Representation Theory I

September 21, 2015

Remark. This is the live-texed notes of Representation Theory I course held by Mátyás Domokos at CEU in the fall trimester of 2014. Every error and typo in the text is mine.

Plans:

1. Construct the irreducible representations of S_n
2. Find the polynomial representations of $GL_n(\mathbb{C})$
3. It depends on.

Literature:

1. C. Procesi: Lie groups - An approach through invariants and representations (Springer, 2007)
2. W. Fulton, J. Harris: Representation Theory - A first course
3. I. G. Macdonald: Symmetric Functions and Hall polynomials

First Lecture, 3rd of October

1 General notions of representations theory

Definition 1.1. A *representation* (ρ, V) of a group G is a vector space V over a field $\mathbb{F}$ together with a group homomorphism $G \to GL(V)$.

Notation: If the representation (ρ, V) is fixed then we just write $gv := (\rho(g))(v) \in V$.

Definition 1.2. Given a representation (ρ, V) of G. A subspace $W \leq V$ is a *ρ-invariant subspace* if $\rho(g)W \subseteq W$ for all $g \in G$. In this case, we can have another representation:

$$\rho_W : G \to GL(W) \qquad g \mapsto \rho(g)|_W \in GL(W)$$

called *subrepresentation.*

Similarly, if W is still an invariant subspace, then we can define the corresponding *factor-representation*

$$\rho_{V/W} : G \to GL(V/W)$$

where $\rho_{V/W}(g)(v + W) := \rho_V(g)(v) + W$.

Definition 1.3. A representation (ρ, V) is *simple* (or *irreducible*) if it has no non-trivial invariant subspaces (only $\{0\}$ and V).

Remark 1.4. Given a representation (ρ, V) of G, we can consider the $\mathbb{F}$-algebra

$$A := \mathrm{Span}_{\mathbb{F}}\{\rho(g) \mid g \in G\} \leq \mathrm{End}_{\mathbb{F}}(V)$$

Over this algebra, V becomes an A-module.

Definition 1.5. A *group algebra* $\mathbb{F}G$ of a group is a vector space with $\mathbb{F}$-basis $\{e_g \mid g \in G\}$ and with the multiplication rule

$$\sum a_g e_g \cdot \sum b_h e_h = \sum_{l \in G} \left(\sum_{gh=l} a_g b_h \right) e_l$$

Remark 1.6. If It does not cause confusion then e_g will be simply denoted by g.

Remark 1.7. By definition of $\mathbb{F}G$, a group homomorphism $\rho(G) \to GL(V)$ extends uniquely to an $\mathbb{F}$-algebra homomorphism $\mathbb{F}G \to \mathrm{End}_{\mathbb{F}}(V)$. This gives a one-one correspondence of $\mathbb{F}G$-modules and the representations of G on an $\mathbb{F}$-vector space.

Definition 1.8. Let R be a ring. Then $e \in R$ is called an *idempotent* if $e^2 = e$.

Example 1.9. For example a projection of a linear space is idempotent.

Proposition 1.10. *Let $e \in R$ be an idempotent. Then ${}_R R = Re \oplus R(1-e)$.*

Proof. Indeed, for all $r \in R$: $r = re + r(1-e)$ and the intersection cannot contain anything because if $xe = y(1-e)$ then $xe = xee = y(1-e)e = y(e-e^2) = 0$. □

Proposition 1.11. *Let $e \in R$ be an idempotent. Then $\mathrm{End}_R(Re) \cong (eRe)^{op}$ as rings.*

Proof. Given an element $exe \in eRe$ for some $x \in R$, then we get an endomorphism

$$\mu_{exe} : Re \to Re \qquad y \mapsto yexe$$

It is clear that it commutes with multiplication from the left. Conversely, if $\varphi \in \mathrm{End}_R(Re)$ then we can find the "corresponding" exe the following way:

$$\varphi(e) = \varphi(e^2) = e\varphi(e)$$

Therefore, if we represent $\varphi(e)$ as xe for some $x \in r$ then we got $xe = exe \in eRe$. Then it is easy to check that $\mu_{\varphi(e)} = \varphi$ using only the definition of endomorphism. Similarly, $\mu_{exe}(e) = exe$ so it is indeed a bijection. We now only have to investigate how they multiply. But one can check that

$$\mu_{exe}\mu_{eye} = \mu_{(eye)(exe)}$$

therefore μ gives an anti-homomorphism so the statement is verified. □

Remark 1.12. Similarly, if $e, f \in R$ are two idempotent then we can also consider the space $\mathrm{Hom}_R(Re, Rf) = eRf$ as abelian groups.

Definition 1.13. The idempotents $e, f \in R$ are *orthogonal*, if $ef = 0 = fe$. Besides, an idempotent $e \in R$ is *primitive*, if it cannot be expressed as a sum of orthogonal idempotents (apart from the trivial "sum" $e = e + 0$).

Proposition 1.14. *Let R be a finite dimensional semisimple $\mathbb{F}$-algebra. If $e \in R$ is a primitive idempotent then Re is a simple submodule. Conversely, if S is a simple R-submodule in R then S is generated by some primitive idempotent.*

Proof. (In the first statement we do not use semisimplicity) Suppose that $e \in R$ is a primitive idempotent and write $Re = M \oplus N$ where M, N are nonzero submodules. Let $\pi \in \mathrm{End}_R(Re)$ be the projection $Re \to M$ with kernel N. Then $0 \neq \pi = \pi^2 \in eRe$ where $e \neq \pi$. Therefore, π and $e - \pi$ are orthogonal idempotents. It contradicts the assumption of primitivity.

Conversely, let S be a simple R-submodule. Then, by semisimplicity, $R \cong S \oplus T$ so we can decompose $1 = e + (1-e) \in S \oplus T$ where $e \in S$ and $1 - e \in T$. These are idempotents because

$$e = e1 = e(e + (1-e)) = e^2 + 0$$

where $e \in S$, $1-e \in T$ so their product $e(1-e) = (1-e)e \in S \cap T = 0$. Therefore, the statement is valid. □

Remark 1.15. By Maschke's theorem: $\mathbb{C}S_n$ or $\mathbb{Q}S_n$ are semisimple algebras.

Lemma 1.16. *Let e be an idempotent in an $\mathbb{F}$-algebra R. If $\dim_{\mathbb{F}}(eRe) = 1$ then e is primitive.*

Proof. Suppose $eRe = \mathbb{F}e$ and $e = a + b$ where a and b are orthogonal idempotents. Then $ea = (a+b)a = a^2 + 0 = a$. Similarly, $ae = a$ so $a \in eRe$. Therefore, $a = \lambda e$. The same can be said about b so $b = \mu e$ and – by orthogonality – either λ or μ has to be zero. □

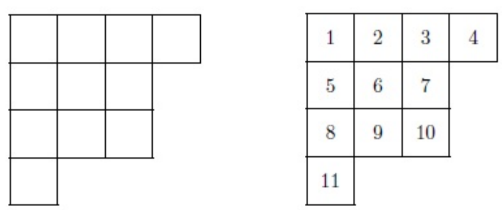
2 The representation theory of S_n

2.1 Young Tableau

Remark 2.1. The isomorphism classes of irreducible representations over $\mathbb{C}$ of S_n are in one-to-one correspondence with the conjugacy class of S_n. (It works for any finite group, see Topics in Algebra) Moreover, the conjugacy classes of S_n are in one-to-one correspondence with the partitions of n, i.e. the sequences $\lambda = (\lambda_1, \ldots, \lambda_k)$ where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 1$ are integers and $\sum_i \lambda_i = n$.

Question. *How can we find the idempotents corresponding to a partition?*

Definition 2.2. A *Young diagram* is a table of squares where the first row has λ_1 boxes, the second has λ_2 and so on. A *Young tableau* is this table filled with the numbers $\{1, 2, \ldots, n\}$ in arbitrary order, each used once. The *shape* of the Young tableau is the corresponding Young diagram without numbers. Notation: $\lambda \vdash n$ means that λ is a Young diagram of size n.



Example for a Young diagram and a Young tableau

Remark 2.3. To emphasize that these are quite natural objects, we call attention to the fact that a Young tableau is nothing else but a – more or less unique – visual "representation" of a partition of the set $\{1, 2, \ldots, n\}$ while a Young diagram stands for its S_n-orbit.

Definition 2.4. Let T be a Young tableau. Then we define distinguished subgroups of S_n:

$$S_n \geq R_T := \{\pi \in S_n \mid i \text{ and } \pi(i) \text{ are in the same row of } T \text{ for all } 1 \leq i \leq n\}$$

$$S_n \geq C_T := \{\pi \in S_n \mid i \text{ and } \pi(i) \text{ are in the same column of } T \text{ for all } 1 \leq i \leq n\}$$

Besides, we define *distinguished elements* of $\mathbb{Q}S_n$ namely $s_T = \sum_{\pi \in R_T} \pi$, $a_T = \sum_{\pi \in c_T} (-1)^\pi \pi$ and $c_T = s_T a_T$.

Definition 2.5. *Order* partitions of n lexicographically: let λ and μ be Young diagrams: then $\lambda > \mu$ if $\lambda_1 > \mu_1$ or if $\lambda_1 = \mu_1$ and $\lambda_2 > \mu_2$ and so on.

Definition 2.6. Let T be a tableau of shape λ. Then we apply a $\pi \in S_n$ on it in an obvious way so we get another tableau πT of the same shape. So we defined a permutation action of G on tableaus of the shape λ.

Lemma 2.7. *Let λ, μ be two Young diagrams and assume that $\lambda \geq \mu$. Let S be a tableau of shape λ and T be a tableau of shape μ. Then exactly one of the following holds:*

1. Either there are two numbers $i \neq j$ that appear in the same row of S and in the same column of T.

2. Or $\lambda = \mu$ and there exists $p \in R_S$ and $q = C_T$ such that $pS = qT$.

In particular, if $\lambda = \mu$ then the statement says that "the only obstruction to rearrange a tableau into another one by a row-permutation and a column permutation is the trivial obstruction i.e. the first case".

Proof. If $\lambda_1 > \mu_1$ then by the Pigeon-hole principle, we have two entries in the first row of S must belong to the same column. If $\lambda_1 = \mu_1$ and the entries in the first row are in different column in T (i.e. the first case does not hold for the entries of the first row of S) then each column has exactly one of them and we can permute these entries into the first row by C_T. In other words, there exists a $q \in C_T$ such that the entries of the first row of S are in the first row of qT.

Now, we can use induction on the number of rows of S and by the same procedure, we either stop at two entries for which the first case holds, or if not then $\lambda = \mu$ and we get a $q \in C_T$ such that the set of elements of the rows in S and in qT are the same, so there exists a $p \in R_T$ such that $pS = qT$. □

Corollary 2.8.

1. If $\lambda > \mu$ and S is a tableau of shape λ and T is a tableau of shape μ. Then for all $z \in S_n$ there exist transpositions $u \in R_S$ and $v \in C_T$ such that $uz = zv$.

2. If $z \notin R_T C_T$ then there exist transpositions $u \in R_T, v \in C_T$ such that $uz = zv$.

Proof. In the first statement, only the first case is possible so $(ij) = u = v$ will do the job. For the second statement, apply the lemma on $S = z^{-1}T$ and T. If the second option holds then $pz^{-1}T = qT$ for some $p \in R_{z^{-1}T} \overset{\text{Homework}}{=} z^{-1}R_T z$ and $q \in C_T$. Then $q = pz^{-1} = z^{-1}p'$ where $p' = zpz^{-1} \in R_T$. Therefore, $z = p'q^{-1} \in R_T C_T$ and that contradicts the assumption. Hence, we are in the first case so there exist $i \neq j$ such that $(ij) \in R_{z^{-1}T} \cap C_T = z^{-1}R_T z \cap C_T$ so $v = (ij)$ and $u = zvz^{-1}$ will do the job. □

Proposition 2.9. *Let S and T be Young tableaux of shape $\lambda > \mu$. Assume that for $b \in \mathbb{Q}S_n$ we have $pb = b$ for all $p \in R_S$ and $bq = (-1)^q b$ for all $q \in C_T$. Then $b = 0$.*

Besides, if for some $b \in \mathbb{Q}S_n$ we have $pb = b$ for all $p \in R_T$ and $bq = (-1)^q b$ for all $q \in C_T$ then a is a scalar multiple of c_T.

Remark 2.10. This proposition makes more sense after reading the next theorem: the facts that $c_U R c_T = 0$ if and only if $U \neq T$ and $c_T R c_T = \mathbb{Q} c_T$ will depend on this.

Proof. Let $b = \sum_{z \in S_n} b_z z$ and suppose the assumptions of the first statement. Choose u and v as in the first part of Corollary 2.8. By the assumption $ub = b$ therefore $b_{uz} = b_z$. The other equation says that $bv = -b$ so $b_{zv} = -b_z$. However, then $b_z = b_{uz} = b_{zv} = -b_z \Rightarrow b_z = 0$ so all the coefficients are zero.

For the other part, consider an $b = \sum_{z \in S_n} b_z z$. The same argument as before shows that $b_z = 0$ for all $z \notin R_T C_T$. Now, let $z = pq$ for some $p \in R_T$ and $q \in C_T$. Then

$$b_z = b_{pq} = b_q = (-1)^q b_1$$

using the assumed equations on p and on q so we found that $b = b_1 c_T$. □

Second Lecture, 10th of October

Theorem 2.11. *Let us denote $\mathbb{Q}S_n$ by R. Then the following are true:*

1. $c_T R c_T = c_T R a_T = s_T R c_T = s_T R a_T = \mathbb{Q} c_T$

2. $c_T^2 = p(\lambda) c_T$ *for some positive integer* $p(\lambda)$

3. $\dim_{\mathbb{Q}}(R c_T) = \frac{n!}{p(\lambda)}$

4. *Let U and V be tableaux of shapes $\lambda > \mu$ respectively. Then $s_U R a_V = 0 = a_V R s_U$.*

5. *If U, V are tableaux of different shapes then $c_U R c_V = 0$.*

Proof. For the first, it is enough to prove that $s_T R a_T = \mathbb{Q}c_T$ and $c_T R c_T \neq 0$ because all the sets are in $s_T R a_T$ and all of them contain $c_T R c_T$. By the definition of s_T, we have $z s_T = s_T$ for all $z \in R_T$. Similarly, $a_T z = (-1)^z a_T$ for all $z \in C_T$ by the definition. Therefore, for all $a \in s_T R a_T$ we have $za = a$ for all $z \in R_T$ and $(-1)^z a z = a$ for all $z \in C_T$. By the second part of Proposition 2.9, we get that there is at most "one" (i.e. up to scalar multiple) such element, namely the elements of $\mathbb{Q}c_T$. To get the statement, indirectly assume that $c_T R c_T = 0$. Then $R c_T R$, the generated non-trivial ideal by c_T would be nilpotent and that is impossible since in a semisimple algebra (where we are by Maschke's theorem) there are no nilpotent ideals.

To prove statement 2) we have to use statement 1): $c_T^2 \in c_T R c_T = \mathbb{Q}c_T$. I.e. $c_T^2 = p(\lambda) c_T$ for some $p(\lambda) \in \mathbb{Q}$, so we have to prove that it is in fact a positive integer.

Take the character Tr of the regular permutation representation of S_n, i.e.

$$\mathrm{Tr}(a) = \begin{cases} n! & a = 1_{S_n} \\ 0 & \text{otherwise} \end{cases}$$

In other words, $\mathrm{Tr}(a)$ is $n!$-times the coefficient of 1_{S_n} in that group algebra element. Now, we can apply it on $c_T^2 = p(\lambda) c_T$ so we get $\mathrm{Tr}(c_T^2) = p(\lambda)\mathrm{Tr}(c_T)$. The left hand side is $n!$-times the coefficient of 1_{S_n} in c_T^2 which is an integer because $c_T, c_T^2 \in \mathbb{Z}S_n$. The right hand side is $p(\lambda)$-times $n!$ because the coefficient of 1_{S_n} in c_T is one. Therefore, dividing by $n!$ we got that $p(\lambda)$ is the (integer) coefficient of 1_{S_n} in c_T^2.

We still have to prove the positivity but that will follow from the third part. At least we know that it is non-zero: indeed, c_T^2 is non-zero because $\mathrm{Tr}(c_T) = n!$ (since the coefficient of 1_{S_n} in c_T is one) which is the same as the trace of the linear transformation induced by the right multiplication by c_T on R, but a nilpotent linear transformation cannot have non-zero trace, hence $p(\lambda) \neq 0$.

To prove statement 3) consider $e_T := \frac{1}{p(\lambda)} c_T$. Then this e_T is obviously an idempotent. In fact, the right multiplication by this element is exactly the projection to the subspace $R e_T = R c_T$. However, the trace of a projection (which is now the same as the character value Tr of the corresponding group algebra element) is the dimension of its image. In our case, $\dim_{\mathbb{Q}}(R c_T) = \mathrm{Tr}(e_T) = \frac{n!}{p(\lambda)} \in \mathbb{Z}_{>0}$. (Consequence: $p(\lambda)$ is a positive divisor of $n!$.)

Statement 4) basically follows from the first part of Proposition 2.9: $s_U R a_V = 0$ because the elements $s_U r a_V$ satisfy the assumptions of the proposition. For the reversed product, note that $a_V R s_U \subseteq R a_V R s_U R$ where the latter is a two-sided ideal. However, $(R a_V R s_U R)^2 \subseteq R(s_U R a_V)R = 0$ so it is a nilpotent ideal which means that $R a_V R s_U R = a_V R s_U = 0$ by semisimplicity (in a semisimple ring, there are no nilpotent ideals).

The last statement is obvious by part 4), i.e. $c_U R c_V = s_U a_U R s_V a_V \subseteq s_U R a_V = 0$, if $\lambda > \mu$. If we have $\lambda < \mu$ then we have to apply the same procedure as in 4). □

Remark 2.12. We have not seen yet that $p(\lambda)$ only depends on the shape, as the notation "states". It will follow from the next proposition.

Theorem 2.13. *Let T be a tableau of shape λ. Then*

1. *The elements $e_T = \frac{c_T}{p(\lambda)}$ for each tableau T are primitive idempotents in R.*

2. *The left ideals $R e_T$ give all isomorphism classes of irreducible R-modules.*

3. *Moreover, $R e_T$ is isomorphic to $R e_{T'}$ if and only if the shape of T is the same as the shape of T'. (In other words, $T' = zT$ for some $z \in S_n$)*

4. *These representations remain irreducible even when considered as representations of $\mathbb{C}S_n$ and not as of $\mathbb{Q}S_n$.*

Corollary 2.14. *$p(\lambda)$ depends only on λ and not on the numbering.*

Proof. (of the corollary) We saw in the proof of Proposition 2.11 that

$$Coeff_{1_{S_n}}(c_T^2) = \frac{1}{n!}\mathrm{Tr}(c_T^2) = p(\lambda)$$

Now, if we interchange the tableau T with some $T' = zT$ then $c_{T'} = zc_Tz^{-1}$ so $\mathrm{Tr}(c_{T'}^2) = \mathrm{Tr}(zc_T^2z^{-1})$, i.e. the coefficient of 1 is the same. □

Proof. (of theorem 2.13) It is a trivial computation that e_T is idempotent. The primitivity follows from $\mathrm{End}(Re_T, Re_T) \cong e_TRe_T = c_TRc_T = \mathbb{Q}c_T = \mathbb{Q}e_T$ which means that there are no more idempotents in the endomorphism ring, so it cannot be decomposed.

To prove 2) we first have to realize that the first part means that Re_T is irreducible. They are non-isomorphic for different shapes because $\mathrm{Hom}(Re_T, Re_{T'})^{op} = e_TRe_{T'} = c_TRc_{T'}$ which is zero by proposition 2.11. Conversely, if the shape is the same, i.e. $T' = zT$ then $c_{T'} = zc_Tz^{-1}$. Therefore, $\mathrm{Hom}(Re_T, Re_{T'}) = e_TRe_{T'} = e_TRze_Tz^{-1} \ni e_Tz^{-1} \neq 0$ so the hom-space is non-trivial. However, these are simple modules so, if there is a non-zero homomorphism then they are isomorphic by Schur's lemma.

This way, we found so many representation so there cannot be more: these representations can either split or not when changing the base field from $\mathbb{Q}$ to $\mathbb{C}$ but each of them give at least one distinct representation. However, over $\mathbb{C}$ we know that the number non-isomorphic representations of a group algebra is the same as the conjugacy classes. Which is, in this case, the same as the number of partitions, i.e. the number of possible shapes of Young diagrams. We found a representation for every different shape, so there cannot be more isomorphism classes. The only thing what can happen is that a $\mathbb{Q}$-irreducible module can split into several isomorphic $\mathbb{C}$-irreducibles. The reason why it cannot happen is based on the first part of 2.11: namely that the endomorphism rings of the $\mathbb{Q}$-irreducibles are $\mathrm{End}_R(Re_T)^{op} \cong c_TRc_T \cong \mathbb{Q}$ and not some finite dimensional $\mathbb{Q}$-division algebra. It is enough because if we look at the Wedderburn decomposition of $\mathbb{Q}S_n$ we only get that $\mathbb{Q}S_n \cong \oplus_{i=1}^q M_{n_i}(D_i)$ where these D_i division rings can be "synthetized" by the endomorphism rings of the irreducibles modules. Now, when all the endomorphism rings are $\mathbb{Q}$ we have $\mathbb{Q}S_n \cong \oplus_{i=1}^q M_{n_i}(\mathbb{Q})$ so when "changing the base field" – what is made precise by taking the algebra $\mathbb{Q}S_n \otimes_{\mathbb{Q}} \mathbb{C} \cong \mathbb{C}S_n$ – it yields $\mathbb{C}S_n \cong \oplus_{i=1}^q M_{n_i}(\mathbb{Q}) \otimes \mathbb{C} \cong \oplus_{i=1}^q M_{n_i}(\mathbb{C})$. So it follows that the $I_\lambda \otimes \mathbb{C}$ has the same length decomposition into irreducibles as I_λ has so $Rc_T \otimes \mathbb{C}$ has to remain irreducible. So we got that the representations of S_n do not split if we extend the field to $\mathbb{C}$ as we stated. □

Remark 2.15. All the proofs carry over for $\tilde{e_T}$ which is defined from $\tilde{c_T} = a_Ts_T$ instead of s_Ta_T. So we could get the decomposition in a different way. To explore this, let us realize that $\mathbb{Q}S_n$ has a (linear, anti-homomorphic) involution $a \mapsto a^*$. Namely,

$$\left(\sum a_z z\right) = \left(\sum a_z z^{-1}\right)$$

This involution gives the connection between e_T and $\tilde{e_T}$ as well.

Exercise 2.16. $Re_T \cong R\tilde{e_T}$.

Remark 2.17. $R\tilde{e_T} \subseteq Rs_T$. The latter Rs_T is exactly the induced representation of the trivial representation of the subgroup $S_{\lambda_1} \times S_{\lambda_2} \times \cdots \times S_{\lambda_k} \leq S_n$ where $\lambda_i \geq \lambda_j \geq 0$ if $i < j$ and $\sum \lambda_i = n$. This also gives a method to calculate the characters of the irreducible representations.

2.2 Characters by the Schur polynomials

Motivation: We would like to to compute the irreducible characters of S_n.

Definition 2.18. Let $\rho : G \to GL(V)$ be a representation of a finite group G on a finite dimensional $\mathbb{C}$-vector space V. The *character* of ρ is the function $\chi_g : G \to \mathbb{C}$; $g \mapsto \mathrm{Tr}(\rho(g))$.

Remark 2.19. By generalities, we know that $\chi_\lambda(\mu) \in \mathbb{Z}$ for all $\mu \in S_n$ because every $\mathbb{C}$-representation can be given by rational matrices so $\chi_\lambda(\mu) \in \mathbb{Q}$ but they are also algebraic integers (see Topics in algebra) so $\chi_\lambda(\mu)$ is an integer.

Definition 2.20. S_n acts on $\mathbb{Z}[x_1, \dots, x_n]$ as $s(f)(x_1, \dots, x_n) = f^s(x_1, \dots, x_n) = f(x_{s(1)}, \dots, x_{s(n)})$. By this, we can define the *symmetric polynomials* as $\mathbb{Z}[x_1, \dots, x_n]^{S_n} = \{f \in \mathbb{Z}[x_1, \dots, x_n] \mid f^s = f \ \forall s \in S_n\}$.

The usual notation for $\mathbb{Z}[x_1, \dots, x_n]^{S_n}$ is Λ^n. This has a natural decomposition by the degree $\Lambda^n = \oplus_{d=0}^{\infty} \Lambda_d^n$ where Λ_d^n denotes the symmetric homogeneous degree d polynomials. As an abelian group, Λ_d^n is a free abelian group, spanned by m_λ's that are the S_n-orbit sums of a monomial $x_1^{\lambda_1} \cdot \dots \cdot x_n^{\lambda_n}$ where λ is a partition of rank d and with height at most n (i.e. there are at most n non-zero partitions in λ) .

Definition 2.21. An $f \in \mathbb{Z}[x_1, \dots, x_n]$ is *alternating* if $f^s = \operatorname{sign}(s) f$ for all $s \in S$.

Example 2.22. The Vandermonde determinant is alternating.

Proposition 2.23. *A polynomial $f \in \mathbb{Z}[x_1, \dots, x_n]$ is alternating if and only if $f = v \cdot g$ for some $g \in \Lambda^n$.*

Proof. $\Leftarrow$ is obvious, an alternating times a symmetric is always alternating: $(vg)^s = v^s g^s = \operatorname{sign}(s) v \cdot g$. For the other direction: Make the substitution $x_2 \mapsto x_1$. Then by the transposition $s = (12)$ we get that $f|_{x_2 \mapsto x_1} = f^s|_{x_2 \mapsto x_1} = (-1) f|_{x_2 \mapsto x_1}$ i.e. it is zero after the substitution. This means that $x_1 - x_2$ divides f in $\mathbb{Z}[x_1, \dots, x_n]$. The same holds for $x_i - x_j$ for all $i < j$. Since $\mathbb{Z}[x_1, \dots, x_n]$ is a unique factorization domain, we get that $v = \prod_{i<j}(x_i - x_j) \mid f$. So there exists a decomposition $f = v \cdot g$ where g must be symmetric (since alternating divided by alternating is symmetric). $\square$

Notation: For $\alpha_1 > \alpha_2 > \dots > \alpha_n \geq 0$ the polynomial A_α is defined as

$$A_\alpha(x_1, \dots, x_n) = \sum_{s \in S_n} \operatorname{sign}(s) x_1^{\alpha_{s(1)}} \cdot x_2^{\alpha_{s(2)}} \cdot \dots \cdot x_n^{\alpha_{s(n)}}$$

i.e. it is the alternating orbit sum of $x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}$. Obviously, this means that $s(A_\alpha) = \operatorname{sign}(s) A_\alpha$.

Proposition 2.24. $\{A_\alpha \mid \alpha_1 > \dots > \alpha_n \geq 0\}$ *is a free $\mathbb{Z}$-module basis in the $\mathbb{Z}$-module of alternating polynomials inside $\mathbb{Z}[x_1, \dots, x_n]$. (Straightforward by antisymmetrization.)*

Definition 2.25. Take a partition of rank d which has height at most n (i.e. it has at most n rows). Let us denote the sequence $(n-1, n-2, \dots, 2, 1, 0)$ by δ, i.e. $v = A_\delta$ (easy exercise). Notation:

$$S_\lambda(x_1, \dots, x_n) := \frac{A_{\lambda+\delta}}{v} \in \mathbb{Z}[x_1, \dots, x_n]$$

named the *Schur polynomials.*

Remark 2.26. An alternative form of them is

$$S_\lambda(x_1, \dots, x_n) = \frac{\det((x_i^{\lambda_j + n - j}))}{\det((x_i^{n-j}))}$$

Corollary 2.27. *The set $\left\{S_\lambda(x_1, \dots, x_n) \mid \lambda \vdash d, \ \operatorname{ht}(\lambda) \leq n\right\}$ is a free $\mathbb{Z}$-module basis of Λ_d^n.*

Proof. By the definition of the Schur polynomials and Proposition 2.23, it is enough to prove that $A_{\lambda+\delta}$'s form a $\mathbb{Z}$-module basis of the alternating polynomials of degree $\binom{n}{2} \cdot d$ of n variables. However, that is true by Proposition 2.24 so we got the corollary. $\square$

Remark 2.28. In summary, there is a quite natural linear bijection between Λ^n and the alternating polynomials in n variables and the Schur polynomials are exactly the natural basis of the alternating polynomials pulled back to Λ^n (and parametrized by the partitions).

Third Lecture, 17th of October

Theorem 2.29. (Cauchy formula) *For the Schur-polynomials the following holds:*

$$\frac{1}{\prod_{i=1}^n \prod_{j=1}^m (1-x_iy_j)} = \sum_{\lambda \text{ partition}} S_\lambda(x_1,\dots,x_n)S_\lambda(y_1,\dots,y_m) \in \mathbb{Z}[[x_1,\dots,x_n,y_1,\dots y_m]]$$

Remark 2.30. On the right hand side, the sum runs over the partitions that have height at most $\min\{n,m\}$ (since the Schur polynomials only defined in that case).

Remark 2.31. We will discuss the representation-theoretic meaning of the statement later in Corollary 3.25.

Proof. It is sufficient to deal with the case when $n = m$ because of the following lemma.

Lemma 2.32. *Let λ be a partition with $\mathrm{ht}(\lambda) \le n$. Then*

$$S_\lambda(x_1,\dots,x_n)|_{x_n=0} = \begin{cases} 0 & \text{if } \lambda_n > 0 \\ S_\lambda(x_1,\dots,x_{n-1}) & \text{if } \lambda_n = 0 \end{cases}$$

Proof. Follows from $A_\alpha(x_1,\dots,x_{n-1},0) = A_{\alpha|_{n-1}}(x_1,\dots,x_{n-1}) \cdot \prod_{i=1}^{n-1} x_i$ if $\alpha_1 > \dots > \alpha_{n-1} > \alpha_n = 0$. Computations omitted. □

The idea of the proof of the Cauchy formula is to express the quantity in question as a determinant of a matrix and the two sides of the formula are the results of computing this determinant in two different ways. In details:

Lemma 2.33. *Let us use the notation:*

$$A := \left(\left(\frac{1}{1-x_iy_j}\right)\right)_{i,j} \qquad \textit{for this we have:} \qquad \det(A) = \frac{V(x)V(y)}{\prod_{i,j=1}^n (1-x_iy_j)}$$

Proof. Subtract the first row of A from the i-th row for $i = 2,\dots,n$. Then we get B where $B_{ij} = \frac{1}{1-x_iy_j} - \frac{1}{1-x_1y_j} = \frac{y_j(x_i-x_1)}{(1-x_iy_j)(1-x_1y_j)}$ for $i \ge 2$. Therefore, we can take out $(x_i - x_1)$ from the i-th row and $\frac{1}{1-x_1y_j}$ from the j-th column. Hence,

$$\det A = \det B = \prod_{i=2}^n (x_i - x_1) \cdot \prod_{j=1}^n \frac{1}{1-x_1y_j} \cdot \det C$$

where the first row of C is constant 1, and $c_{ij} = \frac{y_j}{1-x_iy_j}$ for all $i \ge 2$. Now, subtract the first column from the j-th column for $j = 2,\dots,n$ so we get a matrix D such that $\det C = \det D$ where $d_{11} = 1$, $d_{1j} = 0$ for all $i \ge 2$. Besides, for $i,j \ge 2$

$$d_{ij} = \frac{y_j}{1-x_iy_j} - \frac{y_1}{1-x_iy_1} = \frac{y_j - y_1}{(1-x_iy_j)\cdot(1-x_iy_1)}$$

So by again some row-operations, we get a similar matrix E such that $\det D = \det E$, $d_{ij} = e_{ij}$ for $i,j \ge 2$ but $e_{1j} = e_{i1} = 0$ for all $i,j \ge 2$. (This step is clearly unnecessary since if the first row is $(1,0,0,\dots,0)$ then the first column does not affect the determinant. Nevertheless, it makes the next take-out step less confusing.) Now, we can take out $y_j - y_1$ from the j-th column ($j > 1$) and $1 - x_iy_1$ from the i-th row ($i > 1$). So the result is

$$\det A = \frac{\prod_i (x_i - x_1)\prod(y_j - y_1)}{(1-x_1y_1)\prod_{i\ge 2}(1-x_iy_1)\prod_{j\ge 2}(1-x_1y_j)} \cdot \det\left(A|_{i,j\ge 2}\right)$$

This proves the stated formula, by induction. □

Back to the theorem of Cauchy formula: The proof is to compute $\det A$ in another way. The fraction $\frac{1}{1-x_iy_j} \in \mathbb{Z}[[x_1,\dots,y_m]]$ can be written as $\sum_{k=0}^{\infty}(x_iy_j)^k \in \mathbb{Z}[x_1,\dots,x_n,y_1,\dots,y_n]$. Therefore, by the (row-)multilinearity of the determinant, we get that

$$\det(A) = \sum_{k_1=0}^{\infty}\cdots\sum_{k_n=0}^{\infty}\det(B_{k_1,\dots,k_n})$$

where $B_{k_1,\dots,k_n} = (((x_iy_j)^{k_i}))_{n\times n}$. But we can compute these determinants as well:

$$\det(B_{k_1,\dots,k_n}) = \prod_i x_i^{k_i}\cdot\det(y_j^{k_i}) = \begin{cases}\prod_i x_i^{k_i}A_{\lambda+\delta}(y) & \text{if } k_i's \text{ are distinct}\\ 0 & \text{if not}\end{cases}$$

where λ is the partition corresponding to the ordered sequence of $k_i - i$'s. Therefore, summing $\det(B_{k_1,\dots,k_n})$ for the same k_i's but in all permutations, it yields $A_{\lambda+\delta}(x)\cdot A_{\lambda+\delta}(y)$. Hence,

$$\det(A) = \sum_{\lambda \text{ partition}} A_{\lambda+\delta}(x)\cdot A_{\lambda+\delta}(y)$$

The statement follows by simple rearrangement. □

Definition 2.34. Let $k\in\mathbb{N}$. Then $p_k(x_1,\dots,x_n) := \sum_{i=1}^n x_i^k \in \mathbb{Z}[x_1,\dots,x_n]_k^{S_n} = \Lambda_k^n$. By this, for a partition $\mu\vdash d$ we can define $p_\mu = p_\mu(x_1,\dots,x_n) = \prod_{i=1}^{\mathrm{ht}(\mu)} p_{\mu_i}(x_1,\dots,x_n)\in\Lambda_d^n$.

Alternatively, we can use another notation for partitions: $\mu = w^{k_w}\cdot\dots\cdot 2^{k_2}\cdot 1^{k_1}$ stands for the Young diagram with k_i-many length i rows ($w=\mu_1$ the "width" of the diagram). With this notation $p_\mu(x_1,\dots,x_n) = \prod_{i=1}^{w} p_i^{k_i}(x_1,\dots,x_n)$

Example 2.35. For $\mu = (1,1,\dots,1)$ (corresponding to the conjugacy class of $1\in S_d$) we get $p_\mu(x_1,\dots,x_n) = \left(\sum_{i=1}^n x_i\right)^d$. For $\mu=(d)$ it is $p_{(d)}(x_1,\dots,x_n) = p_d(x_1,\dots,x_n) = \sum_{i=1}^n x_i^d$.

Remark 2.36. This will not be a $\mathbb{Z}$-basis, but a $\mathbb{Q}$-basis of Λ_d^n by the Newton-formulas. It states in principle that the power-sums are a $\mathbb{Q}$-algebra generating system of Λ^n but now we took all the n-variable monomials with power-sum arguments that yield degree d polynomials (that is why the definition of p_μ's are reasonable) so they indeed constitute a $\mathbb{Q}$-basis.

Definition 2.37. For $\lambda,\mu\vdash n$ define $c_\lambda(\mu)\in\mathbb{Z}$ by the equation

$$p_\mu(x) = \sum_{\lambda\vdash n} c_\lambda(\mu)S_\lambda(x)\in\Lambda_n^n$$

which is indeed a definition because S_λ is a $\mathbb{Z}$-basis in Λ_n^n for $\lambda\vdash n$.

Remark 2.38. In this subsection we only use $p_\mu(x_1,\dots,x_n)$ where $\mu\vdash n$. The general notion for $\mu\vdash d$ (and so $\lambda\vdash d$ but $x=(x_1,\dots,x_n)$) will be used to understand the relation between S_d and $GL_n(\mathbb{C})$ in the next section.

Definition 2.39. Let us define χ_λ as the character of the representation of S_n on the irreducible module Ra_Ts_T when T is a tableau of shape λ and the ring is $R=\mathbb{Q}S_n$.

Remark 2.40. In the previous subsection we were working with Rs_Ta_T but by Exercise 2.16, it does not make any substantial difference to interchange them by Ra_Ts_T.

Notation: $\chi_\lambda(\mu)$ is the value of χ_λ on any element of the conjugacy class of S_n containing the permutations with cycle type μ.

Theorem 2.41. (Frobenius character formula) $\chi_\lambda(\mu) = c_\lambda(\mu)$

Proof. The plan is the following:

1. By the Cauchy-formula:

$$\sum_{\lambda \text{ partition}} S_\lambda(x) S_\lambda(y) = \frac{1}{\prod_{i,j=1}^n (1 - x_i y_j)} \overset{\text{see later}}{=} \sum_\mu \frac{1}{n(\mu)} p_\mu(x) p_\mu(y)$$

   where on the right hand side, we try to determine the coefficients of the product in a basis of $\mathbb{Z}[x_1, \dots, x_n, y_1, \dots, y_n]$. It turns out that the coefficients are the inverses of $n(\mu) \in \mathbb{Z}^+$ where the definition is

$$n(\mu) = \prod_{j=1}^{w} k_j! \cdot j^{k_j}$$

   which is the size of the centralizer of an element of type $\mu = w^{k_w} \cdot \dots \cdot 2^{k_2} \cdot 1^{k_1}$ where we used the alternative – but now quite expressive – notation for partitions: $w^{k_w} \cdot \dots \cdot 2^{k_2} \cdot 1^{k_1}$ stands for the Young diagram with k_i-many length i rows and $w = \mu_1$ the "width" of the diagram.

2. By the first step, we can easily prove that $\{c_\lambda \mid \lambda \vdash n\}$ is an orthonormal base in the class-function space $\mathrm{Cent}(S_n) \subseteq \mathrm{Fun}(\mathbb{C}S_n, \mathbb{C})$.

   Remark 2.42. The space $\mathrm{Cent}(S_n)$ has a standard inner product:

$$(f_1, f_2) = \frac{1}{|S_n|} \sum_{g \in S_N} \overline{f_1(g)} f_2(g)$$

   and orthonormality in the above statement is meant with respect to this. In another form, we can collect the elements in conjugacy classes, therefore

$$(f_1, f_2) = \sum_{\mu \vdash n} \frac{1}{n(\mu)} \overline{f_1(\mu)} f_2(\mu)$$

   where we used the orbit-stabilizer equation, so the size of a conjugacy class is $|S_n : C_{S_n}(\mu)| = \frac{|S_n|}{|C_{S_n}(\mu)|} = \frac{|S_n|}{n(\mu)}$.

3. (Independent from the first two step) Let $\lambda \vdash n$ be a partition, the subgroup $S_\lambda := S_{\lambda_1} \times \dots \times S_{\lambda_h} \leq S_n$ is called the Young subgroup of λ. Let us denote by β_λ the permutation character of S_n acting on S_n/S_λ. We can determine the values of β_λ.

4. Using step 3, we prove that $p_\lambda(x) = \sum_{\lambda \vdash n} \beta_\lambda(\mu) \cdot m_\lambda(x)$.

5. After some investigation, we get that the base change matrix $\{S_\lambda \mid \lambda \vdash n\} \rightsquigarrow \{m_\lambda \mid \lambda \vdash n\}$ is unitriangular and is a $\mathbb{Z}$-matrix.

6. By the previous two steps we get that c_λ is a virtual character, i.e. it is an integer linear combination of the characters.

7. But if c_λ are virtual characters and give a normed basis then this means these are ± 1 times the irreducible characters.

8. By 5) we get that that these are exactly the irreducible characters

9. In fact, the bijection is not arbitrary but $c_\lambda = \chi_\lambda$.

To prove the first step, take the formal logarithm of the left hand side of Cauchy formula:

$$\log\left(\prod_{i,j=1}^{n}\frac{1}{1-x_iy_j}\right)=\sum_{i,j=1}^{n}-\log(1-x_iy_j)=\sum_{i,j=1}^{n}\sum_{h=1}^{\infty}\frac{(x_iy_j)^h}{h}=\sum_{h=1}^{\infty}\sum_{i,j=1}^{n}\frac{(x_iy_j)^h}{h}=\sum_{h=1}^{\infty}\frac{1}{h}\cdot p_h(x)p_h(y)$$

where $\log(r) := \log(1-(1-r)) = -\sum_{h=1}^{\infty}\frac{1}{h}(1-r)^h$ defined on the elements of $\mathbb{Q}[[x_1,\dots,x_n,y_1,\dots y_m]]$ that have no constant term. One can check that this is a definition (because every term appears at most finitely many times). By this definition, it is easy to prove the usual identity of log: $\log\prod=\sum\log$. (If the identity would not hold for all power series, then it would neither hold on some polynomials. Then if the resulting power series are different then their real evaluation should be different somewhere too. However, that is not the case.) Similarly, $\exp\left(\log\left(\frac{1}{1-f}\right)\right)=\frac{1}{1-f}$ where exp is again defined by the usual power series so we can get back the original formula from the above argument.

Therefore,

$$\prod_{i,j=1}^{n}\frac{1}{1-x_iy_j}=\exp\left(\sum_{h=1}^{\infty}\frac{p_h(x)p_h(y)}{h}\right)=\sum_{k=0}^{\infty}\frac{1}{k!}\left(\sum_{h=1}^{\infty}\frac{p_h(x)p_h(y)}{h}\right)^k=$$

$$=\sum_{k=0}^{\infty}\frac{1}{k!}\sum_{1^{k_1}2^{k_2}\dots l^{k_l}\vdash k}\binom{k}{k_1,k_2,\dots,k_l}\cdot\prod_{j=1}^{l}\frac{p_j(x)^{k_j}p_j(y)^{k_j}}{j^{k_j}}\overset{\text{def's}}{=}\sum_{\mu\text{ partition}}\frac{1}{n(\mu)}p_\mu(x)p_\mu(y)$$

Step 2: Use the n-th degree of step 1 and write p_μ as the linear combination of S_λ's:

$$\sum S_\lambda(x)S_\lambda(y)=\sum_{\mu\vdash n}\frac{1}{n(\mu)}p_\mu(x)p_\mu(y)=\sum_{\mu\vdash n}\frac{1}{n(\mu)}\left(\sum_{\nu\vdash n}c_\nu(\mu)S_\nu(x)\right)\left(\sum_{\eta\vdash n}c_\eta(\mu)S_\eta(y)\right)=$$

$$=\sum_{\nu,\eta\vdash n}S_\nu(x)S_\eta(y)\sum_{\mu\vdash n}\frac{1}{n(\mu)}c_\nu(\mu)c_\eta(\mu)=\sum_{\eta,\nu\vdash n}S_\nu(x)S_\eta(y)\cdot(c_\nu,c_\eta)$$

Therefore – because $S_\nu(x)S_\eta(y)$ is a $\mathbb{Z}$-basis in $\mathbb{Z}[x_1,\dots,x_n,y_1,\dots,y_m]$ – we got that

$$(c_\nu,c_\mu)=\begin{cases}1 & \text{if } \nu=\mu\\ 0 & \text{else}\end{cases}$$

i.e. they are orthonormal.

Step 3: First, let us investigate what is the character of a permutation representation.

Remark 2.43. Let $H<G$ be a finite group. Then G acts by left multiplication on G/H. We can consider the corresponding linear representation on $\mathrm{Fun}(G/H,\mathbb{C})$ by $\rho_g : h\mapsto h\circ g^{-1}$. It is easy to see that the character of ρ is exactly the number of fixed points of g on G/H, or in formula

$$\chi(g)=|\{xH \mid gxH=xH\}|=|\{xH \mid x^{-1}gx\in H\}|\overset{\text{enumeration}}{=}\frac{|\mathrm{Conj}(g)\cap H|\cdot|C_G(g)|}{|H|}=\sum_i\frac{|O_i|\cdot|C_G(g)|}{|H|}$$

where $O_1,\dots,O_l$ are the conjugacy classes in H that $\mathrm{Conj}(g)$ splits into.

Fourth Lecture, 18th of October

By the remark, the character can be expressed as

$$\beta_\lambda(\mu)=\sum_i\frac{|C_G(g)|\cdot|O_i|}{|H|}=\sum_i\frac{n(\mu)\cdot|O_i|}{|H|}$$

where we want to find the sizes of O_i's. For this purpose, let us define the following sum notion on the partition:

Definition 2.44. Let $\lambda \vdash n$, $\mu \vdash m$ where $\lambda = \prod_i i^{k_i}$ and $\mu = \prod_j j^{l_j}$ as usual. Then

$$\lambda \oplus \mu := \prod_i i^{k_i + l_i}$$

In this case, if g has conjugacy class μ then we can take the intersection of $H := S_\lambda$ with the G-conjugacy class of g. Since S_λ is a product of symmetric groups, the conjugacy classes of elements of H are in bijection with sequences of partitions. Therefore, considering $C_G(g) \cap H$ we get several decompositions of $\mu = \mu^{(1)} \oplus \cdots \oplus \mu^{(h)}$ in the above defined sense (one for each H-conjugacy class) where $(\mu^{(1)}, \ldots, \mu^{(h)})$ is the sequence of partitions corresponding to the H-conjugacy class. We will use the following notations:

$$\mu^{(i)} = \prod_j j^{r_{ji}} \qquad (i = 1, \ldots, h)$$

$$\text{with} \quad \sum_{i=1}^{h} r_{ji} = r_j \quad \text{if} \quad \mu = \prod_s s^{r_s}$$

Now, we can get a formula for the cardinality of the conjugacy class $(\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(h)})$ in H:

$$m_{\mu^{(1)},\ldots,\mu^{(h)}} \overset{\text{Orb}-\text{Stab}}{=} \prod_{i=1}^{h} \frac{\lambda_i!}{n(\mu^{(i)})} = \prod_{i=1}^{h} \frac{\lambda_i!}{\prod_j r_{ji}! \cdot j^{r_{ji}}} =$$

$$= \prod_j \frac{r_j!}{r_j! \cdot j^{\sum_{i=1}^{h} r_{ji}}} \prod_{i=1}^{h} \frac{\lambda_i!}{r_{ji}!} = \frac{1}{n(\mu)} \left(\prod_{i=1}^{h} \lambda_i! \right) \prod_j \binom{r_j}{r_{j1}, \ldots, r_{jh}}$$

What can be substituted into the expression of β:

$$\beta_\lambda(\mu) = \sum_{O\,\text{orbit}} \frac{n(\mu) \cdot |O|}{|H|} = \frac{n(\mu)}{\prod_i \lambda_i!} \sum_{\mu = \mu^{(1)} \oplus \cdots \oplus \mu^{(h)},\ \mu^{(i)} \vdash \lambda_i} m_{\mu^{(1)},\ldots,\mu^{(h)}} = \sum_{\mu = \mu^{(1)} \oplus \cdots \oplus \mu^{(h)},\ \mu^{(i)} \vdash \lambda_i} \prod_j \binom{r_j}{r_{j1}, \ldots, r_{jh}}$$

Step 4: We prove that the coefficient of $x_1^{\lambda_1} \ldots x_n^{\lambda_n}$ in $p_\mu(x)$ is the same as this above expression for β.

$$p_\mu(x) = p_1^{r_1}(x) p_2^{r_2}(x) \cdots = \sum_{r_1 = \sum_i r_{1i}} \sum_{r_2 = \sum_i r_{2i}} \cdots \sum_{r_m = \sum_i r_{mi}} \prod_j \binom{r_j}{r_{j1}, \ldots, r_{jn}} x_1^{r_{j1}} \ldots x_n^{r_{jh}} =$$

Therefore, in short, we got that

$$p_\mu(x) = \sum_{\lambda \vdash n} \beta_\lambda(\mu) m_\lambda(x)$$

Step 5: Let us define the coefficients of m_λ in the base of Schur polynomials and vice verse

$$m_\lambda(x) = \sum_{\mu \vdash n} l_{\lambda,\mu} S_\mu(x) \qquad S_\lambda(x) = \sum_{\mu \vdash n} k_{\lambda,\mu} m_\mu(x)$$

(The $k_{\lambda,\mu}$ are usually called Kostka numbers.)

Proposition 2.45. *The matrices $((l_{\lambda,\mu}))_{\lambda,\mu}$ and $((k_{\lambda,\mu}))_{\lambda,\mu}$ are unitriangular integer-matrices if we order the indexes by the lexicographic order.*

Proof. It is enough to prove that $((k_{\lambda,\mu}))_{\lambda,\mu}$ is unitriangular integer-matrix because they are inverses of each other and if one's determinant is one then the other is an integer matrix as well. For a polynomial $f \in \mathbb{Z}[x_1, \ldots, x_n]$ let us denote its main coefficient by $l(f)$. This is multiplicative, i.e. $l(fg) = l(f)l(g)$. So by $A_{\lambda+\delta} = S_\lambda \cdot V$ we get

$$l(S_\lambda(x)) = l(A_{\lambda+\delta}(x))/l(V) = x^{\lambda+\delta}/x^\delta = x^\lambda$$

using the multi-index notations. So $((k_{\lambda,\mu}))_{\lambda,\mu}$ is in fact unitriangular and the statement of the proposition follows. ☐

Corollary 2.46. of the previous proposition.

1. $\beta_\lambda = c_\lambda + \sum_{\nu>\lambda} k_{\nu,\lambda} c_\nu$ *and* $c_\lambda = \beta_\lambda + \sum_{\nu>\lambda} l_{\nu,\lambda}\beta_\nu$ *(Step 6 and Step 7)*
2. $\{c_\lambda \mid \lambda \vdash n\}$ *are the irreducible characters of* S_n *(Step 8)*
3. $c_\lambda = \chi_\lambda$ *(Step 9)*

Proof. To get the first corollary, consider the definition of c_λ's:

$$\sum_\lambda c_\lambda(\mu) S_\lambda = p_\mu = \sum_\lambda \beta_\lambda(\mu) m_\lambda = \sum_\lambda \sum_\nu \beta_\lambda(\mu) l_{\lambda,\mu} S_\nu = \sum_\nu \left(\sum_\lambda \beta_\lambda(\mu) l_{\lambda,\mu}\right) S_\nu$$

therefore $c_\lambda(\mu) = \sum_\nu \beta_\nu(\mu) l_{\nu,\mu}$ because the Schur polynomials form a basis. The second equation follows the same way.

We got that c's are virtual characters (i.e. integer-linear combinations of characters) since β_λ's are characters. However, they form an orthonormal basis. It can happen only if c_λ is ± 1 times an irreducible character (see First and Second Orthogonal Identity of group representations). The fact that c_λ is indeed a character follows from part part a):

Last statement: By the first consequence, c_λ is a summand of β_λ but not of β_μ for $\mu > \lambda$. Moreover, this property characterizes c_λ among the irreducible characters. (One can recursively determine c_λ in the inverse lexicographic order). And the χ_λ irreducible representations have the same property.

Indeed, if the tableau T of shape λ is endowed with the "natural numbering" then $R_T = S_\lambda$, so $Rs_T \cong \mathbb{C}(S_n/S_\lambda)$ as a $\mathbb{C}S_n$ module. What is more, $\mathrm{Hom}(Ra_U s_U, Rs_T) \cong a_U s_U Rs_T \subseteq a_U Rs_T = 0$ if shape$(U) <$ shape$(T) = \mu$ by Proposition 2.11. It means that the simple module $Ra_U s_U$ is not a direct summand of $Rs_T \cong \mathbb{C}(S_n/S_\lambda)$ for shape$(U) < \mu$ while $Ra_T s_T$ is obviously a direct summand. This means exactly that the above characterizing property is true for the irreducible characters χ. Therefore, $c_\lambda = \chi_\lambda$ for all $\lambda \vdash n$. ☐

We finished the proof of the theorem by this corollary. ☐

2.3 Hook formula

Question. *What are the dimensions of the irreducible representations of* S_n*?*

Remark 2.47. By character theory, one can know that we only have to determine $\chi_\lambda(1)$.

Remark 2.48.

$$p_\mu = \sum_{\nu\vdash n} c_\nu(\mu) S_\nu \quad / \cdot V = A_\delta$$

$$p_\mu \cdot V = \sum_{\nu \vdash n} c_\nu(\mu) A_{\nu+\delta}$$

For a fixed λ partition this has only one summand where the exponents are in a decreasing order with $\lambda+\delta$-type exponent and that is $x^{\lambda+\delta}$. Therefore, $c_\lambda(\mu)$ =the coefficient of $x^{\lambda+\delta}$ in $p_\mu \cdot V$.

By the above Remark and the Frobenius Theorem 2.41, the value of $\chi_\lambda(1)$ is the coefficient of $x^{\lambda+\delta}$ in

$$p_1^n \cdot V = \left(\sum_i^n x_i\right)^n \cdot V = \sum_{\nu\vdash n} \chi_\nu(1) \cdot A_{\nu+\delta}(x_1, \dots, x_n)$$

Let us denote $\lambda+\delta$ by $(l_1 > l_2 > \cdots > l_n)$. Then the previous formula gets the form:

$$\chi_\lambda(1) = \text{the coefficient of } \prod x_i^{l_i} \text{ in } \left(\sum x_i\right)^n \cdot \sum_{s\in S_n} \mathrm{sign}(s) x_i^{s(i)-1}$$

where the right hand side can be rearranged as

$$\sum_{k_1+\dots+k_n=n} \binom{n}{k_1,\dots,k_n} \cdot \prod_{i=1}^{n} x_i^{k_i} \cdot \sum_{s\in S_n} \operatorname{sign}(s) x_i^{s(i)-1}$$

so we can now easily collect the monomials that give $\prod x_i^{l_i}$. Therefore,

$$\chi_\lambda(1) = \sum_{s\in S_n;\ l_i-s(i)+1\geq 0} \operatorname{sign}(s) \frac{n!}{\prod_{i=1}^{n}(l_i - s(i)+1)!} = \frac{n!}{\prod_{i=1}^{n}(l_i!)} \sum_{s\in S_n} \operatorname{sign}(s) \prod_{i=1}^{n} \prod_{k=0}^{s(i)-2} (l_i - k)$$

where extending the indeces of the sum to all the elements of S_n will not cause any problem since then we multiply by zero as well in that summand. Now, we can express the sum as

$$\sum_{s\in S_n} \operatorname{sign}(s) \prod_{i=1}^{n} \prod_{k=0}^{s(i)-2} (l_i - k) = \begin{vmatrix} 1 & l_1 & l_1(l_1-1) & \dots & l_1(l_1-1)\dots(l_1-n+2) \\ 1 & l_2 & l_2(l_2-1) & \dots & l_2(l_2-1)\dots(l_2-n+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & l_n & l_n(l_n-1) & \dots & l_n(l_n-1)\dots(l_n-n+2) \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & l_1 & l_1^2 & \dots & l_1^{n-1} \\ 1 & l_2 & l_2^2 & \dots & l_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & l_n & l_n^2 & \dots & l_n^{n-1} \end{vmatrix} = \prod_{i<j} (l_i - l_j)$$

Corollary 2.49. *So we got a formula on the dimension:*

$$\chi_\lambda(1) = \frac{n!}{\prod_{i=1}^{n}(l_i!)} \cdot \prod_{1\leq i<j\leq n} (l_i - l_j)$$

After some simplification it has a more compact form in combinatorial terms:

Theorem 2.50. (Hook-formula)

$$\chi_\lambda(1) = \frac{n!}{\prod_{x\in\{boxes\ of\ T\}} h_x}$$

where h_x is the number of boxes that are to the right or below the box x including the box itself. This number h_x is called hook length. (E.g. In the diagram below a box of hook length 5 is shaded and each box is labeled with its hook length.)

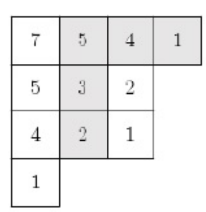


Illustration of hook length

Proof. Combinatorial computations. Details omitted. □

Remark 2.51. Let $F : \mathrm{Virt}(S_n) \to \Lambda_n$ be an isomorphism $\chi_\lambda \mapsto S_\lambda$ of abelian groups. However, $S_\lambda \in \Lambda_n \leq \Lambda$ (for the definition of Λ and Λ_n the space of symmetric functions, see the third book in the Literature) where the latter is a ring. So we can ask what is the pullback of this multiplication to $\mathrm{Virt}(S_n)$? It turns out that (without proof)

$$F\big(\mathrm{Ind}_{S_n \times S_m}^{S_{n+m}}(\chi_\lambda \otimes \chi_\mu)\big) = S_\lambda \cdot S_\mu \in \Lambda_{n+m}$$

Now this is again a Schur polynomial so it can be expressed by the higher degree Schur polynomials:

$$S_\lambda \cdot S_\mu = \sum_\nu c_{\lambda,\mu}(\nu) S_\nu$$

The appearing $c_{\lambda,\mu}$ coefficients are called the Littlewood-Richardson coefficients.

Fifth Lecture, 31th of October

3 The representation theory of $GL_n(\mathbb{C})$

3.1 Schur-Weyl duality

Let V be an n-dimensional vector space over $\mathbb{C}$ (but it works over $\mathbb{Q}$ too). Let us denote $T^d(V) = V^{\otimes d}$. Let

$$\rho : GL(V) \to GL(T^d(V))$$
$$g \mapsto g^{\otimes d}$$

Theorem 3.1. (The goal of this lecture.) *This representation is completely reducible.*

The symmetric group S_d is also naturally represented on $T^d(V)$ by

$$\sigma : S_d \to GL(T^d(V))$$

$$s \mapsto \big(v_1 \otimes \cdots \otimes v_d \mapsto v_{s^{-1}(1)} \otimes \cdots \otimes v_{s^{-1}(d)}\big)$$

One can easily check that it is a representation. Also, it is easy to see that it commutes with the previous action of $GL(V)$. In other words, the two action centralize each other on $GL(T^d(V))$.

In fact, "they are each other's centralizer". To make it precise, define

$$U := \mathrm{Span}_\mathbb{C}\{\rho(g) \mid g \in GL(V)\} \overset{\text{subalgebra}}{\subseteq} \mathrm{End}_\mathbb{C} T^d(V) \supseteq GL(T^d(V))$$

(This U is called the Schur algebra. In the complex case it will be semisimple, but over other fields it has more interesting structure.) Similarly, consider $\sigma(\mathbb{C}S_d) \subseteq \mathrm{End}_\mathbb{C}(T^d(V))$.

Theorem 3.2. *U and $\sigma(\mathbb{C}S_d)$ are each other's centralizer in the algebra $\mathrm{End}_\mathbb{C} T^d(V)$.*

Lemma 3.3. *Let W be a finite dimensional $\mathbb{C}$-vector space. Write $T^d(W)^{S_d}$ for the fixed point subspace of S_d. Then $T^d(W)^{S_d} = \mathrm{Span}_\mathbb{C}\{v \otimes \cdots \otimes v \mid v \in W\}$.*

Proof. It is clear that

$$T^d(W)^{S_d} = \Big\{ \sum_{s \in S_d} v_{s^{-1}(1)} \otimes \cdots \otimes v_{s^{-1}(d)} \mid v_1, \ldots, v_d \in V \Big\}$$

because $\supseteq$ is trivial and $\subseteq$ follows from the existence of the projection $\sigma\big(\frac{1}{d!}\sum_{s \in S_d} s\big)$ of $T^d(W)$ onto $T^d(W)^{S_d}$. Now, we prove that the right hand side is $\mathrm{Span}_\mathbb{C}\{v \otimes \cdots \otimes v \mid v \in W\}$.

Take $v_1, \ldots, v_d \in V$ and $\lambda_1, \ldots, \lambda_d \in \mathbb{C}$. Then $v^{\otimes d} := (\lambda_1 v_1 + \cdots + \lambda_d v_d)^{\otimes d}$ is in

$$\Big\{ \sum_{s \in S_d} v_{s^{-1}(1)} \otimes \cdots \otimes v_{s^{-1}(d)} \mid v_1, \ldots, v_d \in V \Big\}$$

because if we consider it as a polynomial in $\lambda_1, \dots, \lambda_d$ then every coefficient is a linear combination of elements of the above symmetrized forms. Conversely, this linear expression of $v^{\otimes d}$ by symmetrized elements can be "inverted" i.e. we can express the symmetrized elements by $v^{\otimes d}$'s. To get the "invertibility" we implicitly use the following lemma

Lemma 3.4. *Let $f_1, \dots, f_N$ be any function from $Y \to \mathbb{C}$ such that $f_1, \dots, f_N$ are linearly independent as functions. Then there exist $y_1, \dots, y_N \in Y$ such that $\big(f_i(y_j)\big)_{i \le N}$, $j = 1, \dots, N$ are linearly independent vectors. (The proof is exercise.)*

Now, we can apply it on the functions $\mathbb{C}^N \to \mathbb{C}$, $(\lambda_1, \dots, \lambda_d) \mapsto \prod_{i=1}^d \lambda_i^{\alpha_i}$ where $\alpha_i \in \mathbb{N}$ and $\sum_i \alpha_i = d$. $\square$

Proposition 3.5. $C_{\mathrm{End}_{\mathbb{C}} T^d(V)}(\sigma(\mathbb{C}S_d)) \overset{def}{=} \mathrm{End}_{\mathbb{C}S_d}(T^d(V)) = U$

Claim 3.6. $\mathrm{End}_{\mathbb{C}} T^d(V) \cong T^d(\mathrm{End}_{\mathbb{C}}(V))$

Proof. Both have n^{2d} dimensions as a vector space and we have a natural map

$$\kappa : T^d(\mathrm{End}_{\mathbb{C}}(V)) \to \mathrm{End}_{\mathbb{C}} T^d(V)$$

$$L_1 \otimes \cdots \otimes L_d \mapsto \big(v_1 \otimes \cdots \otimes v_d \mapsto L_1 v_1 \otimes \cdots \otimes L_d v_d\big)$$

It is clear that it is a $\mathbb{C}$-linear algebra homomorphism. So if we prove that κ is injective then it have to be an isomorphism by knowing the dimensions. And it is, because it brings the basis $E_{i_1 j_1} \otimes \cdots \otimes E_{i_d j_d}$ into independent linear transformations. In another way, $\mathrm{End}_{\mathbb{C}}(V) \cong V \otimes V^*$ as algebras with the product $V \otimes V^* \otimes V \otimes V^* \to V \otimes V^*$; $v \otimes \varphi \otimes w \otimes \psi \mapsto \varphi(w) \cdot v \otimes \psi$. So

$$T^d\big(\mathrm{End}_{\mathbb{C}}(V)\big) \cong \big(V^* \otimes V\big)^{\otimes d} \cong (V^{\otimes d})^* \otimes V^{\otimes d} \cong \mathrm{End}_{\mathbb{C}}(T^d(V))$$

$\square$

Now, let us introduce

$$\hat{\sigma} : S_d \to GL(T^d(\mathrm{End}_{\mathbb{C}} V))$$

$$s \mapsto (L_1 \otimes \cdots \otimes L_d \mapsto L_{s^{-1}(1)} \otimes \cdots \otimes L_{s^{-1}(d)})$$

This gives a representation of S_d on $T^d(\mathrm{End}_{\mathbb{C}} V)$. In fact, this is the product representation $\sigma \cdot \sigma^*$ using the identification κ.

Claim 3.7. Under the isomorphism $\kappa : T^d(\mathrm{End}_{\mathbb{C}}(V)) \to \mathrm{End}_{\mathbb{C}} T^d(V)$ the subspace $T^d(\mathrm{End}_{\mathbb{C}} V)^{S_n}$ is mapped onto $\mathrm{End}_{\mathbb{C}S_d}(T^d(V))$.

Proof. Indeed, if

$$\hat{\sigma}(s)(L_1 \otimes \cdots \otimes L_d)(v_1 \otimes \cdots \otimes v_d) = \big(L_{s^{-1}(1)} \otimes \cdots \otimes L_{s^{-1}(d)}\big)(v_1 \otimes \cdots \otimes v_d) = (L_1 \otimes \cdots \otimes L_d)(v_1 \otimes \cdots \otimes v_d)$$

then applying κ gives

$$L_{s^{-1}(1)}(v_1) \otimes \cdots \otimes L_{s^{-1}(d)}(v_d) = L_1(v_1) \otimes \cdots \otimes L_d(v_d)$$

where

$$\kappa\big(\hat{\sigma}(s)(L_1 \otimes \dots L_d)\big)(v_1 \otimes \cdots \otimes v_d) = L_{s^{-1}(1)}(v_1) \otimes \cdots \otimes L_{s^{-1}(d)}(v_d) =$$

$$= \sigma(s)\big(L_1(v_{s^{-1}(1)}) \otimes \cdots \otimes L_d(v_{s^{-1}(d)})\big) = \sigma(s)^{-1} \kappa\big(L_1 \otimes \cdots \otimes L_d\big)\sigma(s)(v_1 \otimes \cdots \otimes v_d)$$

κ is an isomorphism so we got exactly the claim because for a fixed $s \in S_d$ and $L \in T^d(\mathrm{End}_{\mathbb{C}}(V))$ we got $\hat{\sigma}(s)L = L$ if and only if $\sigma(s) \circ \kappa(L) = \kappa(L) \circ \sigma(s)$. $\square$

Remark 3.8. In another way, one can notice that $\mathrm{End}_{\mathbb{C}S_n}(T^d(V)) = \mathrm{End}_{\mathbb{C}}(T^d(V))^{S_n}$ by the definitions, if we consider the action $\tau : S_n \to GL(\mathrm{End}_{\mathbb{C}}(T^d(V)))$, $g(\varphi) = \sigma(g) \circ \varphi \circ \sigma(g^{-1})$. So the claim states that κ identifies $T^d(\mathrm{End}_{\mathbb{C}} V)^{S_n}$ with $\mathrm{End}_{\mathbb{C}}(T^d(V))^{S_n}$ what is a consequence of that the isomorphism $T^d\big(\mathrm{End}_{\mathbb{C}}(V)\big) \overset{\kappa}{\cong} \mathrm{End}_{\mathbb{C}}(T^d(V))$ is an isomorphism of S_n-modules since $\hat{\sigma} = \sigma \cdot \sigma^* = \kappa^{-1}(\tau)$.

Proof. (of Proposition 3.5) To get the statement, we can apply the previous claim:

$$\mathrm{End}_{\mathbb{C}S_d}(T^d(V)) \overset{\kappa}{\to} T^d\big(\mathrm{End}_{\mathbb{C}}(V)\big)^{S_n} \overset{\text{Lemma}}{=} \mathrm{Span}_{\mathbb{C}}\{T^d(g) \mid g \in \mathrm{End}_{\mathbb{C}}(V)\} = \mathrm{Span}_{\mathbb{C}}\{T^d(g) \mid g \in GL(V)\} = U$$

□

To prove the other direction of Theorem 3.2, we have to use some general statements about finite dimensional semisimple $\mathbb{C}$-algebras.

Proposition 3.9. *Suppose S is a finite dimensional simple A-module where A is any $\mathbb{C}$-algebra. Consider the A-module $S^m = S \oplus \cdots \oplus S$. Then $\mathrm{End}_A(S \oplus \cdots \oplus S) \cong \mathbb{C}^{m\times m}$. (See Topics in algebra homework.)*

By the notation $\mathbb{C}^{m\times m} \cong B := \mathrm{End}_A(S^m) \cong \mathrm{End}_A(S \otimes \mathbb{C}^m)$ we get that $\mathbb{C}^m$ is a B-module in the natural way. Now, let A be a finite dimensional semisimple $\mathbb{C}$-algebra: $A = A_1 \times \cdots \times A_q$ where – by Wedderburn-Artin, we know that – $A_i \cong \mathbb{C}^{n_i\times n_i}$. Then $S_i = \mathbb{C}^{n_i}$ are simple A_i-modules.

Let M be a finite dimensional A-module. Then $M \cong \oplus_{i=1}^q S_i^{m_i} = \oplus_{i=1}^q S_i \otimes T_i$ where $T_i = \mathbb{C}^{m_i}$ which is naturally a $B_i = \mathrm{End}_A(S_i^{m_i}) \cong \mathbb{C}^{m_i \otimes m_i}$-module. Then

$$\mathrm{End}_A(M) = \mathrm{End}_A(\oplus_{i=1}^q S_i^{m_i}) \cong \oplus_{i=1}^q \mathrm{End}_A(S_i^{m_i}) = B_1 \times \cdots \times B_q$$

Interchanging the roles of A and B we see that $A = \mathrm{End}_B(M)$ automatically. So a simple A-module gives a symmetric relation between A and the endomorphism algebra. In fancy, snobbish terminology, $\oplus_{i=1}^q \mathbb{C}^{n_i} \otimes \mathbb{C}^{m_i}$ is a balanced $A := \oplus_{i=1}^q \mathbb{C}^{n_i\times n_i}$-module.

Proof. (of Theorem 3.2) By Proposition 3.5 we got the first direction of the statement. Conversely, $\mathbb{C}S_d$ is semisimple so $\sigma(\mathbb{C}S_d)$ is semisimple too and we can apply the previous observations resulting that U is semisimple as well and they are each other's centralizer. □

Question. *What is the direct decomposition of $T^d(V)$ as a $GL(V)$-module? The previous general discussion answers this in principle but first, we have to understand $\sigma(\mathbb{C}S_d)$, or equivalently, we have to understand* $\mathrm{Ker}\sigma$*.*

We know that $\mathbb{C}S_d = \oplus_{\lambda\vdash d} I_\lambda$ where $I_\lambda \cong M_{n_\lambda}(\mathbb{C})$ and n_λ is determined by the Hook-formula. So to find $\mathrm{Ker}\sigma$ we just have to evaluate σ on the idempotents, the Young symmetrizers.

Proposition 3.10. *Let $n = \dim V$. Then*

$$\mathrm{Ker}\,\sigma = \sum_{\lambda\vdash d \ \ \mathrm{ht}(\lambda)\geq n+1} I_\lambda$$

Proof. Choose a basis $e_1, \ldots, e_n$. If $d \geq n+1$ then $S_{n+1} \leq S_d$ naturally, as the stabilizer of $n+2, n+3, \ldots$.

$$\Big(\sum_{s\in S_{n+1}} (-1)^s s\Big)(e_{i_1} \otimes \cdots \otimes e_{i_d}) = \sum_{s\in S_n} (-1)^s e_{i_{s^{-1}(1)}} \otimes \cdots \otimes e_{i_{s^{-1}(d)}} =$$

$$= \Big(\sum_{s\in S_n} (-1)^s e_{i_{s(1)}} \otimes \cdots \otimes e_{i_{s(n+1)}}\Big) \otimes e_{i_{n+2}} \otimes \cdots \otimes e_{i_d}$$

However, the sum is already zero because for some k and l we get that $e_{i_k} = e_{i_l}$ because the dimension is smaller so we get every term with zero coefficient. So the $f = \sum_{s\in S_{n+1}} (-1)^s s$ is in the kernel. To prove that all ideals I_λ such that $\mathrm{ht}(\lambda) \geq n+1$ are in the kernel, it is enough to find a single element in each ideal that is in the kernel by irreducibility. A tableau of $\mathrm{ht}(\lambda) \geq n+1$ with a numbering where the first column contains $\{1, 2, \ldots, n+1\}$ gives an element $a_T s_T \in I_\lambda$ such that $f a_T s_T = (n+1)! a_T s_T$ because $(-1)^s s a_T = a_T$ if $s \in C_T$ and f is the "signed sum" of such elements. Therefore, $0 \neq f a_T s_T \in \operatorname{Ker} \sigma$ and so

$$\operatorname{Ker} \sigma \supseteq \sum_{\lambda \vdash d \ \mathrm{ht}(\lambda) \geq n+1} I_\lambda$$

Conversely, suppose $\mathrm{ht}(\lambda) \leq n$. Now, it is enough to find a single element in I_λ that is not in the kernel of σ by irreducibility. Take the standard ordered numbering of T (i.e. the first row is $1, 2, \ldots$). This T gives an element $a_T s_T \in I_\lambda$. Then consider the following element of $T^d(V)$:

$$e_1 \otimes \cdots \otimes e_1 \otimes e_2 \otimes \cdots \otimes e_n$$

where we took λ_1-times e_1, λ_2-times e_2 and so on. Then we have

$$\sigma(a_T s_T)(e_1 \otimes \cdots \otimes e_1 \otimes e_2 \otimes \cdots \otimes e_n) = |C_T| \cdot \sigma(a_T)(e_1 \otimes \cdots \otimes e_1 \otimes e_2 \otimes \cdots \otimes e_n) =$$

because every row-stabilizer acts trivially on this tensor. Continuing this computation gives

$$= |C_T| \cdot (e_1 \otimes \cdots \otimes e_1 \otimes e_2 \otimes \cdots \otimes e_n) + \text{other permutations}$$

because if we take a nontrivial column stabilizer element then we cannot get the same combination. Therefore, it is non-zero so we got the statement. ☐

Corollary 3.11. *The direct decomposition of $T^d(V)$ as a $GL(V)$-module can be determined.*

Sixth Lecture, 7th of November

Corollary 3.12. (One form of Schur-Weyl duality) *By Proposition 3.10, as an $S_d \times GL(V)$-module, $T^d(V)$ decomposes as*

$$\bigoplus_{\lambda \vdash d,\, \mathrm{ht}(\lambda) \leq n} V_\lambda \otimes W_\lambda$$

where V_λ is the irreducible S_d-module labeled by λ and W_λ is an irreducible $GL(V)$-module (where $W_\lambda \ncong W_\mu$ for $\lambda \neq \mu$).

Question. *What is W_λ more explicitly?*

Take a fixed tableau T of shape λ. Then the V_λ's are isomorphic to $\mathbb{C}S_d \cdot c_T$ as S_d-modules. By that, one can guess the following proposition:

Proposition 3.13. $W_\lambda \cong c_T \cdot T^d(V)$

Proof. It is clear that $c_T \cdot T^d(V)$ is a $GL(V)$-module because it is a vector space and by the Double Centralizing Theorem 3.2: for $g \in GL(V)$ we have

$$g c_T T^d(V) = c_T g T^d(V) \subseteq c_T T^d(V)$$

So

$$c_T T^d(V) = c_T \bigoplus_{\mu \vdash d,\, \mathrm{ht}(\mu) \leq n} V_\mu \otimes W_\mu = \bigoplus_{\mu \vdash d,\, \mathrm{ht}(\mu) \leq n} c_T V_\mu \otimes W_\mu =$$

where c_T is the (scalar multiple of the) idempotent corresponding to V_λ so almost all the terms will give zero. And if it is non-zero, it has to be one-dimensional by Proposition 2.11

$$= c_T V_\lambda \otimes W_\lambda \cong \mathbb{C} \otimes W_\lambda \cong W_\lambda$$

So we got the statement. ☐

Example 3.14. Let $\lambda = (d)$ the "1-row" partition and let T be a Young tableau of shape λ with ordered numbering. Then

$$c_T = \sum_{s \in R_T = S_d} s$$

so $c_T(v_1 \otimes \cdots \otimes v_d) = \sum_{s \in S_d} v_{s(1)} \otimes \cdots \otimes v_{s(d)}$ and

$$W_{(d)} = \mathrm{Span}_{\mathbb{C}}\Big\{ \sum v_{s(1)} \otimes \cdots \otimes v_{s(d)} \Big\} = \text{subspace of symmetric tensors}$$

The dimension of this is exactly $\binom{n+d-1}{n-1}$ where $\dim V = n$ because the symmetrization of $(e_1 \otimes e_1 \otimes \cdots \otimes e_1) \otimes (e_2 \otimes \cdots \otimes e_2) \otimes \cdots \otimes (e_n \otimes \cdots \otimes e_n)$ is a basis in it if you take this vector for every n-partition of d.

Example 3.15. Let $\lambda = (1, 1, 1, \ldots, 1) = (1^d)$, the "1-column" partition and let T be a Young tableau of shape λ with ordered numbering. Then

$$c_T = \sum_{s \in R_T = S_d} (-1)^s s$$

so the generated subspace

$$c_T T^d(V) = \mathrm{Span}_{\mathbb{C}}\{ \sum_{s \in S_d} (-1)^s v_{s(1)} \otimes \cdots \otimes v_{s(d)} \mid v_1, \ldots, v_n \in V\}$$

the subspace of alternating tensors.

3.2 Schur functors

Take a λ. Fix a tableau T of shape λ. Define the endofunctor S_λ on the category of $\mathbb{C}$-vectorspaces: for a vector space V:

$$S_\lambda(V) = c_T T^d(V)$$

and for a linear map $L : V \to W$

$$S_\lambda(L) : S_\lambda(V) \to S_\lambda(W)$$

where $S_\lambda(L)(v_1 \otimes \cdots \otimes v_d) = T^d(L)|_{c_T T^d(V)}(v_1 \otimes \cdots \otimes v_d) = (Lv_1 \otimes \cdots \otimes Lv_d)$ which in fact gives an $S_\lambda(W)$-element because the action of S_n can be interchanged with $T^d(L)$.

The $S_\lambda(V)$'s are exactly the previous W_λ's, by definition. One can also see the functor S_λ is a representation of $GL(V)$ since it induces a map $GL(V) \to GL(S_\lambda(V))$.

Question. *What are the characters of these W_λ?*

Let $\Psi : GL_n(\mathbb{C}) \cong GL(V) \to GL(W)$ be a morphism where V will be fixed and W is an arbitrary finite dimensional $\mathbb{C}$-vector space. Now, we can define a corresponding central function on $GL(V)$ by:

$$GL(V) \to \mathbb{C}$$

$$g \mapsto \mathrm{Tr}(\Psi(g))$$

If we assume that $M(\Psi)$ contains only continuous functions (or in other terminology, Ψ is continuous) then it is a continuous representations.

Note that diagonalizable elements U in $GL_n(\mathbb{C})$ constitute a (Zariski-)dense subset. Indeed, $A \in GL_n(\mathbb{C})$ has distinct eigenvalues if and only if $\mathrm{Discriminant}(\det(xI - A)) \neq 0$. This means it is enough to know $\mathrm{Tr} \circ \Psi$ on this dense subset because "taking the trace" is a continuous function.

By the definition of U for all $g \in U$ there exists an $h \in GL_n(\mathbb{C})$ such that hgh^{-1} is diagonal (not only diagonalizable). If we apply $\mathrm{Tr} \circ \Psi$ on these:

$$\mathrm{Tr}(\Psi(g)) = \mathrm{Tr}(\Psi(h)\Psi(g)\Psi(h^{-1})) = \mathrm{Tr}(\Psi(hgh^{-1}))$$

so it is enough to know $\mathrm{Tr} \circ \Psi$ on the diagonal elements. This observation gives rise to the following definition:

Definition 3.16. The character ch_ρ of the representation Ψ is

$$\mathrm{ch}_\rho = \mathrm{Tr} \circ \Psi|_{D_n}$$

where $D_n = \{\mathrm{diag}(z_1, \dots, z_n \mid z_i \in \mathbb{C}^\times)\}$.

Remark 3.17. The above investigation shows that the restriction map (which is a $\mathbb{C}$-algebra homomorphism!):

$$\text{Continuous central functions on } GL_n(\mathbb{C}) \quad \longrightarrow \quad \text{Symmetric functions on } D_n$$

is an injection (where the symmetry of the resulting function came from the fact that two diagonal matrices are conjugate if and only if their diagonal elements are the same up to permutation).

Proposition 3.18. *Let $\rho : GL_n(\mathbb{C}) \to GL(T^d(V))$ then $M(\rho)$ is the set of degree d symmetric polynomial functions on $GL_n(\mathbb{C}) \subseteq \mathbb{C}^{n\times n}$ (i.e. polynomial functions in the coordinates).*

Proof. Let $e_1, \dots, e_n$ be a basis in V so $g \cdot e_j = \sum_i g_{ij} e_i$. So

$$\rho(g)(e_{j_1} \otimes \cdots \otimes e_{j_d}) = ge_{j_1} \otimes \cdots \otimes ge_{j_d} = \Big(\sum_{i_1=1}^{n} g_{i_1 j_1} e_{i_1}\Big) \otimes \cdots \otimes \Big(\sum_{i_d=1}^{n} g_{i_d j_d} e_{i_d}\Big) = \sum_{i_1,\dots,i_d} \prod_{k=1}^{d} g_{i_k j_k} e_{i_1} \otimes \cdots \otimes e_{i_d}$$

so with respect to the standard basis of $T^d(V)$ the matrix elements of ρ are exactly the length d products of the coordinate functions on $\mathbb{C}^{n\times n}$. $\square$

Theorem 3.19. $\mathrm{Tr}(S_\lambda(\mathrm{diag}(z_1, \dots, z_n)) = S_\lambda(z_1, \dots, z_n)$ *where on the right hand side S_λ stands for the Schur polynomials while on the left hand side it is the Schur functor. (Having different arguments saves us from confusing these two.)*

Reminder: $S_\lambda(\mathrm{diag}(z_1, \dots, z_n)) = \rho(\mathrm{diag}(z_1, \dots, z_n))|_{S_\lambda(V) = c_T T^d(V)}$ by definition, where $\rho : GL(V) \to GL(T^d(V))$.

Proposition 3.20. *For $(s, X) \in S_d \times GL(V)$,*

$$\mathrm{Tr}_{T^d(V)}((\sigma \otimes \rho)(s, X)) = \prod_{i=1}^{n} \mathrm{Tr}(X^{\mu_i})$$

where the cycle length of s are $\mu_1, \dots, \mu_n$.

Proof. This will follow from the following:

Claim 3.21. For arbitrary $X_1, \dots, X_d \in \mathrm{End}_{\mathbb{C}}(V)$

$$\mathrm{Tr}_{T^d(V)}(\sigma(s) \circ (X_1 \otimes \cdots \otimes X_d)) = \prod_{j=1}^{k} \mathrm{Tr}_{c_j}(X_1, \dots, X_d)$$

where $s = c_1 \dots c_k$ as a cycle decomposition and Tr_{c_j} is defined as $\mathrm{Tr}_{c_j} = \mathrm{Tr}(X_{i_1} X_{i_2} \dots X_{i_k})$ if $c_j = (i_1 i_2 \dots i_k)$ as a cycle.

Proof. Let us use the identification $\mathrm{End}_{\mathbb{C}}(V) \cong V \otimes V^*$. By the multilinearity of both sides, it is enough to verify the formula where X_i's are rank 1 transformations, i.e. $X_i = u_i \otimes \Psi_i$. So

$$\sigma(s) \circ (X_1 \otimes \cdots \otimes X_d)(v_1 \otimes \cdots \otimes v_d) = \sigma(s)\Big(\big(\prod_{i=1}^{d} \psi_i(v_i)\big)(u_1 \otimes \cdots \otimes u_d)\Big) =$$

$$= \prod_{i=1}^{n} \psi_i(v_i)(u_{s^{-1}(1)} \otimes \cdots \otimes u_{s^{-1}(d)})$$

This computation show that $\sigma(s)\circ X_1\otimes\cdots\otimes X_d=\Big(u_{s^{-1}(1)}\otimes\Psi_1\Big)\otimes\cdots\otimes\Big(u_{s^{-1}(d)}\otimes\Psi_d\Big)$. Hence,

$$\mathrm{Tr}(\sigma(s)\circ(X_1\otimes\cdots\otimes X_d))=\mathrm{Tr}\Big(\Big(u_{s^{-1}(1)}\otimes\Psi_1\Big)\otimes\cdots\otimes\Big(u_{s^{-1}(d)}\otimes\Psi_d\Big)\Big)=\prod_{i=1}^d\Psi_i\big(u_{s^{-1}(i)}\big)=\prod_{i=1}^d\Psi_{s(i)}\big(u_i\big)$$

where we used the fact that $\mathrm{Tr}(u\otimes\Psi)=\Psi(u)$. So we only have to prove that the right hand side is the same for rank one transformations:

$$\mathrm{Tr}_c\Big(\big(u_1\otimes\Psi_1\big),\dots,\big(u_d\otimes\Psi_d\big)\Big)=\mathrm{Tr}_V\Big(\big(u_{i_1}\otimes\Psi_{i_1}\big)\cdot\dots\cdot\big(u_{i_k}\otimes\Psi_{i_k}\big)\Big)=$$

$$=\Psi_{i_2}(u_{i_1})\Psi_{i_3}(u_{i_2})\dots\Psi_{i_k}(u_{i_1})=\prod_{j=1}^k\Psi_{s(i_j)}(u_{i_j})$$

where we used the fact that the multiplication of $V\otimes V^*$ induced by the matrix-multiplication on $\mathrm{End}(V)$ is exactly $V\otimes V^*\otimes V\otimes V^*\to V\otimes V^*$; $u\otimes\phi\otimes v\otimes\psi\mapsto\psi(v)\cdot(u\otimes\psi)$. So the two sides equal. □

The Proposition follows by choosing $X_i=X$ for all i. □

Seventh Lecture, 14th of November

Proof. of Theorem 3.19: Because of

$$T^d(V)\cong\bigoplus_{\lambda\vdash d;\,\mathrm{ht}(\lambda)\le n}V_\lambda\otimes S_\lambda(V)\qquad\text{as } S_d\times GL(V)\text{-modules}$$

if we take an $s\in S_d$ and $\mathrm{diag}(x_1,\dots,x_n)=X\in GL(V)\cong GL_n(\mathbb{C})$ (we have proved that it is enough to consider diagonal matrices) then we can turn to their traces:

$$\mathrm{Tr}_{T^d(V)}(s,X)=\sum_{\lambda\vdash d;\,\mathrm{ht}(\lambda)\le n}\mathrm{Tr}_{V_\lambda}(s)\cdot\mathrm{Tr}_{S_\lambda(V)}(X)=$$

$$=\sum_{\lambda\vdash d;\,\mathrm{ht}(\lambda)\le n}\chi_\lambda(\mu)\cdot\mathrm{Tr}_{S_\lambda(V)}(X)$$

where χ_λ is the invertible character of S_d of type λ and μ is the cycle type of s. Now, on the left hand side we have

$$\mathrm{Tr}_{T^d(V)}(s,X)=\prod_i\mathrm{Tr}(X^{\mu_i})$$

by Proposition 3.20 if s has cycle type $\mu=(\mu_1,\mu_2,\dots)$. Here, X^{μ_i} has simple structure, since

$$X^{\mu_i}=\begin{pmatrix}x_1^{\mu_i}&0&0\\0&\ddots&0\\0&0&x_n^{\mu_i}\end{pmatrix}$$

so $\mathrm{Tr}(X^{\mu_i})=\sum_{j=1}^n x_j^{\mu_i}$, and $\prod_i\mathrm{Tr}(X^{\mu_i})=\prod P_{\mu_i}(x_1,\dots,x_n)$. Taking all together, we got

$$P_\mu(x_1,\dots,x_n)=\sum_{\lambda\vdash d;\,\mathrm{ht}(\lambda)\le n}\chi_\lambda(\mu)\mathrm{Tr}_{S_\lambda(V)}(X)$$

for all $\mu\vdash d$.

On the other hand, by the Frobenius character formula (Theorem 2.41) the same equality holds with the Schur-polynomials, i.e.

$$P_\mu(x_1,\dots,x_n)=\sum_{\lambda\vdash d;\,\mathrm{ht}(\lambda)\le n}\chi_\lambda(\mu)S_\lambda(x_1,\dots,x_n)$$

holds for all $\mu \vdash d$ partition. So it holds for all λ with height at most n and for all $\mu_1 \leq n$. If we can prove that the resulting square-matrix of the coefficient (by the transposition of the Young diagrams one can see that it is indeed a square-matrix)

$$\Big(\big(\chi_\lambda(\mu)\big)\Big)_{\lambda \vdash d,\, \mathrm{ht}(\lambda) \leq n}^{\mu \vdash d,\, \mu_1 \leq n}$$

is invertible and then $S_\lambda(x_1, \ldots, x_n) = \mathrm{Tr}_{S_\lambda(V)}(X)$ follows.

To prove the invertibility, notice that the degree d power-sums $P_d(x_1, \ldots, x_n)$ are a $\mathbb{Q}$-algebra generating system of ${}_{\mathbb{Q}}\Lambda_d^n$ by the Newton-Waring formulas, so $P_\mu(x_1, \ldots, x_n)$'s are a $\mathbb{Q}$-basis in ${}_{\mathbb{Q}}\Lambda_d^n$ (see Remark 2.36). Therefore, – because the number of Schur-polynomials in degree n and the number of P_μ's are the same – $P_\mu(x_1, \ldots, x_n)$'s must constitute a basis as well, as the Schur polynomials do. Hence, the above matrix $\big((\chi_\lambda(\mu))\big)$ cannot be singular i.e. it is invertible. □

Remark 3.22. In the case of S_d we used the explicit character formula to compute the dimensions of the irreducible representations (see Hook-formula, Theorem 2.50). Now, we do the same but for $GL(V)$-modules.

Corollary 3.23.

$$\dim_{\mathbb{C}} S_\lambda(V) = \frac{\prod_{1 \leq i < j \leq n}(\lambda_i - \lambda_j + j - 1)}{\prod_{1 \leq i < j \leq n}(j - i)}$$

where $\lambda = (\lambda_1, \ldots, \lambda_n)$*,* $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$*.*

Proof. We just need to substitute 1 in the Schur-polynomials. However, there is a problem with it: in the definition

$$S_\lambda(x_1, \ldots, x_n) = \frac{\det\big((x_j^{\lambda_i + n - i})\big)_{i,j=1}^n}{\det\big((x_j^{n-i})\big)_{i,j=1}^n}$$

both the numerator and the denominator vanishes in 1. The trick is to instead substitute $(1, q, q^2, \ldots, q^{n-1})$ into the variables:

$$S_\lambda(1, q, q^2, \ldots, q^{n-1}) = \frac{\det\big(((q^{j-1})^{\lambda_i + n - i})\big)_{i,j=1}^n}{\det\big(((q^{j-1})^{n-i})\big)_{i,j=1}^n} =$$

$$= \frac{\det\big(((q^{\lambda_i + n - i})^{j-1})\big)_{i,j=1}^n}{\det\big(((q^{n-i})^{j-1})\big)_{i,j=1}^n} =$$

so both become Vandermonde matrices for which we know the value of their determinant. Therefore,

$$S_\lambda(1, q, q^2, \ldots, q^{n-1}) = \frac{\prod_{1 \leq k < l \leq n}(q^{\lambda_k + n - k} - q^{\lambda_l + n - l})}{\prod_{1 \leq k < l \leq n}(q^{n-k} - q^{n-l})} = \frac{\prod_{1 \leq k < l \leq n}(q^{(\lambda_k + n - k) - (\lambda_l + n - l)} - 1)q^{\lambda_l + n - l}}{\prod_{1 \leq k < l \leq n}(q^{(n-k)-(n-l)} - 1)q^{n-l}}$$

where we can cancel out the $q - 1$ terms. By

$$\left.\frac{q^a - q^b}{q - 1}\right|_{q=1} = a - b$$

we get the stated formula, i.e.

$$S_\lambda(1, 1, \ldots, 1) = \frac{\prod_{1 \leq i < j \leq n}(\lambda_i - \lambda_j + j - 1)}{\prod_{1 \leq i < j \leq n}(j - i)}$$

so the proof is complete. □

Goal: Let us understand the Cauchy formula in the language of GL-characters.

Proposition 3.24. *Straightforward generalization of the Cauchy-formula:*

$$\prod_{i=1}^{n}\prod_{j=1}^{m}\frac{1}{(1-x_iy_jz)}=\sum_{d=0}^{\infty}z^d\sum_{\lambda\vdash d}S_\lambda(x_1,\dots,x_n)S_\lambda(y_1,\dots,y_m)$$

Proof. Substitute $x_i'=x_iz$ in the original formula. ☐

Corollary 3.25. *The representation theoretic meaning of the formula is: let $V:=\mathbb{C}^n$ and $W:=\mathbb{C}^m$ as $GL(V)$ and $GL(W)$-modules. Then the irreducible decomposition of the symmetric tensor-power of $V\otimes W$ as a $GL(V)\times GL(W)$-module is:*

$$S^d(V\otimes W)\cong\bigoplus_{\lambda\vdash d}\Big(S_\lambda(V)\otimes S_\lambda(W)\Big)$$

where S_λ is the Schur functor and $S^d=S_{(d)}$ is the special case of the Schur functor, the functor of symmetric polynomials.

One can notice that this is a multiplicity-free $GL(V)\times GL(W)$-module.

Proof. We need to show that the character of $S^d(V\otimes W)$ as a $GL(V)\times GL(W)$-module is the z-degree d part of the formula on the left hand side. For this, take a basis $e_1,\dots,e_n$ of V and a basis $f_1,\dots,f_m$ of W. Let us introduce the notation: $h_{ij}=e_i\otimes f_j\in V\otimes W$. The basis of $S^d(V\otimes W)$ is all the possible products

$$\prod_{i=1,\dots,n;\ j=1,\dots,m}h_{ij}^{\alpha_{ij}}\quad\text{with the assumption }\sum_{i,j}\alpha_{ij}=d$$

i.e. all the commuting monomials of degree d. Now, we can take the matrices $X=\mathrm{diag}(x_1,\dots,x_n)$ and $Y=\mathrm{diag}(y_1,\dots,y_m)$. Then

$$(X,Y)h_{ij}=(Xe_i,Yf_j)=(x_ie_i,y_jf_j)=x_iy_jh_{ij}$$

so the action of (X,Y) on an arbitrary degree d element is

$$(X,Y)\prod_{i,j}h_{ij}^{\alpha_{ij}}=\prod_{i,j}(x_iy_jh_{ij})^{\alpha_{i,j}}=\Big(\prod_{i,j}(x_iy_j)^{\alpha_{ij}}\Big)\cdot\Big(\prod_{i,j}h_{ij}^{\alpha_{ij}}\Big)\qquad\text{so}$$

$$\mathrm{Tr}_{S^d(V\otimes W)}(X,Y)=\sum_{\alpha=((\alpha_{ij}));\ \sum_{i,j}\alpha_{ij}=d}\ \prod_{i,j}(x_iy_j)^{\alpha_{ij}}=\text{coefficient of }z^d\text{ in }\prod_{i=1}^{n}\prod_{j=1}^{m}\frac{1}{(1-x_iy_jz)}$$

and we got that the two representations in the statement have the same character by the Cauchy formula. If we could prove that these representations are completely reducible and that complete reducibility implies that only isomorphic representations have the same character then we would be done. The former statement is not proven yet, see Corollary 3.30. However, the latter statement is already known to be true: First, recall that in Topics in Algebra course we have proven that for a finite dimensional irreducible representation ρ we have $M(\rho)\cap\mathrm{Cent}(G,\mathbb{C})=\mathbb{C}\mathrm{ch}_\rho$ and the $M(\rho)$'s are independent for distinct irreducible ρ's (and the proof remained valid for not necessarily finite groups). Therefore, if $\mathrm{ch}_\varphi=\mathrm{ch}_\psi$ for two completely reducible representations φ and ψ then $M(\varphi)=M(\psi)\subseteq\mathrm{Fun}(G,\mathbb{C})$ what means $\varphi\cong\psi$ by $M(\varphi)\cong_{G\times G}\varphi^*\otimes\varphi\cong_{1\times G}\varphi+\varphi+\dots+\varphi$.

The representation on the right is completely reducible because it is given as a direct sum of irreducibles. However, we do not know that the one on the left hand side is completely reducible. (But it is also true.) ☐

Remark 3.26. This representation is important in algebraic geometry because we have a commonly used algebra

$$S(V\otimes W)=\bigoplus_{d=0}^{\infty}S^d(V\otimes W)$$

that are basically the polynomials functions on $(V \otimes W)^* \cong \mathbb{C}^{n\times m}$. The group $GL_n \times GL_m$ acts on $(V \otimes W)^*$ by $(X,Y)M = XMY^{-1}$ so we can understand the structure of $S^d(V \otimes W)$ as a $GL_n \times GL_m$-representations. What is more, $GL_n \times GL_m$ has an induced action on the coordinate ring $\mathcal{O}(\mathbb{C}^{n\times m}) = \mathrm{Pol}(\mathbb{C}^{m\times n}) = S(V \otimes W)$. So we have just understood the representation of $GL_n \times GL_m$ on the coordinate ring of $\mathbb{C}^{n\times m}$.

3.3 Polynomial representations of $GL_n(\mathbb{C})$

Definition 3.27. Let $\rho : GL_n(\mathbb{C}) \to GL(W)$ be a representation where we assume that $\dim_{\mathbb{C}} W < \infty$. ρ is called a *polynomial representation* if

$$M(\rho) \subseteq \mathrm{Pol}(GL_n(\mathbb{C})) := \{f|_{GL_n(\mathbb{C})} \mid f \in \mathrm{Pol}(\mathbb{C}^{n\times n})\} \cong \mathrm{Pol}(\mathbb{C}^{n\times n})$$

Besides, ρ is called *homogeneous* of degree d if $M(\rho) \subseteq \mathrm{Pol}(GL_n(\mathbb{C}))_d$ of degree d homogeneous polynomials on $GL_n(\mathbb{C})$.

Remark 3.28. Last week we have seen that $\mathrm{Pol}(GL_n(\mathbb{C}))_d = M(T^d(\mathbb{C}^n))$, see Proposition 3.18.

Proposition 3.29. *A polynomial representation of $GL_n(\mathbb{C})$ decomposes as the direct sum of homogeneous representations.*

Proof. Let $\rho(g) = \big((\rho_{ij}(g))\big)_{i,j}$ where $\rho_{ij} \in \mathrm{Pol}(GL_n(\mathbb{C}))$ for all i,j. These matrix elements decompose as

$$\rho_{ij} = \sum_d \rho_{ij}^d$$

where $\rho_{ij}^d \in \mathrm{Pol}(GL_n(\mathbb{C}))_d$. Let us use the notation $\rho_d(g) = \big((\rho_{ij}^d(g))\big)_{i,j}$.

$$\rho(\lambda \cdot \mathbf{1}_{GL_n}) = \sum_d \lambda^d P_d \qquad \lambda \in \mathbb{C}$$

where $\rho_d(\mathbf{1}_{GL_n}) = P_d \in \mathrm{End}_{\mathbb{C}}(W)$. The plan is to prove that these P_d's are module homomorphisms that are orthogonal projections and their image is a homogeneous representation of degree d. This will prove the statement.

Now, we can apply the previous equation on $\lambda\mu$:

$$\rho\big((\lambda\mu)\mathbf{1}_{GL_n}\big) = \sum_d (\lambda\mu)^d P_d$$

However,

$$\rho\big((\lambda\mu)\mathbf{1}_{GL_n}\big) = \rho\big(\lambda\mathbf{1}_{GL_n}\big)\rho\big(\mu\mathbf{1}_{GL_n}\big)$$

because ρ is a representation. So these two equations give

$$\sum_d (\lambda\mu)^d P_d = \sum_d \lambda^d P_d \sum_e \mu^e P_e \qquad \forall \lambda,\mu \in \mathbb{C}$$

Therefore

$$P_d \cdot P_e = \begin{cases} 0 & \text{if } d \neq e \\ P_d & \text{if } d = e \end{cases}$$

In particular $P_d^2 = P_d \in \mathrm{End}_{\mathbb{C}}(W)$ so these constitute an orthogonal system of projections. In addition, $P_d \in \mathrm{End}_{GL_n(\mathbb{C})}(W)$ holds as well because $g = \mathbf{1}_{GL_n} \cdot g$ so

$$\left(\sum_d \lambda^d P_d\right) \cdot \rho(g) = \rho(\lambda\mathbf{1}_{GL_n(\mathbb{C})}) \cdot \rho(g) = \rho(\lambda g) = \rho(g) \cdot \rho(\lambda\mathbf{1}_{GL_n}) = \rho(g) \cdot \sum_d \lambda^d P_d$$

i.e. $P_d \cdot \rho(g) = \rho(g) \cdot P_d$ for all $g \in GL_n(\mathbb{C})$ and all d.

This argument shows that $W = \oplus_d P_d(W)$ as a GL_n-module. Moreover,

$$\rho(\lambda g)P_d(w) = \lambda^d \rho(g) P_d(w)$$

for all $w \in W$ by the above calculation. So the restriction of the representation to the subspaces $P_d(W)$ is homogeneous of degree d. □

$\mathrm{Pol}(GL_n(\mathbb{C}))_d = M(\rho)$ where $\rho : GL_n(\mathbb{C}) \to T^d(\mathbb{C}^n)$ is the usual tensor-power representation. However, we know the decomposition of ρ so

$$\mathrm{Pol}(GL_n(\mathbb{C}))_d = M(\rho) = \sum_{\lambda \vdash d;\ \mathrm{ht}(\lambda) \leq n} M(S_\lambda)$$

where $S_\lambda : GL_n(\mathbb{C}) \to GL(S_\lambda(\mathbb{C}^n))$. Therefore, considering all the degrees:

$$\mathrm{Pol}(GL_n(\mathbb{C})) = \sum_{\lambda \text{ partition};\ \mathrm{ht}(\lambda) \leq n} M(S_\lambda)$$

as $GL_n(\mathbb{C}) \times GL_n(\mathbb{C})$-modules. Therefore,

$$\mathrm{Pol}(GL_n(\mathbb{C})) \cong \bigoplus_{\lambda \text{ partition};\ \mathrm{ht}(\lambda) \leq n} S_\lambda(\mathbb{C}^n)^* \otimes S_\lambda(\mathbb{C}^n)$$

since as a $GL_n(\mathbb{C}) \times GL_n(\mathbb{C})$-modules $M(S_\lambda) \cong S_\lambda(\mathbb{C}^n)^* \otimes S_\lambda(\mathbb{C}^n)$ as we have seen this in the Topics in Algebra course.

Corollary 3.30. *The module* $\mathrm{Pol}(GL_n(\mathbb{C}))$ *is completely reducible in both sense (as a* $GL_n \times GL_n$*-module and as a* GL_n*-module as well). Hence, every polynomial representation is completely reducible.*

Proof. The above decomposition of $\mathrm{Pol}(GL_n(\mathbb{C}))$ proves its complete reducibility. Now, if ψ is a polynomial representation then $\psi + \psi + \cdots + \psi \cong_G M(\psi) \hookrightarrow_G \mathrm{Pol}(GL_n(\mathbb{C}))$ so ψ is a submodule of a completely reducible module what proves the statement. □

Remark 3.31. On $GL_n(\mathbb{C})$ we can give a natural affine variety structure by

$$GL_n(\mathbb{C}) = \{\det \neq 0\} \subseteq \mathbb{C}^{n \times n}$$

so we get that the coordinate ring of $GL_n(\mathbb{C})$ is

$$\mathcal{O}(GL_n(\mathbb{C})) = \mathbb{C}[x_{ij} \mid i, j \leq n]_{\det((x_{ij}))}$$

where the lower index means localization.

If we use the notation $G = GL_n(\mathbb{C})$ then the following containments hold

$$\mathrm{Fun}(G, \mathbb{C}) \supseteq \mathcal{O}(G) \supseteq \mathrm{Pol}(G) = \mathbb{C}[x_{ij} \mid i, j \leq n]$$

where every term is a $G \times G$-module. Moreover, the containments should be understood as submodules.

Eighth Lecture, 21th of November

4 Topological groups

Definition 4.1. A *topological group* G is a group endowed with a topology such that the maps

$$G \to G \qquad \text{and} \qquad G \times G \to G$$
$$x \mapsto x^{-1} \qquad\qquad (x, y) \mapsto xy$$

are continuous (where we consider the product topology on G).

Remark 4.2. Some also include Hausdorff-ness in the definition.

Example 4.3.

1. Every group with the discrete topology is a topological group.
2. If V is a finite dimensional vector space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ then $GL(V) \cong GL(n, \mathbb{K}) \subseteq \mathbb{K}^{n^2}$ is a topological group with the subspace topology.
3. Any subgroup of a topological group with the induced topology is also a topological group.

Definition 4.4. A *compact group* is a topological group which is compact as a topological space.

Example 4.5.

1. Every finite group with the discrete topology is compact.
2. The orthogonal group $O(V) \cong O_n(\mathbb{R})$ for a finite dimensional $\mathbb{R}$-vector space together with a scalar product. (It is compact because in $\mathbb{C}^{n\times n}$ it is closed and bounded.)
3. The unitary group $U(W) \cong U_n(\mathbb{C})$ for a finite dimensional $\mathbb{C}$-vector space together with a scalar product. (It is compact for the same reason.)
4. Any closed (!) subgroup of a compact group is compact.

Convention: In the following, the base field is $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ when not further specified. Besides, if G is a topological group then by representation in this chapter we mean continuous representation, i.e.

Definition 4.6. A finite dimensional representation of a topological group is *continuous* if $\rho : G \to GL(V)$ is a continuous map. Equivalently, if

$$M(\rho) := \mathrm{Span}_{\mathbb{K}}\{\rho_{ij} \mid 1 \leq i, j \leq \dim V\} \subseteq \mathcal{C}(G, \mathbb{K})$$

The main property of the finite dimensional representations is the following:

Theorem 4.7. *Every finite dimensional representation of a compact group is completely reducible.*

It follows from the existence of the Haar integral where this notion named after Alfred Haar is

Theorem 4.8. (Haar) *On a locally compact group there is a left invariant positive integral (i.e. it is nonnegative-on-nonnegative-functions, $\mathbb{R}$-linear and is zero only if $f \equiv 0$). Moreover, it is unique up to a non-zero scalar multiple.*

Definition 4.9. In the theorem, the *left invariance* means that for the functions

$$L_g : \mathcal{C}(G, \mathbb{K}) \to \mathcal{C}(G, \mathbb{K})$$
$$f \mapsto (x \mapsto f(gx))$$

where $g \in G$ is arbitrary we have

$$\int_{x\in G} f \,\mathrm{d}x = \int_{x\in G} L_g(f) \,\mathrm{d}x$$

And similarly with R_g in the case of *right invariance.*

Corollary 4.10. *On a compact group there is an integral which is both left and right invariant and is unique if we require $\int_{x\in G} 1 \, \mathrm{d}x = 1$.*

Proof. Suppose G is compact so by Haar's theorem we have a left-invariant integral on G which is unique up to scalar multiple. Fix $h \in G$ and let us define a new integral on G by

$$\int_G^{'} f := \int_G R_g f$$

One can check that this is also left-invariant so there must exist a positive scalar such that $\int_G R_h f = \lambda(h) \int_G f$. This map λ is a continuous group homomorphism. Suppose that there exists $\lambda(h) \neq 1$. We can assume that in fact it is larger than one possible by interchanging it with h. Then $\lambda(h^n) = \lambda(h)^n \to \infty$ which is a contradiction because the image of the compact set G under λ should be compact which is a contradiction. ☐

Corollary 4.11. *Every finite dimensional representation of a compact group is orthogonal for some scalar product in the case of $\mathbb{K} = \mathbb{R}$ (respectively, unitary in the case of $\mathbb{K} = \mathbb{C}$).*

Proof. Indeed, the proof is – in principle – the same as in the finite case: one can average an arbitrary scalar product on the underlying vector space by the Haar integral and so get an invariant measure by uniqueness. In details, the action of G on V induces an action on the space of symmetric bilinear functions $S^2(V^*)$ on V. Then take any scalar product β on V and define

$$\eta(x, y) = \int_{g\in G} (g\beta)(x, y) \, \mathrm{d}g = \int_{g\in G} \beta(gx, gy) \, \mathrm{d}g$$

which is invariant by the construction. ☐

Proof. (of Theorem 4.7) The previous corollary implies the statement because orthogonal (resp. unitary) representations are always completely reducible since if $W \subseteq U$ is an invariant subspace then $W^\perp$ too. For the details, see Topics in Algebra course notes. ☐

Proof. (alternative proof of Theorem 4.7 without referring to Haar integral) Let $\rho : G \to GL(V)$ be an n-dimensional representation of G. Denote by B the set of symmetric bilinear forms on V and by P the cone of positive definite symmetric bilinear forms. We can identify B with $\mathbb{R}^{\frac{n(n-1)}{2}}$ so it inherits a topology. The cone P is open with respect to this topology and it is even G-invariant. Now, let K_0 be a compact subset of P of positive Lebesgue measure. Take

$$K = \bigcup_{g\in G} g(K_0)$$

It is a compact subset of P still of positive measure (compact because it is the image of the product map $G \times K_0 \to P$ from a compact space). Moreover, $g(K) = K$ for any $g \in G$ by the construction. So we can take the center of mass of K and that has to be a g-invariant element in P. In details, let us define

$$c(K) = \frac{1}{\lambda(K)} \cdot \int_{x\in K} x \, \mathrm{d}\lambda(x)$$

for this we have

$$c(K) = c(gK) = \frac{1}{\lambda(gK)} \cdot \int_{x\in gK} x \, \mathrm{d}\lambda(x) = \frac{1}{|\det(g)|\lambda(K)} \cdot \int_{x\in K} gx \, \mathrm{d}\lambda(gx) =$$

$$= \frac{1}{|\det(g)|\lambda(K)} \cdot \int_{x\in K} gx \cdot |\det(g)| \, \mathrm{d}\lambda(x) = gc(K)$$

so $c(K)$ is in fact G-invariant. What is more, it is also in P by the convexity of P.

Therefore, we again got a G-invariant inner product on V so the statement follows by the same argument as before. ☐

4.1 The Peter-Weyl Theorem

Definition 4.12. Let G be a topological group and $\mathcal{C}(G) := \mathcal{C}(G, \mathbb{K})$. A function $f \in \mathcal{C}(G)$ is *representative* if f is contained in some finite dimensional (left-)G-stable subspace.

Lemma 4.13. *For an $f \in \mathcal{C}(G)$ the following are equivalent:*

1. *f is a representative function*
2. $\mathrm{Span}_{\mathbb{K}}\{L_g f \mid g \in G\}$ *is finite dimensional*
3. $\mathrm{Span}_{\mathbb{K}}\{R_g f \mid g \in G\}$ *is finite dimensional*
4. $\mathrm{Span}_{\mathbb{K}}\{L_{g^{-1}} R_h f \mid g, h \in G\}$ *is finite dimensional (i.e. it is contained in a finite dimensional $G \times G$-stable subspace)*
5. *there exist a $k \in \mathbb{N}$ and $u_i, v_i \in \mathrm{Fun}(G, \mathbb{K})$ for all $1 \le i \le k$ such that*

$$f(xy) = \sum_{i=1}^{k} u_i(x) v_i(y)$$

Moreover, if f is a representative function then these u_i's and v_i's can be chosen to be representative functions.

Remark 4.14. This proves that the notion representative function is left-right symmetric.

Proof. 1) $\iff$ 2) is obvious by definition. 4) $\Rightarrow$ 2) & 3) is also trivial. Besides 5) $\Rightarrow$ 2) & 3) is straightforward to check. So it is enough to prove 2) $\Rightarrow$ 5) & 4).

Suppose 2) holds and let $v_1, \ldots, v_k$ be a basis in $\mathrm{Span}_{\mathbb{K}}\{L_g f \mid g \in G\}$. Then we can define the functions u_i as

$$f(gx) = (L_{g^{-1}} f)(x) = \sum_{i=1}^{k} u_i(g) v_i(x)$$

so we got statement 5).

The appearing u_i's are in fact continuous functions because for fixed x this is a continuous functions and v_i's are linearly independent so there exists $x_1, \ldots, x_k \in G$ such that $\big((v_i(x_j))\big)_{i,j}$ is an invertible $k \times k$-matrix (see Lemma 3.4) hence u_i's can be expressed by finite linear combinations (using the inverse of the mentioned matrix) of continuous functions. Now, we show that these u_i's are even representative functions.

$$\sum_{i=1}^{k} u_i(x) v_i(zy) = f(xzy) = \sum_{i=1}^{k} u_i(xz) v_i(y)$$

where the v_i's are representative functions so we can apply 5) on them so there exist $c_{ij} \in \mathcal{C}(G)$ such that

$$\sum_{i=1}^{k} u_i(x) v_i(zy) = \sum_{i=1}^{k} u_i(x) \sum_{j=1}^{k} c_{ij}(z) v_j(y) =$$

$$= \sum_{j=1}^{k} \left(\sum_{i=1}^{k} u_i(x) c_{ij}(z) \right) v_j(y) = \sum_{i=1}^{k} \left(\sum_{j=1}^{k} u_j(x) c_{ji}(z) \right) v_i(y)$$

so the linear independence of $v_j's$ imply that

$$u_i(xz) = \sum_{j=1}^{k} u_j(x) c_{ji}(z)$$

so that 5) holds for the u_{ij} hence they are representative.

We also showed that

$$f(xzy) = \sum_{i=1}^{k}\sum_{j=1}^{k} u_i(x)c_{ij}(z)v_j(x) \in \mathrm{Span}_{\mathbb{K}}\{c_{ij} \mid i,j = 1,\ldots,k\}$$

so all bi-translations of f is in a finite dimensional invariant subspace. □

Proposition 4.15. *The space $\mathcal{T}_G := \{$representative functions of $G\} \subseteq \mathcal{C}(G)$ is spanned by the space of matrix elements of the finite dimensional continuous representations of G.*

Proof. Let $\rho : G \to GL_n(\mathbb{K})$ be a representation with coordinate functions $\rho_{ij} \in \mathcal{C}(G)$. Then by the rule of matrix multiplication

$$\rho_{ij}(xy) = \sum_{k=1}^{n} \rho_{ik}(x)\rho_{kj}(y)$$

so part 5) of the lemma above holds for ρ_{ij}. Therefore ρ_{ij} is a representative functions. In other words, $\sum_{\rho \text{ fin dim rep}} M(\rho) \subseteq \mathcal{T}_G$.

Conversely, take a representative function f. Let $U := \mathrm{Span}_{\mathbb{K}}\{R_g f \mid g \in G\} \subseteq \mathcal{C}(G)$ a finite dimensional subspace by Lemma 4.13. Take a basis $u_1, \ldots, u_n$ in this subspace. By the same Lemma we have

$$u_i(xy) = \sum_{j=1}^{n} u_j(x)c_{ji}(y)$$

where one can realize that these c_{ji}'s are exactly the matrix elements of R_y on U. Moreover,

$$u_i(g) = u_i(1g) = \sum_{j=1}^{n} u_j(1)c_{ji}(g)$$

so $u_i \in M(\rho)$ where $\rho : G \to GL(U)$; $y \mapsto R_y|_U$ and we got the statement noticing that this representation is continuous because c_{ji}'s are continuous by the 'Moreover' part of the Lemma. □

Corollary 4.16. *Representative functions form a subalgebra in $\mathcal{C}(G)$ becaus of the previous proposition and $M(\rho) \cdot M(\psi) = M(\rho\psi)$ (see Topics in Algebra, week 5 homework).*

Theorem 4.17. *Let G be a compact group. Then*

$$\mathcal{T}_G = \bigoplus_{\rho \in \mathrm{Irr}(G)} M(\rho)$$

where $\mathrm{Irr}(G)$ *is a complete list of finite dimensional irreducible representations of G.*

Remark 4.18. (Sketch) Basically the consequence of the previous theorem and general statements about the independence of $M(\rho)$'s for arbitrary group.

Ninth Lecture, 28th of November

Proof. By Proposition 4.15 we know that $\mathcal{T}_G$ is spanned by $M(\rho)$'s when ρ ranges over the finite dimensional representations of G. However, every finite dimensional representation ρ can be decomposed as an integer linear combination of irreducible representations ρ_i as

$$\rho = \sum_{i=1}^{k} m_i\rho_i$$

and then $M(\rho) = \sum_{i=1}^{k} M(\rho_i)$. So we only have to prove that the sum in $\mathcal{T}_G = \sum_{\rho \in \mathrm{Irr}(G)} M(\rho)$ is direct.

Note that $G \times G$ acts on the $M(\rho)$ spaces irreducibly (see Topics in algebra notes about $M(\rho) \cong V^* \otimes V$) and pairwise non-isomorphicly. Indeed, if $M(\rho) \cong M(\eta)$ then

$$\rho^* \otimes \rho \cong \psi^* \otimes \psi$$

as $G \times G$ modules for some representations $\rho, \psi \in \operatorname{Irr}(G)$. Then as G-modules this means

$$\rho + \cdots + \rho \cong \psi + \cdots + \psi$$

what gives $\rho \cong \psi$. So the following easy lemma completes the proof of the theorem:

Lemma 4.19. *A family of pairwise non-isomorphic irreducible submodules are independent.*

□

Definition 4.20. On a compact group G there exists a (canonical) $G \times G$-invariant scalar product on $\mathcal{C}(G, \mathbb{C})$:

$$\langle f_1, f_2 \rangle := \int_{x \in G} \overline{f_1(x)} f_2(x) \,\mathrm{d}x \qquad f_1, f_2 \in \mathcal{C}(G, \mathbb{C})$$

where $\mathrm{d}x$ stands for the Haar integral and $G \times G$-invariance means for arbitrary $(g, h) \in G \times G$ and $f_1, f_2 \in \mathcal{C}(G, \mathbb{C})$

$$\langle (g,h) f_2, (g,h) f_2 \rangle \overset{def}{=} \int_{x \in G} \overline{f_1(g^{-1}xh)} f_2(g^{-1}xh) \,\mathrm{d}x = \langle f_1, f_2 \rangle$$

Let $\{\rho^\lambda : G \to GL(V^\lambda) \mid \lambda \in \Lambda\}$ be a complete list of representations of the isomorphism classes of the finite dimensional irreducible complex representations of G. Recall that last week we found a G-invariant scalar product on V^λ (what is unique up to scalar multiple). Denote by ρ^λ_{ij} the matrix elements of ρ with respect to an orthonormal basis e in V^λ endowed with the above invariant scalar product. Then we can consider the above representations as $\rho^\lambda : G \to U(\dim V^\lambda, \mathbb{C})$ homomorphisms where U stands for unitary matrices.

Lemma 4.21. *With respect to the canonical scalar product on $\mathcal{C}(G, \mathbb{C})$ we have*

$$\langle \rho^\lambda_{i,j}, \rho^\mu_{k,l} \rangle = \begin{cases} 0 & \text{if } (\lambda, i, j) \neq (\mu, k, l) \\ \frac{1}{\dim V^\lambda} & \text{if } (\lambda, i, j) = (\mu, k, l) \end{cases}$$

The proof is verbatim the same as in the case of finite groups.

Proof. By Schur's lemma if U and W are minimal G-invariant subspaces in V such that $\rho_U \ncong \rho_W$ then U and W are orthogonal with respect to any invariant scalar product on V. (For detailed proof, see Topics in Algebra course). Besides, recall that if $\xi : G \to GL(V)$ is an irreducible representation then up to scalar multiple there is at most one invariant scalar product on V (again, see Topics in Algebra course for the proof).

Therefore, by the first mentioned statement $\langle \rho^\lambda_{i,j}, \rho^\mu_{k,l} \rangle = 0$ if $\lambda \neq \mu$ since $\rho^\lambda_{i,j} \in M(\rho^\lambda)$ and $\rho^\mu_{i,j} \in M(\rho^\mu)$. So for the remaining part of the proof let us use the notation $\rho = \rho^\lambda = \rho^\mu$, $V = V^\lambda$ and $\dim V^\lambda = n$. In these terms we want to prove that

$$\langle \rho_{ij}, \rho_{kl} \rangle = \begin{cases} 0 & \text{if } (i, j) \neq (k, l) \\ \frac{1}{n} & \text{if } (i, j) = (k, l) \end{cases}$$

These functions live in $M(\rho) \cong V^* \otimes V \cong \operatorname{End}_{\mathbb{C}}(V)$ using the $G \times G$-isomorphisms

$$\mu : \operatorname{End}_{\mathbb{C}}(V) \to M(\rho)$$

$$A \mapsto \Big(x \mapsto \operatorname{Tr}\big(A\rho(x)\big) \Big)$$

with the action $(g,h)A = \rho(g)A\rho(h)^{-1}$ on $\mathrm{End}_{\mathbb{C}}(V)$. This μ can pull back the invariant scalar product from $M(\rho)$ to $\mathrm{End}_{\mathbb{C}}(V)$. However, by the secondly mentioned theorem, there exists a $c \in \mathbb{R}_{>0}$ such that this induced invariant scalar product is c times the usual invariant scalar product on $\mathrm{End}_{\mathbb{C}}(V)$ defined by $(A,B) \mapsto \mathrm{Tr}(A^*B)$. In formulas,

$$\langle \mu(A), \mu(B) \rangle = c\mathrm{Tr}(A^*B)$$

for all $A, B \in \mathrm{End}_{\mathbb{C}}(V)$.

Now, if we identify $\mathrm{End}_{\mathbb{C}}(V)$ with $\mathbb{C}^{n\times n}$ by the previously chosen orthonormal basis e then

$$\mu(E_{ij}) = \Big(x \mapsto \mathrm{Tr}\big(E_{ij}(\rho(x))_e\big)\Big) = \big(x \mapsto \rho_{j,i}(x)\big)$$

so $\mu(E_{ij}) = \rho_{ji}$. Since E_{ij} is an orthonormal basis on $\mathbb{C}^{n\times n}$ we got that

$$\langle \rho_{ij}, \rho_{kl} \rangle = \langle \mu(E_{ji}), \mu(E_{lk}) \rangle = c\mathrm{Tr}(E_{ji}E_{lk}) = \begin{cases} 0 & \text{if } (i,j) \neq (k,l) \\ c & \text{if } (i,j) = (k,l) \end{cases}$$

hence it is enough to find c to complete the proof.

Mimicking the technique we used in the case of finite groups, we use the fact that $\rho(x)$ is a unitary matrix for all $x \in G$. In particular,

$$\sum_{i=1}^{n} \overline{\rho_{i1}(g)}\rho_{i1}(g) = 1$$

what we can "sum up" for all $g \in G$:

$$\int_{g\in G} \sum_{i=1}^{n} \overline{\rho_{i1}}\rho_{i1} \, \mathrm{d}x = \int_{x\in G} 1 \, \mathrm{d}x = 1$$

On the other hand,

$$\int_{g\in G} \sum_{i=1}^{n} \overline{\rho_{i1}}\rho_{i1} \, \mathrm{d}x = \sum_{i=1}^{n} \int_{g\in G} \overline{\rho_{i1}}\rho_{i1} \, \mathrm{d}x = \sum_{i=1}^{n} \langle \rho_{i1}, \rho_{i1} \rangle = nc$$

and we got the statement. □

Goal: To show that $\mathcal{T}_G$ is (uniformly) dense in $\mathcal{C}(G,\mathbb{C})$ i.e. dense in the supremum-norm: $\forall f \in \mathcal{C}(G,\mathbb{C})$ $\exists h \in \mathcal{T}_G$ such that $\max|f-h| < \varepsilon$.

Theorem 4.22. (Stone-Weierstrass Theorem) *Let X be a compact topological space and A be an $\mathbb{R}$-subalgebra of $\mathcal{C}(X,\mathbb{R})$, $1 \in A$ such that for all $x \neq y \in X$ there exists an $a \in A$ such that $a(x) \neq a(y)$ then A is uniformly dense in $\mathcal{C}(X,\mathbb{R})$.*

Proof. Let S be the closure of A, we prove that $S = \mathcal{C}(X,\mathbb{R})$. By Weierstrass theorem: if $f \in S$ then $|f| \in S$ since we can approximate the real $|.|$ function. This also means that for $f, g \in S$ we have

$$\min(f,g) = \frac{f+g-|f-g|}{2} \in S \qquad \text{and similarly} \qquad \max(f,g) = \frac{f+g+|f-g|}{2} \in S$$

Now we prove that for any pair $(p,q) \in \mathbb{R}^2$ and $x \neq y \in X$ there exist an element $b \in S$ such that $b(x) = p$ and $b(y) = q$. Indeed, by the assumptions there is an $s \in s$ such that $s(x) \neq s(y)$ so let

$$b = \frac{p-q}{s(x)-s(y)}s + \left(\frac{qs(x)-ps(y)}{s(x)-s(y)}\right)1$$

which satisfies the requirements.

Now, take any $f \in \mathcal{C}(X, \mathbb{R})$. For all $(x, y) \in X \times X$ take $g_{(x,y)} \in S$ such that

$$g_{(x,y)}(x) = f(x)$$

$$g_{(x,y)}(y) = f(y)$$

For a fixed $\varepsilon > 0$ and fixed $x \in X$ there exists an open neighborhood of y: $U_y \subseteq X$ such that for all $z \in U_y$, $g_{(x,y)}(z) > f(z) - \varepsilon$ holds. From the open covering $\cup_{y \in X} U_y$ we can choose a finite covering of X i.e.$X = \cup_{i=1}^n U_{y_i}$ and by this we simultaneously get a finite set of functions

$$\{g_{(x,y_i)} \mid i = 1, \dots\} \subseteq S$$

so S contains their maximum as well:

$$g_x \stackrel{def}{=} \max_i g_{(x,y_i)} \in S$$

for which we have the property

$$g_x(z) > f(z) - \varepsilon \qquad \forall z \in \bigcup_{i=1}^{n} U_{y_i} = X$$

Now, we can do the same compactness argument for open subsets V_x such that $g_x(z) < f(z) + \varepsilon$ for all $z \in V_x$ and by taking minimum we get a function $g = \min g_x \in S$ that is ε-close to f. Here, ε was arbitrarily small and S is closed so $S = \mathcal{C}(X, \mathbb{R})$. ☐

Corollary 4.23. *Let X be a compact topological space and let A be a $\mathbb{C}$-algebra such that $1 \in A \subseteq \mathcal{C}(X, \mathbb{C})$ and A separates the points of X (as in the previous proposition). If, moreover, A is closed with respect to the complex conjugation then A is uniformly dense in $\mathcal{C}(X, \mathbb{C})$.*

Proof. Take $S := \bar{A}$ as in the previous theorem. For an arbitrary $f \in S$ the real and imaginary parts can be expressed as

$$\text{Re} f = \frac{1}{2}(f + \overline{f}) \in S \qquad \text{and} \qquad \text{Im} f = \frac{1}{2i}(f - \overline{f}) \in S$$

So let $S' := S \cap \mathcal{C}(X, \mathbb{R})$. Now, we check the assumptions of the Stone-Weierstrass Theorem 4.22 on this S': it separates X since if $f \in S$ separates two points then either $\text{Re} f \in S'$ or $\text{Im} f \in S'$ separates them as well. Therefore, $S' = \mathcal{C}(X, \mathbb{R})$. Then for an arbitrary $f \in \mathcal{C}(X, \mathbb{C})$, $\text{Re} f, \text{Im} f \in S$ and S is a $\mathbb{C}$-algebra so $f = \text{Re} f + i \text{Im} f \in S$. ☐

Theorem 4.24. *Let $G \subseteq U(n, \mathbb{C})$ be a compact subgroup. (Equivalently, G is a compact linear group, i.e. one that is compact and has a faithful representation in $GL(n, \mathbb{C})$ for some n, so we can endow the space with an inner product so that G is unitary with respect to it.) Then $\mathcal{T}_G$ is generated by the matrix entries and the inverse of the determinant. Moreover, $\mathcal{T}_G$ is uniformly dense in $\mathcal{C}(G, \mathbb{C})$.*

By the "matrix entries" we think about the matrix elements of the defining representation of G on $\mathbb{C}^n$ so they belong to $\mathcal{T}_G$ by definition. Besides, the inverse of the determinant

$$\det{}^{-1} : G \to \mathbb{C}^\times$$

$$g \mapsto \det(g)^{-1}$$

is a 1 dimensional representation of G so $\det^{-1} \in \mathcal{T}_G$ again by definition.

Proof. Let A be the $\mathbb{C}$-subalgebra generated by $\det^{-1}$ and the matrix entries. In the above, we saw that $A \subseteq \mathcal{T}_G$ since the elements of A are matrix elements of a finite dimensional representation. Then, informally speaking, it is enough to show that A cannot be "too small" so we can deduce the statement from the Stone-Weierstrass theorem:

Apply the Stone-Weierstrass Theorem on A: The matrix entries obviously separate points in $G = X$ since the (say, tautological) representation of G is faithful. The constant function is in A for example by taking $\det \cdot \det^{-1}$. So to prove that it is closed under conjugation it is enough to construct the entries of the matrix g^{-1} since then by $\overline{g_{ij}} = (g^{-1})_{ji}$ we can express the conjugates of the matrix elements. The conjugate of the inverse of the determinant is even more easier because $\overline{\det(g)^{-1}} = \det(g)$ where det is trivially in the subalgebra of the matrix entries.

However, the entries of the g^{-1} have an explicit description by the entries of g using Cramér's rule:

$$g_{ij}^{-1} = \frac{1}{\det(g)}(-1)^{i+j}\det(\hat{g}_{ij}) \in A$$

where $\hat{g}_{ij}$ is the matrix g without the i-th row and the j-th column. So g_{ij}^{-1} is indeed in A hence A is closed under conjugation. So – by the Stone Weierstrass Theorem – we get that A is uniformly continuous in $\mathcal{C}(G, \mathbb{C})$.

Now, suppose indirectly that

$$A \subsetneq \mathcal{T}_G = \bigoplus_{\rho \in \mathrm{Irr}(G)} M(\rho)$$

Then there exists a $\psi \in \mathrm{Irr}(G)$ such that $M(\psi) \not\subseteq A$ so it has to be orthogonal as well by general propositions in Topics in Algebra course. But $0 \neq M(\psi) \perp A$ contradicts the uniform density of A in $\mathcal{C}(G, \mathbb{R})$ so the theorem follows. ☐

Summarizing the proved statements:

Theorem 4.25. (Peter-Weyl) *The unitary matrix elements of the irreducible finite dimensional representations (on $\mathbb{C}$) of a compact linear group G constitute a complete orthogonal system in $\mathcal{C}(G, \mathbb{C})$.*

Proof. By Lemma 4.1 they are orthogonal and by Theorem 4.24 their span is dense. ☐

Here, a complete orthogonal system means a sequence $\{f_i \in \mathcal{C}(G, \mathbb{C}) \mid i \in I\}$ such that $\langle f_i, f_j \rangle = 0$ for $i \neq j \in I$ and $\mathrm{Span}_{\mathbb{C}}\{f_i \mid i \in I\}$ is uniformly dense in $\mathcal{C}(G, \mathbb{C})$.

Remark 4.26. The theorem is true without the assumption of linearity but the proof is more technical (using functional analysis).

Remark 4.27. In the theorem we did not state that the complete system is countable but in the linear case it is also true.

By the theorem, any $f \in \mathcal{C}(G, \mathbb{C})$ can be expanded as a uniformly convergent series

$$f = \sum_k a_k f_k \qquad \text{where} \qquad a_k = \frac{\langle f, f_k \rangle}{\langle f_k, f_k \rangle}$$

usually called *Fourier expansion* of f. In the classical case, the group is $\mathbb{S}^1$ so $\mathcal{C}(\mathbb{S}^1, \mathbb{R})$ is basically space of the periodic functions on $\mathbb{R}$. So to determine the Fourier expansion, we have to know the irreducible representations of G. These are

$$\mathrm{Irr}(G) = \{\rho_n \mid n \in \mathbb{Z}\}$$
$$\rho_n : z \mapsto z^n$$

because for example, we know that all the finite dimensional irreducible representations are one dimensional since G is abelian (or it is a consequence of Theorem 4.24). So we can pull back the above expansion to $\mathbb{R}$ by $t \mapsto e^{it} = z$ hence the representation with $\cos(nt)$ and $\sin(nt)$. For this method to work generally, we only use that there are countably many irreducible representations of G:

Proposition 4.28. *For a compact linear group G,* $\mathrm{Irr}(G)$ *is countable.*

Proof. We may assume that $G \subseteq U_n(\mathbb{C})$. Then

$$\mathcal{T}_G = \mathbb{C}[\det^{-1}, g_{ij} \mid i,j \leq n]$$

It has a filtration by the finite dimensional $G \times G$ invariant subspaces

$$\mathcal{T}_G^k = \text{Span}\{\text{product of lenght } \leq k \text{ in } g_{ij}\text{'s}, \ \det^{-1}\}$$

so $\mathcal{T}_G = \cup_k \mathcal{T}_G^k$. If we introduce the notation

$$\text{Irr}(G)_k := \{\rho \in \text{Irr}(G) \mid M(\rho) \subseteq \mathcal{T}_G^k\}$$

or in other words $\mathcal{T}_G^k = \oplus_{\rho \in \text{Irr}(G)_k} M(\rho)$ then

$$\text{Irr}(G) = \bigcup_k \text{Irr}(G)_k$$

since every $M(\rho)$ is included in some $\mathcal{T}_G^k$ by their finite dimensionality. But the space $\mathcal{T}_G^k$ is finite dimensional for all k so $\text{Irr}(G)_k$ has to be a finite set for all k. Hence, $\text{Irr}(G)$ is countable because it has a filtration with finite sets. □

Tenth Lecture, 5th of December

4.2 The groups SU_2 and SO_3

Definition 4.29. The *special unitary group* is defined as

$$SU_2(\mathbb{C}) := \left\{ \begin{pmatrix} z_1 & -\overline{z_2} \\ z_2 & \overline{z_1} \end{pmatrix} \;\middle|\; z_1, z_2 \in \mathbb{C}, \ |z_1|^2 + |z_2|^2 = 1 \right\}$$

which is sitting in the ring of quaternions that can be seen if we define $\mathbb{H}$ as a subring of $M_2(\mathbb{C})$:

$$\mathbb{H} := \left\{ \begin{pmatrix} z_1 & -\overline{z_2} \\ z_2 & \overline{z_1} \end{pmatrix} \;\middle|\; z_1, z_2 \in \mathbb{C} \right\}$$

The defining relation of SU_2 is the same as to require that the determinant is 1 so SU_2 is topologically just the unit sphere in the Euclidean 4-space. Hence, SU_2 is a connected, simply connected topological group as well.

Remark 4.30. So SU_2 acts on $\mathbb{H}$ by left multiplication and this action preserves the quadratic form $\det : \mathbb{H} \to \mathbb{R}$. It implies that the Haar-integral of SU_2 is then the integral determined by the usual 3-volume inherited from $\mathbb{R}^4$, i.e. the integration on the sphere.

Remark 4.31. The isomorphism between the above $\mathbb{H}$ and the quaternions written by the usual formalism is

$$x_0 + x_1 i + x_2 j + x_3 k \mapsto \begin{pmatrix} x_0 + x_1 i & -x_2 + x_3 i \\ x_2 + x_3 i & x_0 - x_1 i \end{pmatrix}$$

which can be obtained by considering the action of $x_0 + x_1 i + x_2 j + x_3 k$ on the quaternion space as a two dimensional $\mathbb{C}$-vector space, spanned by 1 and j.

Definition 4.32. The *3-dimensional special orthogonal group* is

$$SO_3 = SO_3(\mathbb{R}) := \{A \in \mathbb{R}^{3\times 3} \mid A^T A = I, \ \det(A) = 1\}$$

Definition 4.33. There is an important surjective homomorphism $\rho : SU_2 \twoheadrightarrow SO_3$: let

$$\mathbb{E} := \left\{ \begin{pmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & -x_1 \end{pmatrix} \;\middle|\; x_1, x_2, x_3 \in \mathbb{R} \right\} \cong \mathbb{R}^3$$

i.e. the trace zero subset of $\mathbb{H}$. In this case, $\mathbb{E}$ contains only Hermitian matrices. We have an identification of the above space endowed with a tensor: $(\mathbb{E}, -\det) \cong (\mathbb{R}^3, \sum x_i^2)$. Then SU_2 acts on $\mathbb{E}$ by conjugation so it gives a continuous homomorphism (i.e. a representation of SU_2 as a topological group)

$$\rho : SU_2 \to GL(\mathbb{E})$$

$$A \mapsto (X \mapsto AXA^{-1})$$

which has the property $-\det(\rho(A)(X)) = \det(X)$ so ρ acts orthogonally. In formulas,

$$\rho : SU_2 \to O(\mathbb{E}, -\det) \cong O_3(\mathbb{R})$$

However, SU_2 is connected so its image under ρ is connected too soρ has values in $SO_3(\mathbb{R})$ since $O_3(\mathbb{R})$ has two components: $SO_3(\mathbb{R})$ and its complement and $1 \in SO_3(\mathbb{R})$ hence $\mathrm{Im}\rho \subseteq SO_3(\mathbb{R})$.

Proposition 4.34. *$\rho : SU_2 \to SO_3$ is surjective with kernel $\{\pm I\}$.*

Proof. Note that SO_3 acts transitively on $\mathbb{S}^2 \subseteq \mathbb{R}^3$. Moreover, if a subgroup $G \subseteq SO_3$ is transitive on $\mathbb{S}^2$ and contains all rotations through some line then $G = SO_3$. (Indeed, the rotations through a given line is exactly the stabilizer of a point of $\mathbb{S}^2$ in SO_3 hence the above statement.) So we prove that these are satisfied by $G := \rho(SU_2)$.

We know that any $X \in \mathbb{E}$ can be diagonalized by Hermitian property and this diagonalized version of X still has zero trace so it is of the form (or by interchanging the elements in the diagonal can be rearranged to)

$$c \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} =: cH$$

for some $c \in \mathbb{R}_{>0}$. This implies transitivity of G since $H \in \mathbb{S}^2 \subseteq \mathbb{E}^3$. For the other property of G we prove that G contains all the rotations around the line $\mathbb{R} \cdot H \subseteq \mathbb{E}$. Indeed, let $z \in \mathbb{C}$, $|z| = 1$ and define

$$A(z) := \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} \in SU_2$$

then

$$\rho(A(z)) \begin{pmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & -x_1 \end{pmatrix} = \begin{pmatrix} x_1 & z^2(x_2 + ix_3) \\ z^{-2}(x_2 - ix_3) & -x_1 \end{pmatrix}$$

so its action is indeed the rotation by $2 \cdot \arg(z)$ through H. □

Corollary 4.35. *In particular, this means that*

$$SO_3 \cong SU_2/\{\pm I\}$$

as groups. In fact, as topological groups too but we have not yet clarified what we mean by the factor of a topological group.

Definition 4.36. Let G be a topological group and N be a closed normal subgroup. Endow G/N with the topology as follows: $Z \subseteq G/N$ is closed if $\eta^{-1}(Z) \subseteq G$ is closed where $\eta : G \to G/N$ is the natural surjection.

Remark 4.37. Here, the assumption that N has to be closed is to ensure that G/N is Hausdorff.

Remark 4.38. η is continuous by definition, moreover, a function f on G/N is continuous if and only if $f \circ \eta$ is continuous on G. In particular, a representation of G/N is continuous if and only if it lifts to a continuous representation $T \circ \eta$ of G.

Lemma 4.39. *Let $\rho : G \to F$ be a continuous surjective homomorphism of a compact topological group G onto a topological group F. Then $F \cong G/\text{Ker}\rho$ as topological groups.*

Proof. Define the induced homomorphism $\overline{\rho}$ by the following commutative diagram:

$$\begin{array}{ccc} G & \xrightarrow{\rho} & F \\ {\scriptstyle\eta}\downarrow & \nearrow_{\overline{\rho}} & \\ G/\ker\rho & & \end{array}$$

then $\overline{\rho}$ is continuous because $\overline{\rho} \circ \eta = \rho$ is continuous. However, $\overline{\rho}^{-1}$ is continuous because a compact topology has no "Hausdorff weakening". In details, the bijection $\overline{\rho}^{-1}$ is continuous if $\overline{\rho}$ is open, i.e. it brings open subsets into open subsets. Or equivalently, it brings closed subsets into closed subsets which is now true because a closed subset of a compact space is compact and continuous function brings compact subsets into compact subsets. ☐

Corollary 4.40. *The continuous representations of SO_3 are in bijective correspondence in a natural way (i.e. composing with $\rho : SU_2 \to SO_3$) with those continuous representations of SU_3 that map -1 to the identity.*

Goal: To construct the irreducible representations of SU_2.

The group $GL_2(\mathbb{C})$ acts by right multiplication on the space $\mathbb{C}^2$ of row vectors. This induces a representation of $GL_2(\mathbb{C})$ on $\text{Fun}(\mathbb{C}^2, \mathbb{C})$:

$$\left(\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \cdot f \right)(x_1, x_2) = f(g_{11}x_1 + g_{21}x_2, g_{12}x_1 + g_{22}x_2)$$

If we restrict ourselves to $\text{Pol}(\mathbb{C}^2, \mathbb{C}) \subseteq \text{Fun}(\mathbb{C}^2, \mathbb{C})$. This subspace is $GL_2(\mathbb{C})$-invariant. Moreover, we can decompose the space as

$$\text{Pol}(\mathbb{C}^2, \mathbb{C}) = \bigoplus_{n=0}^{\infty} \text{Pol}_n(\mathbb{C}^2, \mathbb{C})$$

where Pol_n stands for the degree n homogeneous polynomials. These subspaces are again $GL_2(\mathbb{C})$-invariant subspaces. Now, set

$$\Phi_n : SU_2(\mathbb{C}) \to GL(\text{Pol}_n(\mathbb{C}^2, \mathbb{C}))$$

as the restriction of the action of $GL_2(\mathbb{C})$ on $\text{Pol}_n(\mathbb{C}^2, \mathbb{C})$. This is a representation of $SU_2(\mathbb{C})$ of degree $n+1$ since $\text{Pol}_n(\mathbb{C}^2, \mathbb{C}) = \text{Span}\{x^k y^{n-k} \mid k \leq n\}$. It is in fact a continuous representation since the action is a polynomial map on $\text{Pol}_n(\mathbb{C}^2, \mathbb{C})$ in the group elements.

Question: Why are they irreducible?

Proposition 4.41. *The representation Φ_n is irreducible for $n \in \mathbb{N}$.*

Proof. In SU_2 define a so called maximal torus by

$$\mathbb{T} := \left\{ A(z) := \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} \;\middle|\; z \in \mathbb{C},\ |z| = 1 \right\}$$

and restrict the representation Φ_n on this. Then

$$\Phi_n|_\pi : \pi \to GL(\text{Pol}_n(\mathbb{C}^2, \mathbb{C}))$$

is a representation of $\mathbb{T}$. Then this decomposes to 1 dimensional representations because $\mathbb{T}$ is abelian. How? To answer we compute its values:

$$\Phi_n(A(z))x^k y^{n-k} = z^{2k-n} x^k y^{n-k}$$

so $\mathbb{C}x^k y^{n-k}$ is a $\mathbb{T}$-invariant subspace of $\mathrm{Pol}_n(\mathbb{C}^2,\mathbb{C})$. Therefore,

$$\mathrm{Pol}_n(\mathbb{C}^2,\mathbb{C}) = \bigoplus_{i=0}^{n} \mathbb{C}x^k y^{n-k}$$

as a $\mathbb{T}$-module. Therefore, any $\mathbb{T}$ invariant subspace of $\mathrm{Pol}_n(\mathbb{C}^2,\mathbb{C})$ is spanned by $x^k y^l$'s since $\mathbb{T}$ acts pairwise non-isomorphicly on the above 1 dimensional subspaces. It means that an SU_2-invariant subspace – which is also $\mathbb{T}$-invariant – is spanned by the monomials. By direct computations one can check that no such invariant proper subspace exists beyond $\{0\}$: e.g. if we apply an element of SU_2 on $x^k y^{n-k}$ where none of its coordinates are zero then x^n has coefficient $g_{11}^k g_{12}^{n-k}$. So in an invariant subspace x^n is always there, but

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} x^n = \frac{1}{\sqrt{2}}(x+y)^n$$

where every monomial has nonzero coefficient so every monomial is included in the invariant subspace. ☐

Question: Why is it a complete list of the irreducible representations of SU_2?

Definition 4.42. If $\varphi : SU_2 \to GL(V)$ is a finite dimensional representation of SU_2 then

$$\mathrm{ch}_\varphi : G \to \mathbb{C} \qquad g \mapsto \mathrm{Tr}(\varphi(g))$$

One can check that $\mathrm{ch}_\varphi \in \mathcal{C}(G,\mathbb{C}) \cap \mathrm{Cent}(G)$ if φ is continuous.

Fortunately, every element of SU_2 is conjugate to a diagonal element of the form $\mathrm{diag}(z, z^{-1})$ for some $z \in \mathbb{C}$. Therefore it is enough to consider the character on $\mathbb{T}$. Formulated in a proposition this means that

Proposition 4.43. *Restriction of central functions on SU_2 to its subgroup $\mathbb{T}$ gives an injection of this algebra $\mathrm{Cent}(SU_2)$ into $\mathrm{Cent}(\mathbb{T}) = \mathcal{C}(\mathbb{T})$. In fact, into*

$$\mathcal{F} := \{f \in \mathcal{C}(\mathbb{T},\mathbb{C}) \mid f(A(z)) = f(A(z^{-1}))\}$$

since $A(z)$ is conjugate to $A(z^{-1})$ by the interchanging of the two diagonal elements.

Proposition 4.44. *Any function in $\mathcal{F}$ can be uniformly approximated by a function from* $\mathrm{Span}_\mathbb{C}\{1,\ z^k + z^{-k} \mid k \in \mathbb{N}_+\}$

Proof. We have already seen after the Peter-Weyl Theorem (4.25) that $\mathrm{Span}_\mathbb{C}\{z^k \mid k \in \mathbb{N}\}$ is uniformly dense in $\mathcal{C}(\mathbb{T},\mathbb{C})$. Therefore, if we approximate a function by g then the $z \mapsto g(z^{-1})$ will be a similarly close approximation so we can get a symmetrized approximation by the elements of $\mathrm{Span}_\mathbb{C}\{1,\ z^k + z^{-k} \mid k \in \mathbb{N}_+\}$. To tell it in another way: we can consider the infinite Fourier expansion and symmetry implies that the paired Fourier coefficients are the same. By that we can get an approximation taking the first n-terms of the expansion. ☐

Note that

$$\mathrm{Span}_\mathbb{C}\{1,\ z^k + z^{-k} \mid k \in \mathbb{N}_+\} = \mathrm{Span}_\mathbb{C}\{\mathrm{ch}_{\Phi_n}\big|_\mathbb{T} \mid n \in \mathbb{N}\}$$

because $\mathrm{ch}_{\Phi_n}(A(z))$ maps $x^k y^{n-k}$ into $z^{2k-n}x^k y^{n-k}$ hence the matrix of $\Phi_n(A(z))$ with respect to the monomial basis of $\mathrm{Pol}_n(\mathbb{C}^2,\mathbb{C})$ is $\mathrm{diag}(z^n, z^{n-2}, \ldots, z^{-n})$. Therefore, $\mathrm{ch}_{\Phi_n}(A(z)) = z^n + z^{n-2} + \cdots + z^{-n}$.

So what we got is that the characters of Φ_n form a uniformly dense subset in $\mathcal{F}$ therefore there cannot be another finite dimensional representation besides Φ_n's because that would be orthogonal to the span of Φ_n's by Theorem 4.17 but still in $\mathcal{F}$ which is impossible for a uniformly dense subset. The statement we just got is summarized in the following theorem.

Theorem 4.45. $\{\Phi_n \mid n \in \mathbb{N}\}$ *is a complete, irredundant list of representatives of the isomorphism classes of finite dimensional irreducible complex representation of the compact group SU_2.*

Remark 4.46. Note that infinite dimensional irreducible representations are still possible.

Corollary 4.47. $\Phi_0, \Phi_2, \Phi_4, \ldots$ *is a complete list of finite dimensional irreducible representation of* SO_3

Proof. To get the finite dimensional irreducible representations of SO_3 we just have to check whether Φ_n brings -1 to the identity. It is easy to see that it is true if and only if n is even. □

Remark 4.48. Attention: in this case, the dimension of the representation is $n+1$ which is odd. So there is no irreducible even dimensional representation of SO_3.

Remark 4.49. This is the simplest case of a general phenomenon that relates the representations of a topological group to the representations of the (unique) simply connected covering group that can be understood easier. (In the case of Lie groups, by turning to the representations of the Lie algebra.)

Eleventh Lecture, 12th of December

4.3 Homogeneous spaces

Suppose G is a compact group, acting on a Hausdorff topological space X, i.e. we have a continuous map $G \times X \to X$; $(g,x) \mapsto gx$ such that $1 \cdot x$ and $g(hx) = (gh)x$. Now, suppose also that the action is transitive. This action is equivalent to the action on the left cosets of a stabilizer, i.e. for a point $p \in X$ we can define the map $\mu : G \to X$; $g \mapsto gp$ and then the left cosets of $H := \mathrm{Stab}_G(p) = \{g \in G \mid gp = p\}$ are exactly the fibers of this map. Denote by G/H the set of left-cosets in G. The above argument gives a bijective correspondence between G/H and X. Besides, the subgroup H is closed (since the action is continuous) so we can get a Hausdorff topology on G/H by the standard way (see 4.36).

Until this point we did not use that the group is compact. So this extra property helps us by making the above bijection a homeomorphism since a continuous bijection from a compact space (as G/H) to a Hausdorff space is always a homeomorphism. So we got that if X is a homogeneous space (i.e. G acts on X continuously and transitively) then we can assume that $X = G/H$ for some closed subset H of G.

Assumption: From now on, the base field is $\mathbb{C}$ and G is assumed to be a compact group.

Now, let us investigate what are the continuous functions on G/H: by the definition of the topology of G/H we know that a function f on G/H is continuous if and only if its composition $f \circ \eta$ with the natural map $\eta : G \to G/H$ is continuous. Therefore,

$$\mathcal{C}(G/H) \cong \{f \in \mathcal{C}(G) \mid f(xh) = f(x)\ \forall h \in H, \forall x \in G\} = \mathcal{C}(G)^H$$

where $\cong$ means that we can identify them (at least as sets but in fact as G-sets) and where we considered the right regular representation of H on $\mathcal{C}(G)$.

We have proved at 4.17 that

$$\mathcal{C}(G) \supseteq \mathcal{T}_G = \bigoplus_{\rho \in \mathrm{Irr}(G)} M(\rho)$$

and it is a uniformly dense subset in $\mathcal{C}(G)$ where $M(\rho) \cong V_{\rho^*} \otimes V_\rho$ (as $G \times G$-modules) with the notation $\rho : G \to GL(V_\rho)$. Now, we can take the H-fixed points:

$$\mathcal{C}(G/H) = \mathcal{C}(G)^H \supseteq \mathcal{T}_G^H = \bigoplus_{\rho \in \mathrm{Irr}(G)} V_{\rho^*} \otimes V_\rho^H$$

where it is important to note that on $\mathcal{T}_G$ H acts by the right regular representation but on V_ρ it acts via ρ. Therefore,

$$\mathrm{Left} - \mathrm{reg}\Big|_{\mathcal{T}_G^H} = \bigoplus_{\rho \in \mathrm{Irr}(G)} \dim(V_\rho^H)\rho^*$$

Besides, $\mathcal{T}_G^H$ is also uniformly dense since the Peter-Weyl Theorem 4.25 says that for any $f \in \mathcal{C}(G)$ we have a unique expansion $f = \sum_{\rho \in \mathrm{Irr}(G)} f_\rho$ for some $f_\rho \in M(\rho)$. Now, suppose that $\mathrm{Right} - \mathrm{rep}(h)f = f$. So apply $\mathrm{Right} - \mathrm{rep}(h)$ to the expansion: then

$$f = \mathrm{Right} - \mathrm{rep}(h)f = \sum_{\rho \in \mathrm{Irrrep}} \mathrm{Right} - \mathrm{rep}(h)f_\rho$$

but $\mathrm{Right} - \mathrm{rep}(h)f_\rho \in M(\rho)$ because these are submodules. So by the uniqueness of the expansion, we have $\mathrm{Right} - \mathrm{rep}(h)f_\rho = f_\rho$ for all $h \in H$ and $\rho \in \mathrm{Irr}(G)$.

So we got that $\mathcal{T}_G^H$ is uniformly dense in $\mathcal{C}(G)^H$.

4.4 The Laplace spherical functions

Let $G := SO_3$. It acts transitively on $X := \mathbb{S}^2 \subseteq \mathbb{R}^3$. Denote by $p = (0, 0, 1) \in X$ and

$$H := \mathrm{Stab}_G(p) = \{\text{rotations through the axis}\} \cong SO_2$$

so we can identify X and G/H and the action of G on X gives a representation on $\mathcal{C}(X) = \mathcal{C}(G/H)$. Generally, the above explained general scheme applies for $G = SO_3$.

Lemma 4.50. *Every finite dimensional nonzero G-invariant subspace in $\mathcal{C}(X)$ contains a nonzero H-invariant element.*

Proof. Let $U \subseteq \mathcal{C}(X)$ be a nonzero finite dimensional SO_3-invariant subspace. Take a nonzero $f \in U$ so there exists a point $x \in X$ such that $f(x) \neq 0$. Then we can bring x into p by an element $g \in G$ by transitivity. For this g we have

$$(gf)(p) = f(g^{-1}p) = f(x) \neq 0$$

so we got a function in U that does not vanish on p.

This means that the functional $U \ni f \mapsto f(p)$ does not vanish on the whole U so we can define the codimension one subspace of U:

$$U_0 := \{f \in U \mid f(p) = 0\} = \ker\{f \mapsto f(p)\}$$

This U_0 is an H-invariant subspace since for any $h \in H$ and $f \in U_0$ we have

$$(hf)(p) = f(h^{-1}p) = f(p) = 0$$

But H is a compact group so U is completely reducible, i.e. there exists an $f_0 \in U$ such that

$$U = U_0 \oplus \mathbb{C}f_0$$

where this $\mathbb{C}f_0$ is H-invariant. We want a little more: that H fix f_0 itself and does not brings it into its scalar multiple. However, it is clear that even that is true because if we define $\lambda : H \to \mathbb{C}$ as

$$(hf_0) = \lambda(h)f_0$$

then obviously λ will be a nonzero group homomorphism of H into $\mathbb{C}^\times$. Then by $\lambda(h)f_0(p) = f_0(h^{-1}p) = f_0(p) \neq 0$ we get that $\lambda(h) = 1$ for all $h \in H$ hence the statement. ☐

Remark 4.51. In the proof, we treated $\mathcal{C}(X)$ as G were acting on it continuously, but we have not even a definition yet for infinite-dimensional continuous action. So when we apply the theorems (namely we state that U is completely reducible) then we should prove that G acts continuously on U but it is easy to check that.

Definition 4.52. Let $A = \mathbb{C}[x_1, x_2, x_3] \subseteq \mathrm{Fun}(\mathbb{R}^3, \mathbb{C})$. Then the action of SO_3 on $\mathbb{R}^3$ induces an action on $\mathrm{Fun}(\mathbb{R}^3, \mathbb{C})$ such that A is an invariant subspace in it. In fact, then G acts on A by the linear change of variables. (Note that it is a left action so g acts as multiplication by the matrix g^{-1} on the vector $[x_1, x_2, x_3]^T$.)

We can decompose this A as its homogeneous parts $A = \oplus_{m=0}^{\infty} A_m$. These A_m's are G-invariant subspaces for all m with dimension $\binom{m+2}{2}$. For technical reasons we introduce a Hermitian scalar product on A such that the monomials constitute an orthogonal basis with respect to it. However, they will not be an orthonormal basis but we set the scalar product in a way that

$$(x_1^{k_1} x_2^{k_2} x_3^{k_3}, x_1^{k_1} x_2^{k_2} x_3^{k_3}) := k_1! k_2! k_3!$$

for all $k_1, k_2, k_3 \in \mathbb{N}$. This scalar product has the useful property that

Lemma 4.53. *The operator* $(f \mapsto x_i \cdot f) \in \mathrm{End}_{\mathbb{C}}(A)$ *is the adjoint of the operator* $\left(f \mapsto \frac{\partial}{\partial x_i} f\right) \in \mathrm{End}_{\mathbb{C}}(A)$ *for* $i = 1, 2, 3$.

Proof. By the symmetry, it is enough to show that

$$\left(\frac{\partial u}{\partial x_1}, v\right) = (u, x_1 \cdot v)$$

where we can assume that u and v are monomials from A with coefficient 1 . So suppose that $u = x_1^{k_1} x_2^{k_2} x_3^{k_3}$. Then $\frac{\partial u}{\partial x_1} = k_1 x_1^{k_1 - 1} x_2^{k_2} x_3^{k_3}$ so both sides are zero unless $v = x_1^{k_1-1} x_2^{k_2} x_3^{k_3}$. In that case

$$\left(\frac{\partial u}{\partial x_1}, v\right) = (k_1 x_1^{k_1-1} x_2^{k_2} x_3^{k_3}, x_1^{k_1-1} x_2^{k_2} x_3^{k_3}) = k_1! k_2! k_3! = (x_1^{k_1} x_2^{k_2} x_3^{k_3}, x_1^{k_1} x_2^{k_2} x_3^{k_3}) = (u, x_1 \cdot v)$$

so the statement follows. ☐

Corollary 4.54. *With the notations* $r^2 = x_1^2 + x_2^2 + x_3^2$ *and* $\Delta = \left(\frac{\partial}{\partial x_1}\right)^2 + \left(\frac{\partial}{\partial x_2}\right)^2 + \left(\frac{\partial}{\partial x_3}\right)^2$. *Then the operator* $(f \mapsto r^2 \cdot f) \in \mathrm{End}_{\mathbb{C}}(A)$ *is the adjoint of* Δ.

Note that these two maps are also G-equivariant. Moreover, they respect the grading: $r^2 A_{m-2} \subseteq A_m$ and $\Delta A_m \subseteq A_{m-2}$.

Definition 4.55. The *Harmonic polynomials* are the elements of $H := A \cap \operatorname{Ker} \Delta$. Similarly, the set of degree m harmonic polynomials is $H_m := A_m \cap H = A_m \cap \operatorname{Ker} \Delta$.

Why is the adjoint relation between Δ and r^2 is useful in our situation? Because $\operatorname{Ker} \Delta = (\operatorname{Im} \Delta^*)^{\perp}$ i.e. we get a decomposition

$$A_m = H_m \oplus \operatorname{Im}(r^2 \cdot) = H_m \oplus r^2 A_{m-2}$$

which is also a G-invariant decomposition. A repetition of this formula yields

$$A_m = H_m \oplus r^2 H_{m-2} \oplus r^4 H_{m-4} \oplus \dots = \bigoplus_{i=0}^{\left[\frac{m}{2}\right]} r^{2i} H_{m-2i}$$

Now, we can relate the two modules of G: the one we want to understand (namely, $\mathrm{Pol}(X)$) and the one we just spoke about (it is A). So let's define the homomorphism $\rho : A \to \mathrm{Pol}(X) \subseteq \mathcal{C}(X)$ as $f \mapsto f|_X$.

Lemma 4.56. $\mathrm{Ker} \rho \cap A_m = \{0\}$ *for all* m *so* $\rho|_{A_m}$ *is injective.*

Proof. $f \in A_m$ if and only if $f(\lambda x) = \lambda^n f(x)$ hence a homogeneous function vanishing on X vanishes everywhere. ☐

Theorem 4.57. *$A_m = H_m \oplus r^2 H_{m-2} \oplus r^4 H_{m-4} \oplus \dots$ is a decomposition of A_m into direct sum of minimal SO_3-invariant subspaces.*

Proof. We know that $A_m \cong \rho(A_m) \subseteq \mathrm{Pol}(X) \subseteq \mathcal{C}(X)$. So by Lemma 4.50 we get that the number of summands in the decomposition of $\rho(A_m)$ into the direct sum of minimal G-invariant subspaces is at most $\dim_{\mathbb{C}} \rho(A_m)^H = \dim_{\mathbb{C}} A_m^H$. About this dimension we will prove in Lemma 4.58 below, that it is $\left[\frac{m}{2}\right] + 1$ which is exactly the number of summands in the above formulated decomposition. Therefore, it is indeed a decomposition into minimal G-invariant subspaces. □

Notation: Let $h(z) \in SO_3$ be the rotation through the x_3 axis by the angle t where $z = e^{it}$. So we just parametrized $H := \{h(z) \mid z \in U_1(\mathbb{C})\}$.

The finite dimensional irreducible representation of H are 1-dimensional, namely the maps $H \to \mathbb{C}$, $h(z) \mapsto z^k$ for $k \in \mathbb{Z}$.

Lemma 4.58. *A_m has a basis of joint eigenfunctions of $H = \{h(z) \mid z \in U_1(\mathbb{C})\}$ and the multiplicity of the representation $h(z) \mapsto z^k$ of H in A_m is*

$$\left[\frac{m-|k|}{2}\right] + 1$$

for all $k \in \{0, \pm 1, \pm 2, \dots, \pm m\}$ and these are all the irreducibles appearing in A_m. In particular that trivial representation (i.e. when $k = 0$) has multiplicity $\left[\frac{m}{2}\right] + 1$, or in other words, $\dim A_m^H = \left[\frac{m}{2}\right] + 1$.

Proof. We just change the variables in a way that the action of H gets diagonalized. So let $u := x_1 - ix_2$ and $\bar{u} = x_1 + ix_2$. It is clear that it is in fact a basis change in A_1 and one can check that

$$h(z)u = zu \qquad h(z)\bar{u} = z^{-1}\bar{u} = \bar{z}\bar{u}$$

It also induces a base change in A_m where we get the basis $\{u^p \bar{u}^q x_3^l \mid p + q + l = m,\ p, q, l \in \mathbb{N}\}$ and H acts as

$$h(z)u^p \bar{u}^q x_3^l = z^{p-q} u^p \bar{u}^q x_3^l$$

so it is in fact an eigenbasis. Hence the dimensions can be directly computed by counting the number of the eigenfunction for a given eigenvalue. One can see that $|p - q| \le p + q = m - l \le m$ so there are no other representations in the decompositions than the stated ones. Moreover, for $k \in \{0, \pm 1, \pm 2, \dots, \pm m\}$ we have

$$|\{(p, q, l) \mid p + q + l = m,\ p, q, l \in \mathbb{N},\ p - q = k\}| = \left[\frac{1}{2}(m - |k|)\right] + 1$$

as we stated. □

Theorem 4.59. *The SO_3-module $\mathrm{Pol}(X)$ decomposes into the orthogonal direct sum of the minimal SO_3-invariant subspaces $U_m = \rho(H_m)$ for $m \in \mathbb{N}$ where $\dim U_m = 2m + 1$ and U_m has an orthogonal basis $Y_{m,0}, Y_{m,\pm 1}, Y_{m,\pm 2}, \dots, Y_{m,\pm m}$ called the Laplace spherical functions that are characterized (up to non-zero scalar multiple) by the property that*

$$h(z)Y_{m,k} = z^k Y_{m,k}$$

Here orthogonality means that with respect to any G-invariant scalar product which exist by for example pushing the Haar integral from $\mathrm{Pol}(G)$ to $\mathrm{Pol}(X) = \mathrm{Pol}(G/H)$.

Remark 4.60. Another natural choice for a G-invariant scalar product is

$$\langle f_1, f_2 \rangle = \int_X \overline{f_1} f_2$$

where we means the standard integration on the unit sphere.

Remark 4.61. Note that we got that $\mathrm{Pol}(X)$ is not only a multiplicity-free representation but every finite dimensional irreducible representations appears exactly once in $\mathrm{Pol}(X)$.