correspondence to the subset of $\mathbb{Z}^2 = X(T_2)$



We will denote the simple GL_n-representation of highest weight λ by S_λ. For example,

$$\mathbb{C}^n = S_{(1,0,\ldots,0)} \quad \text{and} \quad \mathbb{C}^{n*} = S_{(0,\ldots,0,-1)}$$

Consider the action of GL_n on $M_n(\mathbb{C})$ by conjugation. If e_{ij} is the matrix having a 1 at place (i, j) and zeroes elsewhere, then we have

$$M_n(\mathbb{C}) = M_n(\mathbb{C})_0 \oplus \oplus_{i \neq j} M_n(\mathbb{C})_{\epsilon_i - \epsilon_j} \quad \text{with } M_n(\mathbb{C})_{\epsilon_i - \epsilon_j} = \mathbb{C}e_{ij}$$

and $M_n(\mathbb{C})_0$ is the space of diagonal matrices. The weightspaces left invariant by the B_n-action are $\mathbb{C}e_{1n}$ and $\mathbb{C}\mathbb{1}_n$. The first generates the simple representation with highest weight $(1, 0, \ldots, 0, -1)$, the second is the trivial representation (highest weight $(0, \ldots, 0)$. That is, as GL_n-representations

$$M_n(\mathbb{C}) = M_n(\mathbb{C})_{(0,\ldots,0)} \oplus M_n(\mathbb{C})_{(1,0,\ldots,0,-1)} = \mathbb{C}\mathbb{1}_n \oplus M_n^0(\mathbb{C})$$

where $M_n^0(\mathbb{C})$ is the space of all *trace zero* matrices. The weights $\epsilon_i - \epsilon_j$ are called the *roots* of GL_n. The roots $\epsilon_i - \epsilon_j$ with $i < j$ are called *positive*. This allows a reformulation of the ordering in $X(T_n)$: for $\lambda, \mu \in X(T_n)$ we have $\lambda \leq \mu$ if and only if $\mu - \lambda$ can be written as a sum of positive roots.

10c. GL_n-varieties.

Combining reductivity of GL_n with the combinatorial description of the isomorphism classes of simple GL_n-representations allows us to determine all linear actions of GL_n on a finite dimensional vectorspace V. To formalize the notion of GL_n*-variety* we consider the induced action of GL_n on the polynomial functions f on V. Considering the diagram

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "g."'] & \mathbb{C} \\ V \arrow[ur, "g.f"'] & \end{tikzcd}$$

we see that this action is defined by the rule $g.f(v) = f(g^{-1}.v)$. Alternatively, if the dimension of V is l, then the ring of all polynomial maps on V is $\mathbb{C}[V] = \mathbb{C}[x_1, \ldots, x_l]$ where the linear forms $\mathbb{C}x_1 + \ldots + \mathbb{C}x_l$ can be identified with the dual vectorspace V^*. V^* is a GL_n-representation with action $g.\phi$ for $V \xrightarrow{\phi} \mathbb{C}$ in $\mathbb{V}^*$ defined by $(g.\phi)(v) = \phi(g^{-1}.v)$. Hence, the action of GL_n on the polynomial ring coincides with the natural action of GL_n on the *symmetric algebra* of V^*

$$S(V^*) = \oplus_{i=0}^{\infty} S^i\ V^* \simeq \mathbb{C}[x_1, \ldots, x_l]$$

where $S^i\ V^*$ is the subspace of symmetric tensors in $V^* \otimes \ldots \otimes V^*$ (i terms). The action of GL_n on $\mathbb{C}[V]$ is *locally finite*, that is, any finite set of polynomials $\{f_1, \ldots, f_a\} \subset \mathbb{C}[V]$ is contained in a finite dimensional GL_n-subrepresentation W of $\mathbb{C}[V]$. Indeed, let k be such that all f_i have total degree at most k, then they are all contained in the finite dimensional GL_n-subrepresentation $\oplus_{i=0}^{k} S^i\ V^*$.

Recall that the *Hilbert basis theorem* asserts that any ideal $I \triangleleft \mathbb{C}[V]$ is finitely generated. We say that an ideal $I \triangleleft \mathbb{C}[V]$ is GL_n*-stable* if $g.I \subset I$ for all $g \in GL_n$.

Definition 10.3. A GL_n*-variety* is a couple (V, I) where

- V is a finite dimensional GL_n-representation, and
- I is a GL_n-stable ideal of $\mathbb{C}[V] = S(V^*)$.

Examples of GL_n-varieties.

Jordan forms	$V = M_n(\mathbb{C})$
	$g.m = gmg^{-1}$
	$\mathbb{C}[V] = \mathbb{C}[m_{11}, \ldots, m_{nn}]$
	$I = 0$

Examples of GL_n-varieties.

Dynamical systems	$V = M_{n \times m}(\mathbb{C}) \oplus M_n(\mathbb{C}) \oplus M_{p \times n}(\mathbb{C})$ $g.(A,B,C) = (gA, gBg^{-1}, Cg^{-1})$ $\mathbb{C}[V] = \mathbb{C}[a_{11}, \ldots, a_{nm}, b_{11}, \ldots, b_{nn}, c_{11}, \ldots, c_{pn}]$ $I = 0$
Hilbert schemes	$V = \mathbb{C}^n \oplus M_n(\mathbb{C}) \oplus M_n(\mathbb{C})$ $g.(v,X,Y) = (gv, gXg^{-1}, gYg^{-1})$ $\mathbb{C}[V] = \mathbb{C}[v_1, \ldots, v_n, x_{11}, \ldots, x_{nn}, y_{11}, \ldots, y_{nn}]$ $I = (\sum_k x_{ik}y_{kj} \mid \forall\, i,j)$