REPRESENTATION THEORY

Contents

Preface

Basic idea: If k is a field and V is a k-vector space, then the set of k-linear automorphisms of V form a group $\mathrm{GL}(V)$ under composition. The basic idea is to study a general group G by considering group homomorphisms

$$G \xrightarrow{\pi} \mathrm{GL}(V).$$

We think of $g \in G$ as being complicated and of $\pi(g) \in \mathrm{GL}(V)$ as being easier. Indeed, we can use the methods of linear algebra to study $\pi(g) \in \mathrm{GL}(V)$.

Textbook: E. B. Vinberg, Linear representations of groups, Translated from the 1985 Russian original by A. Iacob. Reprint of the 1989 translation. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2010.ISBN: 978-3-0348-0062-4.

Schedule: The plan is to cover one chapter in the textbook each week, beginning with Chapter 0.

1. Basic definitions

Let k be a field (typically $k = \mathbb{R}$ or $k = \mathbb{C}$), and let V be a k-vector space. The group of k-linear automorphisms of V is given by the set

$$\mathrm{GL}(V) = \{f \colon V \to V \mid f \text{ is } k\text{-linear and an isomorphism}\}$$

with group structure given by the map

$$\mathrm{GL}(V) \times \mathrm{GL}(V) \xrightarrow{\ \circ\ } \mathrm{GL}(V)$$

that to (f, g) assigns $f \circ g$.

Definition 1.1. A k-linear representation of a group G is a pair (V, π) of a k-vector space V and a group homomorphism $\pi \colon G \to \mathrm{GL}(V)$.

So elements $g \in G$ are represented by k-linear operators $\pi(g) \colon V \to V$ such that

$$\pi(g \cdot h) = \pi(g) \circ \pi(h)$$

and such that

$$\pi(e) = \mathrm{id}_V .$$

Here $e \in G$ is the identity element.

Suppose that $\dim_k(V) = n < \infty$. A choice of a basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of V determines an isomorphism of groups

$$\mathrm{GL}(V) \xrightarrow{\ \alpha\ } \mathrm{GL}_n(k)$$

that to the k-linear automorphism $f \colon V \to V$ assigns the invertible $n \times n$-matrix $\alpha(f) = (a_{ij})$, whose entries $a_{ij} \in k$ are the unique solutions to the equations

$$f(\boldsymbol{e}_j) = \boldsymbol{e}_1 a_{1j} + \boldsymbol{e}_2 a_{2j} + \cdots + \boldsymbol{e}_n a_{nj}$$

for $1 \leq j \leq n$. Hence, a k-linear representation $\pi \colon G \to \mathrm{GL}(V)$ determines and is determined by the composite group homomorphism

$$G \xrightarrow{\ \pi\ } \mathrm{GL}(V) \xrightarrow{\ \alpha\ } \mathrm{GL}_n(k).$$

We stress that the group isomorphism α depends on the choice of basis! We say that the composite map $\alpha \circ \pi \colon G \to \mathrm{GL}_n(k)$ is a matrix representation of G.

Definition 1.2. A matrix representation of a group G over a field k is a group homomorphism $\pi \colon G \to \mathrm{GL}_n(k)$.

In order to do calculations, it can be convenient to choose a basis of a vector space and calculate in coordinates. However, for theoretical considerations, it is always best to avoid making a choice of basis.

If we study mathematical objects given by sets equipped with some structure, then we should at the same time study the maps between such objects that preserve this structure. If V_1 and V_2 are k-vector spaces, then the maps $f \colon V_1 \to V_2$ that preserve the structure of a k-vector space are the k-linear maps. And if X_1 and X_2 are topological spaces, then the maps $f \colon X_1 \to X_2$ that preserve the structure of a topological space are the continuous maps. The maps that preserve the structure of a k-linear representation of a fixed group G, which we now define, are called the intertwining (or equivariant) maps.

Definition 1.3. If (V_1, π_1) and (V_2, π_2) are k-linear representations of G, then an intertwining map $f\colon (V_1, \pi_1) \to (V_2, \pi_2)$ is a k-linear map $f\colon V_1 \to V_2$ such that

$$f(\pi_1(g)(\boldsymbol{x})) = \pi_2(g)(f(\boldsymbol{x}))$$

for all $g \in G$ and $\boldsymbol{x} \in V_1$.

If (V, π) is a k-linear representation of G, then we will sometime abbreviate

$$g\boldsymbol{x} = \pi(g)(\boldsymbol{x}).$$

So if both (V_1, π_1) and (V_2, π_2) are representations of the same group G, then a k-linear map $f\colon V_1 \to V_2$ is intertwining if and only if

$$f(g\boldsymbol{x}) = gf(\boldsymbol{x})$$

for all $g \in G$ and $\boldsymbol{x} \in V_1$.

Definition 1.4. Let (V_1, π_1) and (V_2, π_2) be k-linear representations of a group G. An intertwining map $f\colon (V_1, \pi_1) \to (V_2, \pi_2)$ is an isomorphism, if there exists an intertwining map $g\colon (V_2, \pi_2) \to (V_1, \pi_1)$ such that $g \circ f = \mathrm{id}_{V_1}$ and $f \circ g = \mathrm{id}_{V_2}$.

We show in the problem set that an intertwining map $f\colon (V_1, \pi_1) \to (V_2, \pi_2)$ is an isomorphism if and only if the map $f\colon V_1 \to V_2$ is a bijection.

Remark 1.5. Let (V_1, π_1) and (V_2, π_2) be two k-linear representations of a group G. We say that (V_1, π_1) and (V_2, π_2) are isomorphic and write $(V_1, \pi_1) \simeq (V_2, \pi_2)$, if there exists an isomorphism $f\colon (V_1, \pi_1) \to (V_2, \pi_2)$. However, note that it is much better to know that "the map $f\colon (V_1, \pi_1) \to (V_2, \pi_2)$ is an isomorphism" than it is to know that "(V_1, π_1) and (V_2, π_2) are isomorphic." Indeed, in the former case, the given isomorphism $f\colon (V_1, \pi_1) \to (V_2, \pi_2)$ tells us *how* to translate between the two representations, whereas in the latter case, we only know that, in principle, such a translation is possible.

We consider examples of representations and begin with the group

$$G = (\mathbb{R}, +)$$

of real numbers under addition. Given $a \in \mathbb{R}$, the exponential function

$$G \xrightarrow{\pi_a} \mathrm{GL}_1(\mathbb{R})$$

defined by $\pi_a(t) = e^{at}$ is a matrix representation. Indeed, we have

$$\pi_a(t+u) = e^{a(t+u)} = e^{at+au} = e^{at}e^{au} = \pi_a(t)\pi_a(u)$$

and

$$\pi_a(0) = e^{a0} = e^0 = 1,$$

as required. This begs the question as to whether every 1-dimensional representation $\pi\colon G \to \mathrm{GL}_1(\mathbb{R})$ of G is of this form. The answer is "Yes," provided that we require the map π to be continuous. We prove the following weaker result:

Lemma 1.6. *Let $G = (\mathbb{R}, +)$ be the additive group of real numbers. For every differentiable 1-dimensional representation $\pi\colon G \to \mathrm{GL}_1(\mathbb{R})$, there exists a unique $a \in \mathbb{R}$ such that $\pi = \pi_a\colon G \to \mathrm{GL}_1(\mathbb{R})$, namely, $a = \pi'(0)$.*

Proof. We will assume the stronger hypothesis that π be differentiable instead of continuous. That π is a representation means that $\pi(0) = 1$ and that for all $t, u \in \mathbb{R}$,

$$\pi(t + u) = \pi(t) \cdot \pi(u).$$

We differentiate the latter equation with respect to u at $u = 0$, which gives the ordinary differential equation

$$\pi'(t) = \pi(t) \cdot \pi'(0).$$

Every solution to the ODE is of the form $\pi(t) = Ce^{at}$, where $a = \pi'(0) \in \mathbb{R}$, and the initial condition $\pi(0) = 1$ implies that $C = 1$. This proves the lemma in the case, where π is differentiable. □

We let $M_n(k)$ be the set of $n \times n$-matrices with entries in k, considered as a ring under matrix addition and matrix multiplication.[1] If $k = \mathbb{R}$ or $k = \mathbb{C}$, then we have the matrix exponential of $A \in M_n(k)$ defined by the series

$$e^A = \sum_{n \geq 0} \frac{1}{n!} A^n \in M_n(k),$$

which converges in operator norm, because

$$\|\frac{1}{n!} A^n\| \leq \frac{1}{n!} \|A\|^n \leq e^{\|A\|}.$$

If $AB = BA$, then $e^{A+B} = e^A e^B$, but this is generally *not* true without this assumption! In particular, the map

$$G = (\mathbb{R}, +) \xrightarrow{\pi_A} \mathrm{GL}_n(k)$$

defined by $\pi_A(t) = e^{tA}$ is a group homomorphism, and hence, an n-dimensional matrix representation of G, where $k = \mathbb{R}$ or $k = \mathbb{C}$.

Lemma 1.7. *If $\pi \colon G = (\mathbb{R}, +) \to \mathrm{GL}_n(k)$ is a differentiable real or complex representation, then $\pi = \pi_A \colon G \to \mathrm{GL}_n(k)$ with $A = \pi'(0) \in M_n(k)$.*

Proof. As before, we obtain the ordinary differential equation

$$\pi'(t) = \pi(t) \cdot \pi'(0)$$

with the initial condition $\pi(0) = E \in M_n(k)$, and it has $\pi = \pi_A$ with $A = \pi'(0)$ as its unique solution. □

Example 1.8. We consider

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

and calculate

$$A^n = \begin{cases} (-1)^m E & \text{if } n = 2m \text{ is even} \\ (-1)^m A & \text{if } n = 2m + 1 \text{ is odd,} \end{cases}$$

[1] The book writes $L_n(k)$ instead of $M_n(k)$.

which shows that

$$\begin{aligned}
e^{tA} &= \sum_{n\geq 0,\ \text{even}} \frac{1}{n!}(tA)^n + \sum_{n\geq 0,\ \text{odd}} \frac{1}{n!}(tA)^n \\
&= \sum_{m\geq 0} \frac{(-1)^m t^{2m}}{(2m)!} E + \sum_{m\geq 0} \frac{(-1)^m t^{2m+1}}{(2m+1)!} A \\
&= \cos t \cdot E + \sin t \cdot A \\
&= \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.
\end{aligned}$$

So we conclude that the map $\pi\colon G = (\mathbb{R}, +) \to \mathrm{GL}_2(\mathbb{R})$ defined by

$$\pi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

is a 2-dimensional real representation of G.

In general, if (V, π) is a k-linear representation of a group G, then we call

$$\ker(\pi) = \{g \in G \mid \pi(g) = \mathrm{id}_V\} \subset G$$

the kernel of π. It is a normal subgroup of G. The representation (V, π) will obviously not be of any help to study the elements in $\ker(\pi) \subset G$. We recall that $\ker(\pi) = \{e\}$ if and only if $\pi\colon G \to \mathrm{GL}(V)$ is injective.

Definition 1.9. A k-linear representation (V, π) of a group G is faithful if the group homomorphism $\pi\colon G \to \mathrm{GL}(V)$ is injective.

Example 1.10. (1) The representation $\pi_a\colon (\mathbb{R}, +) \to \mathrm{GL}_1(\mathbb{R})$ defined by $\pi_a(t) = e^{at}$ is faithful if and only if $a \neq 0$.

(2) The representation $\pi\colon (\mathbb{R}, +) \to \mathrm{GL}_2(\mathbb{R})$ defined by

$$\pi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

is not faithful, since $\ker(\pi) = 2\pi\mathbb{Z} \subset \mathbb{R}$.

Example 1.11. We next let $G = S_n$ be the (finite) symmetric group on n letters. It is defined to be the set of all bijections

$$\{1, 2, \dots, n\} \xrightarrow{\ \sigma\ } \{1, 2, \dots, n\}$$

equipped with the group structure

$$S_n \times S_n \xrightarrow{\ \circ\ } S_n$$

that to (σ, τ) assigns the composite bijection $\sigma \circ \tau$.

If k is any field, then we define the n-dimensional matrix representation

$$S_n \xrightarrow{\ P\ } \mathrm{GL}_n(k)$$

to be the map that to $\sigma \in S_n$ assigns the permutation matrix[2]

$$P(\sigma) = \begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \dots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix} \in \mathrm{GL}_n(k).$$

[2] One much check that $P(\sigma \circ \tau) = P(\sigma) \cdot P(\tau)$ and that $P(e) = E$, which is not difficult, but we will give a high-tech proof later in Example 1.13.

Clearly, the kernel of P is trivial, so $P\colon S_n \to \mathrm{GL}_n(k)$ is a faithful representation. We recall that the determinant defines a group homomorphism

$$\mathrm{GL}_n(k) \xrightarrow{\det} \mathrm{GL}_1(k),$$

and therefore, the composite map

$$S_n \xrightarrow{P} \mathrm{GL}_n(k) \xrightarrow{\det} \mathrm{GL}_1(k)$$

is a 1-dimensional matrix representation of S_n. It is called the sign representation, since the sign of σ, by definition, is given by

$$\mathrm{sgn}(\sigma) = \det(P(\sigma)) \in \{\pm 1\} \subset \mathrm{GL}_1(k).$$

If $2 \neq 0$ in k, then the kernel $\ker(S_n) = A_n \subset S_n$ is the alternating group on n letters. In particular, the sign representation is not faithful, except in trivial cases.

We next consider the regular representation. If X is any set and k is a field, then we view the set of all maps $f\colon X \to k$ as a k-vector space $k[X]$ with vector sum and scalar multiplication defined by

$$(f+g)(x) = f(x) + g(x)$$
$$(a \cdot f)(x) = a \cdot f(x).$$

The k-linear representation $(k[G], L)$ of a group G, where $L\colon G \to k[G]$ is given by

$$L(g)(f)(x) = f(g^{-1}x),$$

is called the left regular representation, and the k-linear representation $(k[G], R)$ of G, where $R\colon G \to \mathrm{GL}(k[G])$ is given by

$$R(g)(f)(x) = f(xg),$$

is called the right regular representation. We check that $(k[G], L)$ and $(k[G], R)$ are representations of G. First, we clearly have

$$L(e) = \mathrm{id}_{k[G]} = R(e),$$

and second, the calculations

$$\begin{aligned} L(gh)(f)(x) &= f((gh)^{-1}x) = f(h^{-1}g^{-1}x) = L(h)(f)(g^{-1}x) \\ &= L(g)(L(h)(f))(x) = (L(g) \circ L(h))(f)(x) \\ R(gh)(f)(x) &= f(xgh) = R(h)(f)(xg) \\ &= R(g)(R(h)(f))(x) = (R(g) \circ R(h))(f)(x) \end{aligned}$$

show that $L(gh) = L(g) \circ L(h)$ and $R(gh) = R(g) \circ R(h)$ as required. The left regular representations give rise to a representation on a subspace $V \subset k[X]$, provided that V is G-invariant in the sense that $L(g)(V) \subset V$ for all $g \in G$. Similarly, for the right regular representation.

Example 1.12. If $G = (\mathbb{R}, +)$, then we have

$$L(t)(f)(x) = f(-t + x) = f(x - t),$$

so the following subspaces are G-invariant:

$$\begin{aligned} V &= \{f \in k[G] \mid f \text{ is a polynomial function}\} \subset k[G], \\ W &= \mathrm{span}(\cos, \sin) \subset \mathbb{R}[G]. \end{aligned}$$

In the case of $W \subset \mathbb{R}[G]$, we have

$$L(t)(\cos)(x) = \cos(-t+x) = \cos t \cos x + \sin t \sin x$$
$$L(t)(\sin)(x) = \sin(-t+x) = -\sin t \cos x + \cos t \sin x,$$

so we recover the representation $\pi \colon G \to \mathrm{GL}_2(\mathbb{R})$ given by

$$\pi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

Finally, we consider premutation representations. Let X be any set, and let $S(X)$ be the group of all bijections $\sigma \colon X \to X$ with the composition group structure defined by $(\sigma \circ \tau)(x) = \sigma(\tau(x))$. For example, the group $S_n = S(\{1, 2, \dots, n\})$ is the symmetric group on n letters. If k is any field, then we may define a k-linear representation $(k[X], \pi)$ of $S(X)$ by [3]

$$\pi(\sigma)(f)(x) = f(\sigma^{-1}(x)).$$

A left action by a group G on a set X is defined to be a group homomorphism $\rho \colon G \to S(X)$. Thus, given a left action by G on X, the composite map

$$G \xrightarrow{\rho} S(X) \xrightarrow{\pi} \mathrm{GL}(k[X])$$

defines a k-linear representation $(k[X], \pi \circ \rho)$ of the group G. We say that a k-linear representation of this form is a permutation representation.

Example 1.13. The identity map $\rho \colon S_n \to S(\{1, 2, \dots, n\})$ is a left action, where

$$\rho(\sigma)(i) = \sigma(i).$$

So we obtain the permutation representation

$$S_n \xrightarrow{\pi = \pi \circ \rho} \mathrm{GL}(k[\{1, 2, \dots, n\}]).$$

Let us calculate the corresponding matrix representation with respect to the basis

$$(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \dots, \boldsymbol{e}_n^*)$$

of $k[\{1, 2, \dots, n\}]$, where $\boldsymbol{e}_i^* \colon \{1, 2, \dots, n\} \to k$ is the map defined by

$$\boldsymbol{e}_i^*(j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

By the definition of π, we have

$$\begin{aligned} \pi(\sigma)(\boldsymbol{e}_i^*)(j) &= \boldsymbol{e}_i^*(\sigma^{-1}(j)) \\ &= \begin{cases} 1 & \text{if } i = \sigma^{-1}(j) \\ 0 & \text{if } i \neq \sigma^{-1}(j) \end{cases} \\ &= \begin{cases} 1 & \text{if } \sigma(i) = j \\ 0 & \text{if } \sigma(i) \neq j \end{cases} \\ &= \boldsymbol{e}_{\sigma(i)}^*(j), \end{aligned}$$

which shows that

$$\pi(\sigma)(\boldsymbol{e}_i^*) = \boldsymbol{e}_{\sigma(i)}^*.$$

[3] The book writes $\sigma_*(f)$ instead of $\pi(\sigma)(f)$.

Hence, we conclude that the matrix that represents the k-linear map

$$k[\{1, 2, \ldots, n\}] \xrightarrow{\pi(\sigma)} k[\{1, 2, \ldots, n\}]$$

with respect to the basis $(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \ldots, \boldsymbol{e}_n^*)$ is the permutation matrix

$$P(\sigma) = \begin{pmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \ldots & \boldsymbol{e}_{\sigma(n)} \end{pmatrix} \in \mathrm{GL}_n(k).$$

So we recover the matrix representation

$$S_n \xrightarrow{P} \mathrm{GL}_n(k)$$

from Example 1.11. In particular, we may conclude that the identities

$$P(\sigma \circ \tau) = P(\sigma) \cdot P(\tau)$$

and $P(e) = E$ hold.

In this lecture, we define semisimple representations and we prove three theorems that we will use repeatedly to show that various representations are semisimple. We apply these theorems to three important examples, all of which are representations of the group $G = \mathrm{GL}(V)$ with V a finite dimensional k-vector space.

Definition 2.1. Let (V, π) be a k-linear representation of a group G. A subspace $U \subset V$ is said to be π-invariant if for all $g \in G$ and $\boldsymbol{u} \in U$, $\pi(g)(\boldsymbol{u}) \in U$.

We note that the subspaces $U = \{0\} \subset V$ and $U = V \subset V$ always are π-invariant.

Example 2.2. Let $G = (\mathbb{R}, +)$ be the additive group of real numbers, and let $(\mathbb{R}[G], L)$ be the left regular representation of G on the real vector space $\mathbb{R}[G]$ of all functions $f \colon \mathbb{R} \to \mathbb{R}$, which, we recall, is defined by

$$L(t)(f)(x) = f(-t + x).$$

We say that a function $f \colon \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) = \sum_{0 \leq n \leq d} a_n x^n$$

with $a_0, a_1, \dots, a_d \in \mathbb{R}$ is a polynomial function of degree $\leq d$, and we claim that the subspace $U_d \subset \mathbb{R}[G]$ of polynomial functions of degree $\leq d$ is L-invariant. Indeed, for $t \in G$ and $f \colon \mathbb{R} \to \mathbb{R}$ as above, we calculate that

$$\begin{aligned} L(t)(f)(x) = f(-t + x) &= \sum_{0 \leq n \leq d} a_n (-t + x)^n \\ &= \sum_{0 \leq n \leq d} a_n \Big(\sum_{0 \leq i \leq n} (-t)^{n-i} x^i \Big) = \sum_{0 \leq i \leq d} \Big(\sum_{i \leq n \leq d} a_n (-t)^{n-i} \Big) x^i, \end{aligned}$$

which shows that $L(t)(f) \in U_d$, as required.

Remark 2.3. Suppose that (V, π) is a k-linear representation of a group G with $\dim_k(V) < \infty$, and let $U \subset V$ be a subspace. We first choose a basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ of U, and then extend it to a basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \dots, \boldsymbol{e}_{m+n})$ of V. In this situation, the subspace $U \subset V$ is π-invariant if and only if the matrix that represents the k-linear map $\pi(g) \colon V \to V$ with respect to this basis is of the form

$$\left(\begin{array}{c|c} A(g) & B(g) \\ \hline O & D(g) \end{array} \right)$$

with $A(g) \in M_m(k)$, $B(g) \in M_{m,n}(k)$, and $D(g) \in M_n(k)$.

We recall from algebra that if V is a k-vector space and $U \subset V$ is a subspace, then the quotient vector space V/U is defined to be the set

$$V/U = \{\boldsymbol{v} + U \subset V \mid \boldsymbol{v} \in V\}$$

equipped with the vector sum $(\boldsymbol{v} + U) + (\boldsymbol{v}' + U) = (\boldsymbol{v} + \boldsymbol{v}') + U$ and the scalar multiplication $(\boldsymbol{v} + U) \cdot a = (\boldsymbol{v} \cdot a) + U$. Moreover, a k-linear map $f \colon V \to V$ with the property that $f(U) \subset U$ gives rise to a k-linear map

$$V/U \xrightarrow{f/U} V/U$$

defined by $(f/U)(\boldsymbol{v}+U) = f(\boldsymbol{v})+U$. If $f_1, f_2 \colon V \to V$ are two such maps, then we have $(f_1 \circ f_2)/U = (f_1/U) \circ (f_2/U)$. Hence, if $\mathrm{GL}(V,U) \subset \mathrm{GL}(V)$ is the subgroup of k-linear automorphisms $f \colon V \to V$ such that $f(U) \subset U$, then the map

$$\mathrm{GL}(V,U) \xrightarrow{-/U} \mathrm{GL}(V/U)$$

that to $f \colon V \to V$ assigns $f/U \colon V/U \to V/U$ is a group homomorphism.[4]

Definition 2.4. Let (V,π) is a k-linear representation of a group G, and let $U \subset V$ is a π-invariant subspace.

(1) The representation (U, π_U), where $\pi_U \colon G \to \mathrm{GL}(U)$ is defined by

$$\pi_U(g)(\boldsymbol{u}) = \pi(g)(\boldsymbol{u})$$

for $\boldsymbol{u} \in U$, is called the subrepresentation of (V,π) on U.

(2) The representation $(V/U, \pi_{V/U})$, where $\pi_{V/U} \colon G \to \mathrm{GL}(V/U)$ is defined by

$$\pi_{V/U}(g)(\boldsymbol{v}+U) = \pi(g)(\boldsymbol{v}) + U$$

for $\boldsymbol{v}+U \in V/U$, is called the quotient representation of (V,π) on V/U.

It is common to abuse language and simply say that π_U is a subrepresentation of π and that $\pi_{V/U}$ is a quotient representation of π.

Remark 2.5. Let (V,π) be a k-linear representation of a group G, and let $U \subset V$ be a π-invariant subspace. Suppose that $\dim_k(V) < \infty$. If we choose a basis $(\boldsymbol{e}_1, \dots \boldsymbol{e}_m)$ of U and extend it to a basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \dots, \boldsymbol{e}_{m+n})$ of V, then the family $(\boldsymbol{e}_{m+1}+U, \dots, \boldsymbol{e}_{m+n}+U)$ is a basis of V/U, and moreover, the matrices that represent the maps $\pi_U(g) \colon U \to U$ and $\pi_{V/U}(g) \colon V/U \to V/U$ with respect to these bases are $A(g)$ and $D(g)$, if $\pi(g) \colon V \to V$ is represented by the matrix

$$\left(\begin{array}{c|c} A(g) & B(g) \\ \hline O & D(g) \end{array}\right)$$

with respect to the basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \dots, \boldsymbol{e}_{m+n})$.

Definition 2.6. A k-linear representation (V,π) of a group G is irredicible (or simple) if $V \neq \{0\}$ and if the only π-invariant subspaces of V are $\{0\} \subset V$ and $V \subset V$.[5]

We note the formal similarity of the definition of an irreducible representation to the definition of a prime number.

Example 2.7. (1) Every 1-dimensional representation if irreducible. In particular, the trivial representation of G on k given by the constant map $\pi \colon G \to \mathrm{GL}(k)$ that to every $g \in G$ assigns $\mathrm{id}_k \in \mathrm{GL}(k)$ is irreducible.

(2) The identity representation $\pi = \mathrm{id}_{\mathrm{GL}(V)} \colon G = \mathrm{GL}(V) \to \mathrm{GL}(V)$ is irreducible.

(3) The 2-dimensional representation $\pi \colon G = (\mathbb{R},+) \to \mathrm{GL}_2(\mathbb{R})$ given by

$$\pi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

[4] It is also common to write $\bar{f}$ instead of f/U.

[5] The assumption $V \neq \{0\}$ is missing in the book.

is irreducible. Indeed, the map $\pi(t)\colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by counterclockwise rotation through t radians around the origin, so it leaves no line through the origin invariant, unless $t \in \ker(\pi) = 2\pi\mathbb{Z}$.

(4) Let $G = (\mathbb{R}, +)$, and let (U_d, π_{U_d}) be the subrepresentation of the left regular representation $(\mathbb{R}[G], L)$ from Example 2.2. If $d \geq 1$, then π_{U_d} is not irreducible, since any $U_e \subset U_d$ with $0 \leq e < d$ is π_{U_d}-invariant and $\{0\} \subsetneq U_e \subsetneq U_d$.

(5) Let $G = S_n$ be the symmetric group on n letters, and let $\pi\colon G \to \mathrm{GL}_n(k)$ be the standard permutation representation on $V = k^n$. The subspaces

$$\begin{aligned} V_0 &= \{\boldsymbol{x} \in V \mid \textstyle\sum_{1\leq i\leq n} x_i = 0\} \subset V \\ V_1 &= k \cdot (1, 1, \dots, 1) \subset V \end{aligned}$$

are π-invariant subspaces of dimension $n-1$ and 1, respectively. Moreover, if $\mathrm{char}(k)$ does not divide n, then $V_1 \cap V_0 = \{\boldsymbol{0}\}$, so in this case, the intertwining map

$$V_0 \oplus V_1 \longrightarrow V$$

induced by the canonical inclusions is an isomorphism. We claim that both the subrepresentations $\pi_0 = \pi|_{V_0}$ and $\pi_1 = \pi|_{V_1}$ are irreducible. This is clear for π_1, since $\dim_k(V_1) = 1$. To prove that also π_0 is irreducible, we let $\{\boldsymbol{0}\} \neq U \subset V_0$ be a π_0-invariant subspace and prove that $U = V_0$. We choose a nonzero vector

$$\boldsymbol{x} = \textstyle\sum_{1\leq i\leq n} \boldsymbol{e}_i x_i \in U.$$

Since $\boldsymbol{x} \notin V_1$, the coordinates x_i are not all equal, and since $U \subset V_0$ is π_0-invariant, we may assume that $x_1 \neq x_2$. But then

$$\pi_0((12))(\boldsymbol{x}) - \boldsymbol{x} = (\boldsymbol{e}_1 - \boldsymbol{e}_2)(x_2 - x_1) \in U,$$

so $\boldsymbol{e}_1 - \boldsymbol{e}_2 \in U$. Again, since $U \subset V_0$ is π_0-invariant, it follows that $\boldsymbol{e}_i - \boldsymbol{e}_j \in U$, for all $1 \leq i < j \leq n$. But this shows that

$$V_0 = \mathrm{span}_k(\boldsymbol{e}_i - \boldsymbol{e}_j \mid 1 \leq i < j \leq n) \subset U,$$

so we conclude that $U = V_0$. Hence, π_0 is irreducible as claimed.

Definition 2.8. A k-linear representation (V, π) of a group G is completely reducible if for every π-invariant subspace $U \subset V$, there exists a π-invariant subspace $W \subset V$ such that the map induced by the canonical inclusions

$$U \oplus W \longrightarrow V$$

is an isomorphism.

If $U, W \subset V$ are as in the definition, then we say that $W \subset V$ is a π-invariant complement of $U \subset V$. We note, in this situation, that the composition

$$W \xrightarrow{\;i\;} V \xrightarrow{\;p\;} V/U$$

of the canonical inclusion and the canonical projection is a k-linear isomorphism, which intertwines between π_W and $\pi_{V/U}$. We also remark that if a π-invariant complement $W \subset V$ of $U \subset V$ exists, then it is typically *not* unique.

Remark 2.9. If (V, π) is a k-linear representation of a group G with $\dim_k(V) < \infty$, then a π-invariant subspace $U \subset V$ admits a π-invariant complement if and only if we can find bases $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ of U and $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \dots, \boldsymbol{e}_{m+n})$ of V such that for all $g \in G$, the matrix that represents $\pi(g)\colon V \to V$ with respect to the latter basis is a block matrix

$$\left(\begin{array}{c|c} A(g) & O \\ \hline O & D(g) \end{array}\right).$$

Example 2.10. We consider two representations $(\mathbb{R}^2, \pi_A)$ of the form

$$G = (\mathbb{R}, +) \xrightarrow{\pi_A} \mathrm{GL}_2(\mathbb{R})$$

where $\pi_A(t) = e^{tA}$ with $A \in M_2(\mathbb{R})$.

We first let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Since $A^2 = O$, we have

$$\pi_A(t) = e^{tA} = E + tA = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

It follows that the subspace

$$U = \mathrm{span}_{\mathbb{R}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \subset \mathbb{R}^2,$$

is π_A-invariant but has no π_A-invariant complement. Therefore, the representation π_A is not completely reducible.

We next let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

so that $A^2 = A$. It follows that

$$\pi_A(t) = e^{tA} = E + (e^t - 1)A = \begin{pmatrix} e^t & e^t - 1 \\ 0 & 1 \end{pmatrix}.$$

In this case, the same subspace

$$U = \mathrm{span}_{\mathbb{R}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \subset \mathbb{R}^2$$

is π_A-invariant, but it now has the π_A-invariant complement

$$W = \mathrm{span}_{\mathbb{R}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \subset \mathbb{R}^2.$$

Moreover, the subspaces $U, W \subset V = \mathbb{R}^2$ are the only 1-dimensional π_A-invariant subspaces, so we conclude that π_A is completely reducible; compare Theorem 2.13.

We now prove three theorems that we will use repeatedly. The theorems are listed as Theorem 1, 2, and 3 in Chapter 1 of the book.

Theorem 2.11. *Let (V, π) be a k-linear representation of a group G, and let $U \subset V$ be a π-invariant subspace. If π is completely reducible, then so is π_U.*

Proof. Let $U_1 \subset U$ be a π_U-invariant subspace. We must show that $U_1 \subset U$ admits a π_U-invariant complement $W_1 \subset U$. Now, since $U_1 \subset V$ is π-invariant, there exists, by the assumption that π is completely reducible, a π-invariant subspace $W \subset V$ such that the map induced by the canonical inclusions

$$U_1 \oplus W \longrightarrow V$$

is an isomorphism. But then

$$U = (U_1 + W) \cap U = (U_1 \cap U) + (W \cap U) = U_1 + (W \cap U),$$

so $W_1 = W \cap U \subset U$ is a π_U-invariant subspace, and the map

$$U_1 \oplus W_1 \longrightarrow U$$

induced by the canonical inclusions is an isomorphism. This shows that $W_1 \subset U$ is a π_U-complement of $U_1 \subset U$ as desired. □

Theorem 2.12. *Let (V, π) be a completely reducible k-linear representation of a group G with $\dim_k(V) < \infty$. There exists π-invariant subspaces $V_1, \dots, V_m \subset V$ such that the map induced by the canonical inclusions*

$$V_1 \oplus \cdots \oplus V_m \longrightarrow V$$

is an isomorphism and such that $\pi_{V_1}, \dots, \pi_{V_m}$ are irreducible.

Proof. We argue by induction on $n = \dim_k(V)$. If $n = 0$, then the statement is trivial, so we assume, inductively, that the statement has been proved for $n < r$ and prove it for $n = r$. We claim that there exists a π-invariant subspace $V_1 \subset V$ such that π_{V_1} is irreducible. Granting the claim, there exists, by the assumption that π is completely reducible, a π-invariant complement $W \subset V$ of $V_1 \subset V$, and since $\dim_k(V_1) \geq 1$, we have

$$\dim_k(W) = \dim_k(V) - \dim_k(V_1) < r.$$

So by the inductive hypothesis, there exist π_W-invariant subspaces $V_2, \dots, V_m \subset W$ such that the map induced by the canonical inclusions

$$V_2 \oplus \cdots \oplus V_m \longrightarrow W$$

is an isomorphism and such that $\pi_{V_2}, \dots, \pi_{V_m}$ are irreducible. It follows that the map induced by the canonical inclusions

$$V_1 \oplus V_2 \oplus \cdots \oplus V_m \longrightarrow V$$

is an isomorphism and the subrepresentations $\pi_{V_1}, \pi_{V_2}, \dots, \pi_{V_m}$ all are irreducible, which proves the induction step. It remains to prove the claim. The set S of nonzero π-invariant subspaces $U \subset V$ is partially ordered under inclusion. It is nonempty, since $V \in S$, and it has a minimal element, since $\dim_k(V) = r < \infty$. Let $V_1 \in S$ be such a smallest element.[6] If $\{0\} \neq U \subset V_1$ is a π_{V_1}-invariant subspace, then we necessarily have $U = V_1$, since otherwise $U \in S$ is smaller than $V_1 \in S$. This shows that π_{V_1} is irreducible, which proves the claim. □

[6] In general, a minimal element $V_1 \in S$ is not unique.

Theorem 2.13. *Let (V, π) be a k-linear representation of a group G, and suppose that there exist π-invariant subspaces $V_1, \dots, V_m \subset V$ such that*

$$V = V_1 + \cdots + V_m$$

and such that $\pi_{V_1}, \dots \pi_{V_m}$ are irreducible. If $U \subset V$ is a π-invariant subspace, then there exist $\{i_1, \dots, i_p\} \subset \{1, \dots, m\}$ such that the map

$$U \oplus V_{i_1} \oplus \cdots \oplus V_{i_p} \longrightarrow V$$

induced by the canonical inclusions is an isomorphism. In particular, π is completely reducible.

Proof. We let S be the set of subsets $\{i_1, \dots, i_p\} \subset \{1, \dots, m\}$ with the property that the map induced by the canonical inclusions

$$U \oplus V_{i_1} \oplus \cdots \oplus V_{i_p} \longrightarrow V$$

is injective. The set S is partially ordered under inclusion. It is nonempty, since $\emptyset \in S$, and it is finite, since there are only finitely many subsets of $\{1, \dots, m\}$, and therefore, it has a maximal element. So we let $\{i_1, \dots, i_p\} \in S$ be a maximal element and prove that map induced by the canonical inclusions

$$U \oplus V_{i_1} \oplus \cdots \oplus V_{i_p} \longrightarrow V$$

is an isomorphism. By the definition of S, we know that the map is injective, so we only need to show that the map is surjective, or equivalently, that

$$V = U + V_{i_1} + \cdots + V_{i_p}.$$

Moreover, since $V = V_1 + \cdots + V_m$, it suffices to show that

$$V_i \subset U + V_{i_1} + \cdots + V_{i_p}$$

for all $1 \leq i \leq m$. If $i \in \{i_1, \dots, i_p\}$, then there is nothing to prove, so suppose that $i \notin \{i_1, \dots, i_p\}$. We consider the maps

$$U \oplus V_{i_1} \oplus \cdots \oplus V_{i_p} \oplus V_i \longrightarrow (U + V_{i_1} + \cdots + V_{i_p}) \oplus V_i \longrightarrow V$$

induced by the canonical inclusions. Since $\{i_1, \dots, i_p\} \in S$, the left-hand map is an isomorphism, and since $\{i_1, \dots, i_p\} \in S$ is maximal, the composite map is *not* injective, so we conclude that the right-hand map is not injective. Therefore, its kernel, which is equal to

$$(U + V_{i_1} + \cdots + V_{i_p}) \cap V_i \subset V_i$$

is nonzero. But π_{V_i} is irreducible, so this implies that

$$(U + V_{i_1} + \cdots + V_{i_p}) \cap V_i = V_i,$$

so $V_i \subset U + V_{i_1} + \cdots + V_{i_p}$ as desired. □

Remark 2.14. A representation (V, π) is defined to be semisimple, if there exists a finite number of π-invariant subspaces $V_1, \dots, V_m \subset V$ such that the map

$$V_1 \oplus \cdots \oplus V_m \longrightarrow V$$

induced by the canonical inclusions is an isomorphism and such that each of the subrepresentations π_{V_i} is irreducible. Thus, Theorems 2.12 and 2.13 shows that a finite dimensional representation is semisimple if and only if it is completely reducible.

Corollary 2.15. *Let (V, π) be a k-linear representation of a group G, and let $U \subset V$ be a π-invariant subspace. If $\dim_k(V) < \infty$ and if π is completely reducible, then also the quotient representation $\pi_{V/U}$ is completely reducible.*

Proof. By Theorem 2.12, there exists π-invariant subspaces $V_1, \dots, V_m \subset V$ such that the map induced by the canonical inclusions

$$V_1 \oplus \cdots \oplus V_m \longrightarrow V$$

is an isomorphism and such that each π_{V_i} is irreducible. We let $\overline{V}_i \subset V/U$ be the image of the composition

$$V_i \longrightarrow V \longrightarrow V/U$$

of the canonical inclusion and the canonical projection and note that $\overline{V}_i$ is zero if and only if $V_i \subset U$. So let $S = \{i_1, \dots, i_p\} \subset \{1, \dots, m\}$ be the subset consisting those $i \in \{1, \dots, m\}$ for which $V_i \not\subset U$. The subspace $V_i \cap U \subset V_i$ is π_{V_i}-invariant, so if $i \in S$, then $V_i \cap U = \{0\}$, because π_{V_i} is irreducible. Therefore, if $i \in S$, then the canonical map $V_i \to \overline{V}_i$ is an isomorphism. This shows that the $\pi_{V/U}$-invariant subspaces $\overline{V}_{i_1}, \dots, \overline{V}_{i_p} \subset V/U$ satisfy the hypothesis of Theorem 2.13, we conclude that $\pi_{V/U}$ is completely reducible, as stated. □

We consider three examples, in all of which $G = \mathrm{GL}(V)$ with V a k-vector space of finite dimension n. We first consider the k-vector space $\mathrm{End}_k(V)$[7] of all k-linear maps $f \colon V \to V$ with vector sum and scalar multiplication defined by

$$(f_1 + f_2)(\boldsymbol{v}) = f_1(\boldsymbol{v}) + f_2(\boldsymbol{v})$$
$$(a \cdot f)(\boldsymbol{v}) = a \cdot f(\boldsymbol{v}).$$

We consider the representation of G on $\mathrm{End}_k(V)$ by left multiplication:

Proposition 2.16. *Let V be a k-vector space of finite dimension n, and define*

$$G = \mathrm{GL}(V) \xrightarrow{\ \lambda\ } \mathrm{GL}(\mathrm{End}_k(V))$$

by $\lambda(g)(f) = g \circ f$. The representation $(\mathrm{End}_k(V), \lambda)$ is completely reducible.

Proof. We choose a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V and define

$$L_j = \{f \in \mathrm{End}_k(V) \mid f(\boldsymbol{v}_i) = \boldsymbol{0} \text{ for } i \neq j\} \subset \mathrm{End}_k(V).$$

It is a λ-invariant subspace. Indeed, if $g \in G$ and $f \in L_j$, then

$$\lambda(g)(f)(\boldsymbol{v}_i) = g(f(\boldsymbol{v}_i)) = \boldsymbol{0}$$

for $i \neq j$, because g is k-linear, so $\lambda(g)(f) \in L_j$. Moreover, the map

$$L_j \xrightarrow{\ h_j\ } V$$

defined by $h_j(f) = f(\boldsymbol{v}_j)$ is an isomorphism. It is also intertwining between λ and the identity representation of G on V. Indeed,

$$h_j(\lambda(g)(f)) = h_j(g \circ f) = (g \circ f)(\boldsymbol{v}_j) = g(f(\boldsymbol{v}_j)) = \mathrm{id}(g)(f(\boldsymbol{v}_j)) = \mathrm{id}(g)(h_j(f)).$$

[7] The book writes $L(V)$ instead of $\mathrm{End}_k(V)$.

Thus, $h_j\colon (L_j, \lambda_{L_j}) \to (V, \mathrm{id})$ is an isomorphism, and since (V, id) is irreducible, so is (L_j, λ_{L_j}). Finally, the map induced by the canonical inclusions

$$L_1 \oplus \cdots \oplus L_n \longrightarrow \mathrm{End}_k(V)$$

is an isomorphism, since every $f \in \mathrm{End}_k(V)$ can be written uniquely as

$$f = f_1 + \cdots + f_n$$

with $f_j \in \mathrm{End}_k(V)$ defined by

$$f_j(\boldsymbol{v}_i) = \begin{cases} f(\boldsymbol{v}_j) & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Hence, Theorem 2.13 shows that $(\mathrm{End}_k(V), \lambda)$ is completely reducible, as stated. □

We next consider the adjoint representation of G on $\mathrm{End}_k(V)$. It is an example of the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ of a reductive group on its Lie algebra.

Proposition 2.17. *Let V be a k-vector space of finite dimension n, and define*

$$G = \mathrm{GL}(V) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathrm{End}_k(V))$$

by $\mathrm{Ad}(g)(f) = g \circ f \circ g^{-1}$. The adjoint representation $(\mathrm{End}_k(V), \mathrm{Ad})$ is completely reducible, provided that $\mathrm{char}(k)$ does not divide n.

Proof. We let $\mathfrak{t} \subset \mathrm{End}_k(V)$ be the 1-dimensional subspace spanned by id_V, and let $\mathfrak{sl}_n \subset \mathrm{End}_k(V)$ be the subspace consisting of the k-linear maps $f\colon V \to V$ with $\mathrm{tr}(f) = 0$. Both subspaces are Ad-invariant. In the case of $\mathfrak{sl}_n$, we use the fact from linear algebra that $\mathrm{tr}(g \circ f \circ g^{-1}) = \mathrm{tr}(f)$. By our assumption that $\mathrm{char}(k)$ does not divide n, we have $\mathrm{tr}(\mathrm{id}_V) = n \neq 0 \in k$, so $\mathfrak{t} \cap \mathfrak{sl}_n = \{0\}$, and hence, the map

$$\mathfrak{t} \oplus \mathfrak{sl}_n \longrightarrow \mathrm{End}_k(V)$$

induced by the canonical inclusions is an isomorphism. It turns out that $\mathrm{Ad}_{\mathfrak{t}}$ and $\mathrm{Ad}_{\mathfrak{sl}_n}$ both are irreducible. This is trivial in the case of the $\mathfrak{t}$, but the proof for $\mathfrak{sl}_n$ is not so simple. We prove this for $n = 2$ in the appendix. So Theorem 2.13 shows that π is completely reducible. □

Finally, we consider a representation of $G = \mathrm{GL}(V)$ on the k-vector space

$$B(V) = \{f\colon V \times V \to k \mid f \text{ is } k\text{-bilinear}\} \simeq \mathrm{Hom}_k(V \otimes_k V, k)$$

of k-bilinear forms on V.

Proposition 2.18. *Let V be a k-vector space of finite dimension n, and define*

$$G = \mathrm{GL}(V) \xrightarrow{\pi} \mathrm{GL}(B(V))$$

by $\pi(g)(f)(\boldsymbol{x}, \boldsymbol{y}) = f(g^{-1}(\boldsymbol{x}), g^{-1}(\boldsymbol{y}))$. The representation $(B(V), \pi)$ is completely reducible, provided that $\mathrm{char}(k) \neq 2$.

Proof (Incomplete). Let $B^{\pm}(V) \subset B(V)$ be the subspaces of symmetric forms and skew-symmetric forms, respectively. We recall that $f \in B^{+}(V)$ if and only if

$$f(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{y}, \boldsymbol{x})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$, and that $f \in B^{-}(V)$ if and only if

$$f(\boldsymbol{x}, \boldsymbol{y}) = -f(\boldsymbol{y}, \boldsymbol{x})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$. Clearly, $B^{\pm}(V) \subset B(V)$ are both π-invariant, and if $\operatorname{char}(k) \neq 2$, then the map induced by the canonical inclusions

$$B^{+}(V) \oplus B^{-}(V) \longrightarrow B(V)$$

is an isomorphism. One can prove that $\pi_{B^{\pm}(V)}$ both are irreducible, but we will not do so here. So Theorem 2.13 shows that π is completely reducible. ☐

Appendix: Direct sum of vector spaces

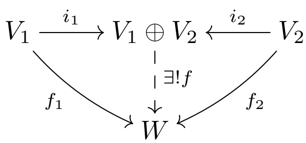
Let k be a field. If V_1 and V_2 are two k-vector spaces, then their direct sum is a triple $(V_1 \oplus V_2, i_1, i_2)$ of a k-vector space $V_1 \oplus V_2$ and two k-linear maps

$$V_1 \xrightarrow{i_1} V_1 \oplus V_2 \xleftarrow{i_2} V_2$$

with the property that if (W, f_1, f_2) is any triple of a k-vector space W and k-linear maps $f_1 \colon V_1 \to W$ and $f_2 \colon V_2 \to W$, then there exists a *unique* k-linear map

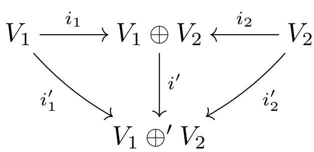
$$V_1 \oplus V_2 \xrightarrow{f} W$$

such that $f_1 = f \circ i_1$ and $f_2 = f \circ i_2$. The commutative diagram

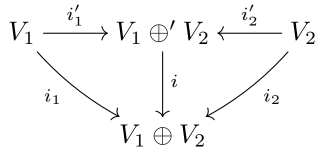


indicates this definition. In this situation, we also write $f_1 + f_2 \colon V_1 \oplus V_2 \to W$ for the unique k-linear map $f \colon V_1 \oplus V_2 \to W$.

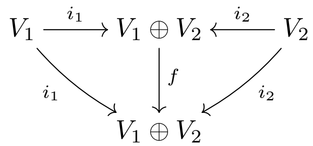
A direct sum $(V_1 \oplus V_2, i_1, i_2)$ of k-vector spaces V_1 and V_2 is not unique, but it is unique, up to *unique* isomorphism, which is just as good, or even better. Indeed, if also $(V_1 \oplus' V_2, i_1', i_2')$ is a direct sum of V_1 and V_2, then since $(V_1 \oplus V_2, i_1, i_2)$ is a direct sum of V_1 and V_2, there is a unique k-linear map i' making the diagram



commute, and similarly, since $(V_1 \oplus' V_2, i'_1, i'_2)$ is a direct sum of V_1 and V_2, there is a unique k-linear map i making the diagram



commute. It follows that both $f = i \circ i'$ makes the diagram



commute, but clearly $f = \mathrm{id}_{V_1 \oplus V_2}$ does so, too, and therefore, by the uniqueness statement in the definition of a direct sum, we conclude that $i \circ i' = \mathrm{id}_{V_1 \oplus V_2}$. The same argument shows that $i' \circ i = \mathrm{id}_{V_1 \oplus' V_2}$, so we have proved that the unique maps i and i' are each other's inverses.[8]

We will often abuse language and say that $(V_1 \oplus V_2, i_1, i_2)$ is *the* direct sum of V_1 and V_2, since we unique isomorphisms between any two choices of a direct sum. We may also abuse notation and write that $V_1 \oplus V_2$ is the direct sum of V_1 and V_2, omitting the k-linear maps i_1 and i_2 that are part of the structure.

Now, suppose that $V_1, V_2 \subset V$ are subspaces of a k-vector space V. The canonical inclusions $j_1 \colon V_1 \to V$ and $j_2 \colon V_2 \to V$ give rise to the unique k-linear map

$$V_1 \oplus V_2 \xrightarrow{\ j\ } V$$

such that $j_1 = j \circ i_1$ and $j_2 = j \circ i_2$. In general, the map j is neither surjective nor injective. The image of j is a subspace of V that we denote by

$$V_1 + V_2 \subset V,$$

and a kernel of j is given by the k-linear map[9]

$$V_1 \cap V_2 \xrightarrow{i_1 - i_2} V_1 \oplus V_2$$

that to $\boldsymbol{x} \in V_1 \cap V_2$ assigns $i_1(\boldsymbol{x}) - i_2(\boldsymbol{x}) \in V_1 \oplus V_2$. In other words, the sequence

$$0 \longrightarrow V_1 \cap V_2 \xrightarrow{i_1 - i_2} V_1 \oplus V_2 \xrightarrow{\ j\ } V_1 + V_2 \longrightarrow 0$$

of k-vector spaces and k-linear maps is exact. In particular, we see that

$$V_1 \oplus V_2 \xrightarrow{\ j\ } V$$

is an isomorphism if and only if $V_1 \cap V_2 = \{\mathbf{0}\}$ and $V_1 + V_2 = V$. Please note that we will *never* use $V_1 \oplus V_2$ to denote a subspace of V.

---

[8] This is general fact in category theory: Objects that are defined by a universal property are unique, up to *unique* isomorphism.

[9] The map $i_2 - i_1 \colon V_1 \cap V_2 \to V_1 \oplus V_2$ is also a kernel of j.

Appendix: The adjoint representation

We include a proof of the following theorem, which we used above.

Theorem 2.19. *If* $\operatorname{char}(k) \neq 2$, *then the adjoint representation*

$$\mathrm{GL}_2(k) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{sl}_2(k))$$

is irreducible.

Proof. We must show that if $U \subset \mathfrak{sl}_2(k)$ is an Ad-invariant subspace, then either $U = \{0\}$ or $U = \mathfrak{sl}_2(k)$. So we assume that $U \neq \{0\}$ and proceed to prove that $U = \mathfrak{sl}_2(k)$. We fix the basis (H, X, Y) of $\mathfrak{sl}_2(k)$, where

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

We claim that $H \in U$ if and only if $X \in U$ if and only if $Y \in U$. First, we have

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} X \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = Y, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} Y \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = X,$$

so if $X \in U$, then $Y \in U$ and vice versa. Second, we use the fact that $X^2 = 0$ so that $1 + X \in \mathrm{GL}_2(k)$ with inverse $1 - X$. Hence, the calculation

$$HX = X, \qquad XH = -X, \qquad XHX = 0$$

shows that

$$(1 + X)H(1 - X) = H - HX + XH - XHX = H - 2X.$$

Therefore, if $H \in U$, then $2X \in U$, and hence $X \in U$, since we are assuming that $2 \neq 0$ in k. Similarly, the calculation

$$YX = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad XY = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad XYX = X$$

shows that

$$(1 + X)Y(1 - X) = Y - YX + XY - XYX = Y + H - X.$$

Therefore, since we have already seen that $Y \in U$ if and only if $X \in U$, we conclude that if $Y \in U$, then $H = (Y + H - X) - Y + X \in U$. This proves the claim.

It remains to prove that at least one of H, X, and Y is in U. Since U is nonzero, there exists $0 \neq A \in \mathfrak{sl}_2(k)$. We write

$$A = aH + bX + cY = \begin{pmatrix} a & b \\ c & -a \end{pmatrix},$$

with $(a, b, c) \neq (0, 0, 0)$. For all $t \in k^*$, we have

$$g(t) = \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_2(k)$$

with inverse $g(t)^{-1} = g(t^{-1})$. Since U is assumed Ad-invariant, the calculation

$$\mathrm{Ad}(g(t))(A) = g(t)Ag(t)^{-1} = \begin{pmatrix} a & tb \\ t^{-1}c & -a \end{pmatrix} = aH + tbX + t^{-1}cY$$

shows that $aH + tbX + t^{-1}cY \in U$ for all $t \in k^*$. We wish to conclude that each of aH, bX, and cY is in U. So we wish to show that the system of linear equations

$$\begin{aligned} aH + rbX + r^{-1}cY &= aH \\ aH + sbX + s^{-1}cY &= bX \\ aH + tbX + t^{-1}cY &= cY \end{aligned}$$

has a solution with $r, s, t \in k^*$. The calculation

$$\det \begin{pmatrix} 1 & r & r^{-1} \\ 1 & s & s^{-1} \\ 1 & t & t^{-1} \end{pmatrix} = -(rst)^{-1}(r-s)(r-t)(s-t)$$

shows that a solution exists, provided that k^* has order at least three. Hence, if this is the case, then aH, bX, and cY are all in U, and since $(a,b,c) \neq (0,0,0)$, it follows that at least one of H, X, and Y is in U, so we are done.

The only missing case is $k = \mathbb{F}_3$, where $k^* = \{\pm 1\}$ only has order 2. In this case, the argument above shows that

$$\operatorname{span}_k(aH + bX + cY, aH - bX - cY) \subset U,$$

so $aH \in U$ and $bX + cY \in U$. If $a \neq 0$, then $H \in U$. Also, if $a = b = 0$, then $c \neq 0$, so $Y \in U$, and similarly, if $a = c = 0$, then $b \neq 0$, so $X \in U$. Hence, it only remains to prove that both the subspaces

$$\begin{aligned} V &= \operatorname{span}_k(\operatorname{Ad}(g)(X+Y) \mid g \in \mathrm{GL}_2(k)) \subset U \\ W &= \operatorname{span}_k(\operatorname{Ad}(g)(X-Y) \mid g \in \mathrm{GL}_2(k)) \subset U \end{aligned}$$

are equal to U. The calculation

$$\operatorname{Ad}(\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix})(X+Y) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}(X+Y)\begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix} = H$$

shows that $H \in V$, so that $V = U$, and the calculation

$$\operatorname{Ad}(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix})(X-Y) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}(X-Y)\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = -H - X - Y$$

shows that $H + X + Y \in W$, so $W = U$, since, by the argument above, $H \in W$. ☐

Remark 2.20. We note that if $\operatorname{char}(k) = 2$, then the adjoint representation

$$\mathrm{GL}_2(k) \xrightarrow{\operatorname{Ad}} \mathrm{GL}(\mathfrak{sl}_2(k))$$

is not irreducible. Indeed, since $H = 1$, the 1-dimensional subspace

$$\operatorname{span}_k(H) \subset \mathfrak{sl}_2(k)$$

is Ad-invariant.

3. Unitarity of finite dimensional complex representations

We recall from last time that a k-linear representation (V, π) of a group G is defined to be completely reducible if for every π-invariant subspace $U \subset V$, there exists a π-invariant subspace $W \subset V$ such that the map

$$U \oplus W \longrightarrow V$$

induced by the canonical inclusions is an isomorphism. In this lecture, we will show that every finite dimensional continuous real or complex representation of a compact topological group is completely reducible.

Definition 3.1. A finite dimensional real (resp. complex) representation (V, π) of a group G is orthogonal (resp. unitary), if there exists an inner product (resp. a hermitian inner product) $\langle -, - \rangle \colon V \times V \to k$ such that

$$\langle \pi(g)(\boldsymbol{v}_1), \pi(g)(\boldsymbol{v}_2) \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$$

for all $g \in G$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.[10]

Remark 3.2. If $f \colon V \to V$ is a linear endomorphism of a finite dimensional real (resp. complex) vector space with inner product (resp. hermitian inner product) $\langle -, - \rangle$, then its adjoint $f^* \colon V \to V$ is the unique linear map such that

$$\langle f^*(\boldsymbol{v}_1), \boldsymbol{v}_2 \rangle = \langle \boldsymbol{v}_1, f(\boldsymbol{v}_2) \rangle$$

for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Hence, in Definition 3.1, the requirement that

$$\langle \pi(g)(\boldsymbol{v}_1), \pi(g)(\boldsymbol{v}_2) \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$$

for all $g \in G$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ is equivalent to the requirement that

$$\pi(g)^* = \pi(g^{-1})$$

for all $g \in G$.

Definition 3.3. Let $(V, \langle -, - \rangle)$ be a finite dimensional real inner product space (resp. hermitian inner product space). The orthogonal group (resp. the unitary group) is the subgroup[11] $\mathrm{O}(V, \langle -, - \rangle) \subset \mathrm{GL}(V)$ (resp. $\mathrm{U}(V, \langle -, - \rangle) \subset \mathrm{GL}(V)$) of all k-linear maps $f \colon V \to V$ with the property that

$$\langle f(\boldsymbol{v}_1), f(\boldsymbol{v}_2) \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$$

for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.

So a finite dimensional real (resp. complex) representation (V, π) is orthogonal (resp. unitary) if and only if the group homomorphism $\pi \colon G \to \mathrm{GL}(V)$ takes values in the subgroup $\mathrm{O}(V, \langle -, - \rangle) \subset \mathrm{GL}(V)$ (resp. $\mathrm{U}(V, \langle -, - \rangle) \subset \mathrm{GL}(V)$) for some inner product (resp. hermitian inner product) $\langle -, - \rangle$ on V.

Proposition 3.4. *Every orthogonal (resp. unitary) representation is completely reducible.*

---

[10] So that (V, π) is orthogonal (resp. unitary) means that it has the *property* that such an inner product (resp. a hermitian inner product) exists. It does not mean that the *structure* of such an inner product (resp. hermitian inner product) has been chosen.

[11] Often $\mathrm{O}(V, \langle -, - \rangle)$ and $\mathrm{U}(V, \langle -, - \rangle)$ are abbreviated $\mathrm{O}(V)$ and $\mathrm{U}(V)$, but we will not do so.

Proof. We let (V, π) be an orthogonal (resp. unitary) representation of a group G and choose an inner product (resp. a hermitian inner product) $\langle -, - \rangle$ on V such that $\pi(g)^* = \pi(g^{-1})$ for all $g \in G$. If $U \subset V$ is a subspace, then its orthogonal complement with respect to $\langle -, - \rangle$ is the subspace defined by

$$U^\perp = \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{u} \rangle = 0 \text{ for all } \boldsymbol{u} \in U\} \subset V.$$

We claim that $U \subset V$ is π-invariant if and only if $U^\perp \subset V$ is π-invariant. Indeed, given any linear endomorphism $f \colon V \to V$, we have

$$f(U) \subset U \quad \Leftrightarrow \quad f^*(U^\perp) \subset U^\perp,$$

and therefore, we conclude that

$$\begin{aligned}
&U \subset V \text{ is } \pi\text{-invariant} \quad \Leftrightarrow \\
&\pi(g)(U) \subset U \text{ for all } g \in G \quad \Leftrightarrow \\
&\pi(g)^*(U^\perp) \subset U^\perp \text{ for all } g \in G \quad \Leftrightarrow \\
&\pi(g^{-1})(U^\perp) \subset U^\perp \text{ for all } g \in G \quad \Leftrightarrow \\
&\pi(g)(U^\perp) \subset U^\perp \text{ for all } g \in G \quad \Leftrightarrow \\
&U^\perp \subset V \text{ is } \pi\text{-invariant},
\end{aligned}$$

as claimed. In particular, every π-invariant subspace $U \subset V$ has a π-invariant complement, namely, $U^\perp \subset V$, so π is completely reducible. □

We first consider finite groups.

Theorem 3.5. *Every finite dimensional real (resp. complex) representation of a finite group is orthogonal (resp. unitary).*

Proof. Let (V, π) be a finite dimensional real (resp. complex) representation of a finite group G. We choose an arbitrary inner product (resp. hermitian inner product) $\langle -, - \rangle_0 \colon V \times V \to k$ and define $\langle -, - \rangle \colon V \times V \to k$ by

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \frac{1}{|G|} \sum_{x \in G} \langle \pi(x)(\boldsymbol{v}_1), \pi(x)(\boldsymbol{v}_2) \rangle_0.$$

It is easy to check that $\langle -, - \rangle$ is an inner product (resp. a hermitian inner product), and we claim that it is π-invariant. Indeed, for all $g \in G$, we have

$$\begin{aligned}
\langle \pi(g)(\boldsymbol{v}_1), \pi(g)(\boldsymbol{v}_2) \rangle &= \frac{1}{|G|} \sum_{x \in G} \langle \pi(x)(\pi(g)(\boldsymbol{v}_1)), \pi(x)(\pi(g)(\boldsymbol{v}_2)) \rangle_0 \\
&= \frac{1}{|G|} \sum_{x \in G} \langle (\pi(x) \circ \pi(g))(\boldsymbol{v}_1), (\pi(x) \circ \pi(g))(\boldsymbol{v}_2) \rangle_0 \\
&= \frac{1}{|G|} \sum_{x \in G} \langle \pi(xg)(\boldsymbol{v}_1), \pi(xg)(\boldsymbol{v}_2) \rangle_0 \\
&= \frac{1}{|G|} \sum_{y \in G} \langle \pi(y)(\boldsymbol{v}_1), \pi(y)(\boldsymbol{v}_2) \rangle_0 \\
&= \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle
\end{aligned}$$

as desired. □

Definition 3.6. A topological group is a group G with a topology such that the maps $\mu\colon G \times G \to G$ and $\iota\colon G \to G$ given by $\mu(g,h) = gh$ and $\iota(g) = g^{-1}$ are continuous. A compact group is a topological group, whose underlying topological space is compact and Hausdorff.

Example 3.7. (1) A finite group with the discrete topology is a compact group.

(2) If V is a finite dimensional real or complex vector space, then $\mathrm{GL}(V)$ is a topological group with the compact-open topology. It is a locally compact group, but it is not compact, unless $V = \{\mathbf{0}\}$.

(3) If $(V, \langle -,- \rangle)$ is a finite dimensional real inner product space (resp. hermitian inner product space), then $\mathrm{O}(V, \langle -,- \rangle)$ (resp. $\mathrm{U}(V, \langle -,- \rangle)$) is a topological group with the compact-open topology. It is a compact group.

We will only consider (real or complex) representations (V,π) of a topological group G that are continuous in the sense that the group homorphism

$$G \xrightarrow{\ \pi\ } \mathrm{GL}(V)$$

is continuous.

Example 3.8. Suppose that $(V, \langle -,- \rangle)$ be a finite dimensional real inner product space (resp. hermitian inner product space). The canonical inclusion

$$\mathrm{O}(V, \langle -,- \rangle) \xrightarrow{\ \pi\ } \mathrm{GL}(V)$$

is continuous and a group homomorphism, so (V,π) is a continuous representation.

Theorem 3.9. *Let G be a compact group, and let (V,π) be a finite dimensional continuous real (resp. complex) representation of G. Then (V,π) is orthogonal (resp. unitary), and hence, completely reducible.*

We will give two different proofs of the theorem. The first proof uses the following deep theorem. This is a important and useful theory, but it will take us too far afield to prove it here. A proof can be found in [2, Chapter 7, §1, No. 2, Theorem 1].

Theorem 3.10. *Let G be a compact group. There exists a map*

$$\begin{aligned} C^0(G,\mathbb{C}) &\longrightarrow \mathbb{C} \\ f &\longmapsto \textstyle\int_G f(x)dx \end{aligned}$$

with the following properties:

(1) *It is linear.*
(2) *It is positive in the sense that if $f \in C^0(G,\mathbb{C})$ takes non-negative real values, then $\int_G f(x)dx \geq 0$, and the integral is zero only if $f = 0$.*
(3) *It is right invariant in the sense that for all $f \in C^0(G,\mathbb{C})$ and $g \in G$,*

$$\int_G f(xg)dx = \int_G f(x)dx.$$

(4) *The constant function $1 \in C^0(G,\mathbb{C})$ with value $1 \in \mathbb{C}$ has integral*

$$\int_G 1\, dx = 1.$$

Remark 3.11. (1) If G is a compact group, then there is a unique measure μ on G called the Haar measure such that $\int_G f(x)dx = \int_G f d\mu$. If G is finite, then

$$\int_G f(x)dx = \frac{1}{|G|}\sum_{x\in G} f(x),$$

and in this case, the Haar measure is called the counting measure.

(2) In fact, parts (1)–(3) of Theorem 3.10 hold for every locally compact group such as $G = (\mathbb{R}, +)$. Moreover, for compact G (but not for locally compact G), part (3) can be replaced by the stronger statement that

$$\int_G f(xg)dx = \int_G f(x)dx = \int_G f(gx)dx$$

for all $f \in C^0(G, \mathbb{C})$ and $g \in G$.

Proof (of Theorem 3.9). We repeat the proof for G finite, replacing sum by integral. So given any choice $\langle -, -\rangle_0$ of inner product (resp. hermitian inner product) on V, we define $\langle -, -\rangle\colon V \times V \to \mathbb{C}$ by

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle = \int_G \langle \pi(x)(\boldsymbol{v}_1), \pi(x)(\boldsymbol{v}_2)\rangle_0 dx,$$

where we use the integral provided by Theorem 3.10. The linearity of the integral implies that $\langle -, -\rangle$ is an inner product (resp. a hermitian inner product), and we claim that it is π-invariant. Indeed, given $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$, we define $f \in C^0(G, \mathbb{C})$ by

$$f(x) = \langle \pi(x)(\boldsymbol{v}_1), \pi(x)(\boldsymbol{v}_2)\rangle_0$$

so that $\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle = \int_G f(x)dx$. The right-invariance of the integral shows that

$$\langle \pi(g)(\boldsymbol{v}_1), \pi(g)(\boldsymbol{v}_2)\rangle = \int_G f(xg)dx = \int_G f(x)dx = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle$$

for all $g \in G$, as claimed. We conclude that π is orthogonal (resp. unitary), so Proposition 3.4 shows that it is completely reducible. ☐

Remark 3.12. We explain the idea in the proof above, assuming that π is a real representation. The representation π induces a representation

$$G \xrightarrow{\rho} \mathrm{GL}(B^+(V))$$

on the space $B^+(V)$ of real symmetric bilinear forms on V defined by

$$\rho(g)(\langle -, -\rangle)(\boldsymbol{v}_1, \boldsymbol{v}_2) = \langle \pi(g^{-1})(\boldsymbol{v}_1), \pi(g^{-1})(\boldsymbol{v}_2)\rangle.$$

The subset $I(V) \subset B^+(V)$ consisting of the real inner products is an open cone, and it is preserved by ρ in the sense that $\rho(I(V)) \subset I(V)$ for all $g \in G$. Thus, given $\langle -, -\rangle_0 \in I(V)$, we have $\rho(G)(\langle -, -\rangle_0) \subset I(V)$, which expresses that the G-orbit through $\langle -, -\rangle_0$ is fully contained in $I(V)$. The π-invariant inner product

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle = \int_G \langle \pi(x)(\boldsymbol{v}_1), \pi(x)(\boldsymbol{v}_2)\rangle_0 dx$$

may thus be seen as an "average" over the G-orbit through $\langle -, -\rangle_0$.

The second proof is to construct the π-invariant inner product $\langle -, -\rangle$ as a center of mass. We recall the definition of the center of mass.

Definition 3.13. Let W be a finite dimensional real vector space, and let μ be a Lebesgue measure[12] on W. Suppose that $K \subset W$ is a Lebesgue measurable subset with positive volume $\mu(K) > 0$. The center of mass of $K \subset W$ is the vector

$$c(K) = \frac{1}{\mu(K)} \int_K \boldsymbol{x} d\boldsymbol{x} \in W.$$

We prove three lemmas in the situation of Definition 3.13.

Lemma 3.14. *If $K \subset W$ is a Lebesgue measurable subset with $\mu(K) > 0$, then*

$$c(f(K)) = f(c(K))$$

for all $f \in \mathrm{GL}(W)$.

Proof. Substituting $\boldsymbol{y} = f(\boldsymbol{x})$ and $d\boldsymbol{y} = \det(f) d\boldsymbol{x}$, we find that

$$c(f(K)) = \frac{1}{\mu(f(K))} \int_{f(K)} \boldsymbol{y} d\boldsymbol{y} = \frac{1}{\det(f)\mu(K)} \int_K \boldsymbol{x} \det(f) d\boldsymbol{x} = c(K)$$

as desired. Here we use that $\det(f)$ is a scalar, independent of $\boldsymbol{x} \in K$. □

We recall that if $K \subset W$ is any subset, then its convex hull is defined to be the subset $\mathrm{conv}(K) \subset W$ that consists of all linear combinations of the form

$$\boldsymbol{x}_0 a_0 + \cdots + \boldsymbol{x}_m a_m \in W$$

with $m \geq 0$, $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m \in K$, $a_0, \ldots, a_m \in [0,1]$, and $a_0 + \cdots + a_m = 1$. We say that a linear combination of this form is a convex combination.

Lemma 3.15. *If $K \subset W$ is compact, then so is $\mathrm{conv}(K) \subset W$.*

Proof. Let $n = \dim_{\mathbb{R}}(W)$. A classical theorem of Carathéodory states that every $\boldsymbol{w} \in \mathrm{conv}(K)$ is a convex combination of at most $n+1$ points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n \in K$. So in fact, the subset $\mathrm{conv}(K) \subset W$ consists of all convex combinations

$$\boldsymbol{x}_0 a_0 + \cdots + \boldsymbol{x}_n a_n \in W$$

with $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n \in K$, $a_0, \ldots, a_n \in [0,1]$, and $a_0 + \cdots + a_n = 1$. It follows that we have a continuous surjection

$$K^{n+1} \times \Delta^n \xrightarrow{p} \mathrm{conv}(K)$$

that to $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n, a_0, \ldots, a_n)$ assigns $\boldsymbol{x}_0 a_0 + \cdots + \boldsymbol{x}_n a_n$. Here

$$\Delta^n \subset [0,1]^{n+1}$$

is the subspace of tuples $(a_0, \ldots, a_n)$ with $a_0 + \cdots + a_n = 1$. So $\mathrm{conv}(K)$ is the image of a compact space by a continuous map, and therefore, it is compact. □

Lemma 3.16. *If $K \subset W$ is a compact*[13] *subset with $\mu(K) > 0$, then*

$$c(K) \in \mathrm{conv}(K).$$

[12] The normalization of μ is irrelevant for this definition.

[13] Every compact subset $K \subset W$ is Lebesgue measuable.

Proof. By the theory of (Lebesgue) integration,

$$c(K) = \lim_{r \to \infty} \frac{1}{\mu(K)} \sum_{i=1}^{r} \boldsymbol{x}_i \mu(K_i),$$

where $K = \coprod_{i=1}^{r} K_i$ is a decomposition of K into r disjoint Lebesgue measurable subsets, and where $\boldsymbol{x}_i \in K_i$ is any point. By definition, we have

$$\frac{1}{\mu(K)} \sum_{i=1}^{r} \boldsymbol{x}_i \mu(K_i) \in \operatorname{conv}(K)$$

for all $r \geq 1$. But by Lemma 3.15, $\operatorname{conv}(K) \subset W$ is a compact subset of a Hausdorff space, and hence, closed, so $c(K) \in \operatorname{conv}(K)$ as stated. $\square$

Proof (of Theorem 3.9). We first suppose that (V, π) is a finite dimensional real representation of the compact group G and show that π is orthogonal. Let $B^+(V)$ be the real vector space of symmetric bilinear forms on V, and let

$$G \xrightarrow{\rho} \operatorname{GL}(B^+(V))$$

be the group homomorphism defined by

$$\rho(g)(\langle -, - \rangle)(\boldsymbol{v}_1, \boldsymbol{v}_2) = \langle \pi(g^{-1})(\boldsymbol{v}_1), \pi(g^{-1})(\boldsymbol{v}_2) \rangle.$$

The pair $(B^+(V), \rho)$ is a representation of G. The subspace $I(V) \subset B^+(V)$ of inner products is an open cone, and it is ρ-invariant in the sense that for all $g \in G$,

$$\rho(g)(I(V)) \subset I(V).$$

We now choose $\langle -, - \rangle_0 \in I(V)$ and $\langle -, - \rangle_0 \in K_0 \subset I(V)$ with K_0 compact, and define $K \subset I(V)$ to be the image of the composite map

$$G \times K_0 \xrightarrow{G \times i} G \times B^+(V) \xrightarrow{\rho} B^+(V),$$

where $i \colon K_0 \to B^+(V)$ is the canonical inclusion. Since both maps are continuous, so is the composite map, and since $G \times K_0$ is compact, so is the image $K \subset B^+(V)$. Moreover, we have $K \subset I(V)$, because $K_0 \subset I(V)$ and because $I(V) \subset B^+(V)$ is ρ-invariant. We have $\mu(K) \geq \mu(K_0) > 0$, so the center of mass

$$\langle -, - \rangle = c(K) \in \operatorname{conv}(K) \subset B^+(V)$$

is defined. But $K \subset I(V)$ and $I(V) \subset B^+(V)$ is convex, being an open cone, so we have $\operatorname{conv}(K) \subset I(V)$, and hence,

$$\langle -, - \rangle = c(K) \in \operatorname{conv}(K) \subset I(V)$$

is an inner product. By Lemma 3.14, it is ρ-invariant, which is equivalent to the statement that $\langle -, - \rangle$ is a π-invariant inner product on V. In particular, (V, π) is orthogonal, and hence, completely reducible by Proposition 3.4.

Finally, if instead (V, π) is a finite dimensional complex representation of G, then we argue in the same way, but with $B^+(V)$ replaced by the *real* vector space $H^+(V)$ of hermitian forms on V, and with $I(V) \subset B^+(V)$ replaced by the open cone $J(V) \subset H^+(V)$ of hermitian inner products. $\square$

Appendix: Hermitian forms and hermitian inner products

Since we have already used the notions of a hermitian form and a hermitian inner product on a complex vector space, let us recall the definition. So let V be a (right) complex vector space. A hermitian form on V is a map

$$V \times V \xrightarrow{\langle -,- \rangle} \mathbb{C}$$

such that for $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in V$ and $a \in \mathbb{C}$, the following hold:[14]

(H1) $\langle \boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2 \rangle = \langle \boldsymbol{x}, \boldsymbol{y}_1 \rangle + \langle \boldsymbol{x}, \boldsymbol{y}_2 \rangle$
(H2) $\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y} \rangle = \langle \boldsymbol{x}_1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}_2, \boldsymbol{y} \rangle$
(H3) $\langle \boldsymbol{x}, \boldsymbol{y} \cdot a \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot a$
(H4) $\langle \boldsymbol{x} \cdot a, \boldsymbol{y} \rangle = \overline{a} \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle$
(H5) $\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}$

Here $\overline{a} \in \mathbb{C}$ is the complex conjugate of $a \in \mathbb{C}$. By (H5), we have in particular that $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \in \mathbb{R}$, and a hermitian form is defined to be a hermitian inner product if, in addition to (H1)–(H5), it has the following positivity property:

(P) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$ unless $\boldsymbol{x} = \boldsymbol{0}$.

As we have also used, the set $H^+(V)$ of hermitian forms on V form a *real* vector space with vector sum and scalar multiplication defined by

$$(\langle -,- \rangle_1 + \langle -,- \rangle_2)(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle_1 + \langle \boldsymbol{x}, \boldsymbol{y} \rangle_2$$
$$(\langle -,- \rangle \cdot a)(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot a.$$

with $\boldsymbol{x}, \boldsymbol{y} \in V$ and $a \in \mathbb{R} \subset \mathbb{C}$. The subset $J(V) \subset H^+(V)$ of hermitian inner products is an open cone. Indeed, while $J(V) \subset H^+(V)$ is closed under vector sum, it is only closed under scalar multiplication by *positive* real numbers a.

Appendix: Caratheodory's theorem

Let us prove Caratheodory's theorem that we used in the second proof above. So we let W be a finite dimensional real vector space, and recall that, by definition, the convex hull of a subset $K \subset W$ is the union

$$\operatorname{conv}(K) = \bigcup_{-1 \leq m < \infty} \operatorname{conv}_m(K) \subset W,$$

where $\operatorname{conv}_m(K) \subset W$ is the subset of all convex combinations of $m+1$ points in W, that is, the subset of all vectors of the form

$$\boldsymbol{x}_0 a_0 + \cdots + \boldsymbol{x}_m a_m$$

with $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m \in K$, $a_0, \ldots, a_m \in [0,1]$ and $a_0 + \cdots + a_m = 1$.

Theorem 3.17. *Let W be a real vector space of finite dimension n, let $K \subset W$ be any subset, and let $-1 \leq d \leq n$ be the dimension of the smallest affine subspace that contains K. In this situation,*

$$\operatorname{conv}(K) = \operatorname{conv}_d(K).$$

[14] We use the physics convention that $\langle -,- \rangle$ is linear in the second variable and conjugate linear in the first variable. Much of the mathematical literature, including the book, uses the opposite convention that $\langle -,- \rangle$ is linear in the first variable and conjugate linear in the second variable.

Proof. By the definition of the convex hull, we may assume that $K \subset W$ is a finite subset. We prove the statement by induction on the cardinality N of K. If $N = 0$, then $K = \emptyset$, so $\operatorname{conv}(K) = \emptyset = \operatorname{conv}_{-1}(K)$, and hence, the statement holds in this case. So we let $N = r > 0$ and assume that the statement has been proved for $N < r$. We let $K \subset W$ be a subset of cardinality r, and write $K = L \cup \{\boldsymbol{x}\}$ as the union of a subset $L \subset W$ of cardinality $r - 1$ and a singleton. Let d and e be the dimensions of the smallest affine subspaces that contain K and L, respectively. Clearly, either $d = e$ or $d = e + 1$. By the inductive hypothesis, we have

$$\operatorname{conv}(L) = \operatorname{conv}_e(L) \ = \bigcup_{1 \leq i \leq s} \Delta_i^e,$$

where each $\Delta_i^e \subset W$ is an e-simplex, whose $e+1$ vertices are elements of L, and this implies that

$$\operatorname{conv}(K) = \operatorname{conv}(L \cup \{\boldsymbol{x}\}) \ = \bigcup_{1 \leq i \leq s} \operatorname{conv}(\Delta_i^e \cup \{\boldsymbol{x}\}).$$

So it will suffice to prove that

$$\operatorname{conv}(\Delta_i^e \cup \{\boldsymbol{x}\}) \subset \operatorname{conv}_d(K)$$

for all $1 \leq i \leq s$. If $d = e+1$ or if $d = e$ and $\boldsymbol{x} \in \Delta_i^e$, then there is nothing to prove. So we assume that $d = e$ and that $\boldsymbol{x} \notin \Delta_i^e$. For every subset $M \subset W$, we have

$$\operatorname{conv}(M \cup \{\boldsymbol{x}\}) = \bigcup_{\boldsymbol{u} \in M} \operatorname{conv}(\{\boldsymbol{u}, \boldsymbol{x}\}),$$

so it suffices to show that for every $\boldsymbol{u} \in \Delta_i^e$, the line segment

$$\operatorname{conv}(\{\boldsymbol{u}, \boldsymbol{x}\}) = \{\boldsymbol{u}a + \boldsymbol{x}b \in W \mid a, b \in [0,1], a + b = 1\}$$

is contained in a d-simplex, whose vertices are elements of K. Since $\boldsymbol{u} \in \Delta_i^e$, we can write $\boldsymbol{u}$ as a convex combination

$$\boldsymbol{u} = \boldsymbol{x}_0 a_0 + \cdots + \boldsymbol{x}_e a_e$$

with $a_0, \dots, a_e \in [0,1]$ and $a_0 + \cdots + a_e = 1$. Thus, every $\boldsymbol{y} \in \operatorname{conv}(\{\boldsymbol{u}, \boldsymbol{x}\})$ can be written as a convex combination

$$\boldsymbol{y} = \boldsymbol{u}a + \boldsymbol{x}b = (\boldsymbol{x}_0 a_0 + \cdots + \boldsymbol{x}_e a_e)a + \boldsymbol{x}b$$

with $a, b \in [0,1]$ and $a + b = 1$. Since $d = e$ and $\boldsymbol{x} \notin \Delta_i^e$, we can arrange that at least one of the $a_0, \dots, a_e$ be equal to zero. By rearranging the $\boldsymbol{x}_i$, if necessary, we can assume that $a_0 = 0$. But then

$$\operatorname{conv}(\{\boldsymbol{u}, \boldsymbol{x}\}) \subset \operatorname{conv}(\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_e, \boldsymbol{x}\}),$$

which is an e-simplex with vertices in K as required. This proves the induction step and the theorem. $\square$

4. Dual representation, tensor product of representations

We first discuss the dual vector space. To do so (and not make mistakes), we will let k be any skew-field. So we do not assume $a \cdot b$ and $b \cdot a$ are equal for $a, b \in k$. A skew-field $k = (k, +, \cdot)$ has an opposite skew-field $k^{\mathrm{op}} = (k, +, \star)$ with the same underlying set and the same addition, but with the new multiplication

$$a \star b = b \cdot a.$$

By a k-vector space, we will always mean a *right* k-vector space. So we agree that scalars multiply from the right and not from the left. We have to do so, if we want matrices (that represent linear maps) to multiply from the left, and I think that we all agree that we want that. Let us recall how this works.

So let $\varphi \colon W \to V$ be a linear map between right k-vector spaces, which we will assume to be finite dimensional, and let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ and $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ be bases for V and W, respectively. Every $\boldsymbol{w} \in W$ and $\boldsymbol{v} \in V$ can be unique written as

$$\boldsymbol{v} = \boldsymbol{v}_1 x_1 + \boldsymbol{v}_2 x_2 + \cdots + \boldsymbol{v}_m x_m$$
$$\boldsymbol{w} = \boldsymbol{w}_1 y_1 + \boldsymbol{w}_2 y_2 + \cdots + \boldsymbol{w}_n y_n$$

with $\boldsymbol{x} = (x_j) \in M_{m,1}(k)$ and $\boldsymbol{y} = (y_i) \in M_{n,1}(k)$. Now, there is a unique matrix

$$A = (a_{ij}) \in M_{m,n}(k)$$

such that for all $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$, $\boldsymbol{v} = \varphi(\boldsymbol{w})$ if and only if $\boldsymbol{x} = A\boldsymbol{y}$, namely, the matrix whose entries a_{ij} are the unique solutions to the linear equations

$$\varphi(\boldsymbol{w}_j) = \boldsymbol{v}_1 a_{1j} + \boldsymbol{v}_2 a_{2j} + \cdots + \boldsymbol{v}_m a_{mj}$$

with $1 \leq j \leq n$. So the jth column in A is the coordinate vector of $\varphi(\boldsymbol{w}_j)$ with respect to the basis $(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$. We say that A is the matrix that represents $\varphi \colon W \to V$ with respect to the bases $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ and $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$.

If $U = (U, +, \cdot)$ is a *left* k-vector space, then we view it as a right k^{op}-vector space $U = (U, +, \star)$ with the same underlying set (of vectors) and the same vector sum, but with the new scalar multiplication $\star \colon U \times k^{\mathrm{op}} \to U$ given by

$$\boldsymbol{u} \star a = a \cdot \boldsymbol{u}.$$

We now discuss the dual vector space. So suppose that $V = (V, +, \cdot)$ is a *right* k-vector space. Its dual is the *left* k-vector space[15]

$$V^* = (\mathrm{Hom}_k(V, k), +, \cdot)$$

with vector sum and *left* scalar multiplication given by

$$(f + g)(\boldsymbol{v}) = f(\boldsymbol{v}) + g(\boldsymbol{v})$$
$$(a \cdot f)(\boldsymbol{v}) = a \cdot f(\boldsymbol{v}).$$

Let us check that $a \cdot f \in V^*$. It is clear that $a \cdot f$ is additive, and the calculation

$$(a \cdot f)(\boldsymbol{v} \cdot b) = a \cdot f(\boldsymbol{v} \cdot b) = a \cdot (f(\boldsymbol{v}) \cdot b) = (a \cdot f(\boldsymbol{v})) \cdot b = (a \cdot f)(\boldsymbol{v}) \cdot b.$$

shows that it is also preserves right multiplication by b, as required. Note that this would not be true, if we instead let a multiply from the right, unless $a \cdot b = b \cdot a$. We agreed to consider this left k-vector space as the right k^{op}-vector space

$$V^* = (\mathrm{Hom}_k(V, k), +, \star),$$

[15] The book writes V' instead of V^*.

with the right scalar multiplication by $a \in k^{\mathrm{op}}$ given by

$$(f \star a)(\boldsymbol{v}) = (a \cdot f)(\boldsymbol{v}) = a \cdot f(\boldsymbol{v}) = f(\boldsymbol{v}) \star a.$$

If $\dim_k(V) < \infty$, then a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ of the right k-vector space V gives rise to a basis $(\boldsymbol{v}_1^*, \dots, \boldsymbol{v}_m^*)$ of the dual right k^{op}-vector space V^* defined by

$$\boldsymbol{v}_i^*(\boldsymbol{v}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

which we call the dual basis.

A k-linear map $\varphi \colon W \to V$ between right k-vector spaces V and W determine a k^{op}-linear map $\varphi^* \colon V^* \to W^*$ between right k^{op}-vector spaces defined by

$$\varphi^*(f)(\boldsymbol{w}) = f(\varphi(\boldsymbol{w})).$$

Moreover, if V and W are finite dimensional, and if $A \in M_{m,n}(k)$ is the matrix that represents $\varphi \colon W \to V$ with respect to bases $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ and $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ of V and W, respectively, then the matrix that represents the map $\varphi^* \colon V^* \to W^*$ with respect to the dual bases $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m^*)$ and $(\boldsymbol{w}_1^*, \dots, \boldsymbol{w}_n^*)$ is the transpose matrix

$$A^t = (a_{ji}) \in M_{n,m}(k^{\mathrm{op}}).$$

If V is a right k-vector space, then its double dual $V^{**} = (V^*)^*$ is also a right k-vector space, so it is possible to compare them. There is a natural k-linear map

$$V \xrightarrow{\delta_V} V^{**}$$

defined by $\delta(\boldsymbol{v})(f) = f(\boldsymbol{v})$. That the map δ_V is natural[16] means that if $\varphi \colon W \to V$ is any k-linear map, then the diagram

$$\begin{array}{ccc} W & \xrightarrow{\delta_W} & W^{**} \\ \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \varphi^{**}} \\ V & \xrightarrow{\delta_V} & V^{**} \end{array}$$

commutes. If $\dim_k(V) < \infty$, then δ_V is an isomorphism. Indeed, if $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ is a basis of V, then $(\boldsymbol{v}_1^{**}, \dots, \boldsymbol{v}_m^{**})$ is a basis of V^{**}, and the calculation

$$\delta_V(\boldsymbol{v}_i)(\boldsymbol{v}_j^*) = \boldsymbol{v}_j^*(\boldsymbol{v}_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

shows that $\delta_V(\boldsymbol{v}_i) = \boldsymbol{v}_i^{**}$.

[16] We use the word "natural" to indicate natural transformations between functors, whereas we use the word "canonical" to indicate some particular or preferred choice. So "natural" has a precise mathematical meaning, whereas "canonical" does not.

Warning 4.1. By contrast, there is *no* preferred way to compare V and V^*. If V is a right k-vector space, then V^* is a right k^{op}-vector space. So to convert V^* into a right k-vector space $\sigma_*(V^*)$, we need a ring homomorphism

$$k \xrightarrow{\sigma} k^{\mathrm{op}}.$$

Such a ring homomorphism may not exist, and if it does, then it may not be unique. For instance, if $k = \mathbb{C}$, then we can choose σ to be the identity map, but we can also choose σ to be the map given by complex conjugation, which is different! Given $\sigma\colon k \to k^{\mathrm{op}}$, we must choose a map of right k-vector spaces

$$V \xrightarrow{b} \sigma_*(V^*).$$

The map b determines and is determined by the map

$$V \times V \xrightarrow{\langle -,- \rangle} k$$

defined by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = b(\boldsymbol{x})(\boldsymbol{y})$, and b is a well-defined and k-linear map if and only if the map $\langle -,- \rangle$ satisfies

(S1) For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$, $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$.
(S2) For all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $a \in k$, $\langle \boldsymbol{x}, \boldsymbol{y} \cdot a \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot a$.
(S3) For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$, $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
(S4) For all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $a \in k$, $\langle \boldsymbol{x} \cdot a, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \star \sigma(a) = \sigma(a) \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

We say that $\langle -,- \rangle$ is a σ-sesquilinear form, and we say that it is non-singular if the map b is an isomorphism. Therefore, in order to compare V and V^*, we must both choose a ring homomorphism $\sigma\colon k \to k^{\mathrm{op}}$ and a σ-sesquilinear form $\langle -,- \rangle\colon V \times V \to k$. Obviously, we should never do so, if we can avoid it! Let us also mention that if $\sigma \circ \sigma = \mathrm{id}_k$, then the σ-sesquilinear form $\langle -,- \rangle$ is said to be σ-hermitian if, in addition, it satisfies:

(H) For all $\boldsymbol{x}, \boldsymbol{y} \in V$, $\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \sigma(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)$.

The requirement (H) is equivalent to the statement that the diagram

$$\begin{array}{ccc} V & \xrightarrow{\quad b \quad} & \sigma_*(V^*) \\ \Big\downarrow{\scriptstyle \delta_V} & & \Big\uparrow{\scriptstyle \sigma_*(b^*)} \\ V^{**} & = & \sigma_*((\sigma_*(V^*))^*) \end{array}$$

commutes.

We now assume that k is a field. Since $a \cdot b = b \cdot a$ for all $a, b \in k$, the identity map is a ring homomorphism $\mathrm{id}_k\colon k \to k^{\mathrm{op}}$. If V is a right k-vector space, then we agree that we will *always* use the identity map $\sigma = \mathrm{id}_k\colon k \to k^{\mathrm{op}}$ to view the right k^{op}-vector space V^* as a right k-vector. In particular, if $k = \mathbb{C}$, then we will *not* use complex conjugation to view V^* as a right $\mathbb{C}$-vector space.

Definition 4.2. Let k be a field, and let (V, π) be a k-linear representation of a group G. The dual representation is the pair (V^*, π^*) of the dual k-vector space V^* and the group homomorphism

$$G \xrightarrow{\pi^*} \mathrm{GL}(V^*)$$

defined by $\pi^*(g) = \pi(g^{-1})^*$.

Let us check that π^* is indeed a group homomorphism. We have

$$\begin{aligned}\pi^*(g \cdot h) &= \pi((g \cdot h)^{-1})^* = \pi(h^{-1} \cdot g^{-1})^* = (\pi(h^{-1}) \circ \pi(g^{-1}))^* \\ &= \pi(g^{-1})^* \circ \pi(h^{-1})^* = \pi^*(g) \circ \pi^*(h)\end{aligned}$$

as required. Also, if V is finite dimensional, and if the matrix

$$A(g) \in M_n(k)$$

represents $\pi(g)\colon V \to V$ with respect to the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$, then the matrix

$$A(g^{-1})^t = (A(g)^{-1})^t \in M_n(k)$$

represents $\pi^*(g)\colon V^* \to V^*$ with respect to the dual basis $(\boldsymbol{v}_1^*, \dots, \boldsymbol{v}_n^*)$.

Example 4.3. Let (V, π) be a finite dimensional real representation of a group G. We claim that if π is orthogonal, then $\pi^* \simeq \pi$. To see this, recall that π is said to be orthogonal if there exists an inner product $\langle -, - \rangle\colon V \times V \to \mathbb{R}$ such that

$$\langle \pi(g)(\boldsymbol{x}), \pi(g)(\boldsymbol{y}) \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$$

for all $g \in G$ and $\boldsymbol{x}, \boldsymbol{y} \in V$. Therefore, the matrix $Q(g) \in M_n(\mathbb{R})$ that represents $\pi(g)\colon V \to V$ with respect to a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ that is orthonormal with respect to $\langle -, - \rangle$ is orthogonal, that is, it satisfies $Q(g) = (Q(g)^{-1})^t$. So the map

$$V \xrightarrow{\ b\ } V^*$$

defined by $b(\boldsymbol{x})(\boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is intertwining between π and π^*. Since $\langle -, - \rangle$ is an inner product, the map b is also an isomorphism of vector spaces, so the claim follows. We note, as in Warning 4.1, that the isomorphism $\pi \simeq \pi^*$ is not canonical, let alone natural, but depends on the choice of inner product.

Example 4.4. If k is a field and if (V, π) is a k-linear representation of a group G, then the map $\delta_V\colon V \to V^{**}$ is intertwining between π and $\pi^{**} = (\pi^*)^*$. Indeed,

$$\pi^{**}(g) = \pi^*(g^{-1})^* = \pi(g)^{**},$$

and the diagram

$$\begin{array}{ccc} V & \xrightarrow{\ \delta_V\ } & V^{**} \\ \downarrow{\scriptstyle \pi(g)} & & \downarrow{\scriptstyle \pi(g)^{**}} \\ V & \xrightarrow{\ \delta_V\ } & V^{**} \end{array}$$

commutes by the naturality of δ.

Theorem 4.5. *Let k be a field, and let (V, π) be a finite dimensional k-linear representation of a group G.*

(1) *π is irreducible if and only π^* is so.*
(2) *π is completely reducible if and only if π^* is so.*

Proof. Indeed, the sequence

$$0 \longrightarrow U \xrightarrow{\ i\ } V \xrightarrow{\ p\ } W \longrightarrow 0$$

is exact (resp./ split exact) if and only if the sequence

$$0 \longrightarrow W^* \xrightarrow{p^*} V^* \xrightarrow{i^*} U^* \longrightarrow 0$$

is exact (resp. split exact). □

Remark 4.6. In elementary particle physics, an elementary particle is an irreducible representation π of the gauge group $\mathcal{G}$. The corresponding antiparticle is the dual (irreducible) representation π^*.

Sums of representations

We will next define (direct) sums of representations. There are two version, the exterior sum denoted "⊞" and the (interior) sum denoted "⊕." First, let

$$V_1 \xrightarrow{i_1} V_1 \oplus V_2 \xleftarrow{i_2} V_2$$

be a direct sum of k-vector spaces V_1 and V_2. Given $\boldsymbol{v}_1 \in V_1$ and $\boldsymbol{v}_2 \in V_2$, we write

$$\boldsymbol{v}_1 \oplus \boldsymbol{v}_2 = i_1(\boldsymbol{v}_1) + i_2(\boldsymbol{v}_2) \in V_1 \oplus V_2.$$

If $f_1 \colon W_1 \to V_1$ and $f_2 \colon W_2 \to V_2$ are k-linear maps, then there is a unique k-linear map $f_1 \oplus f_2 \colon W_1 \oplus W_2 \to V_1 \oplus V_2$ that makes the diagram

$$\begin{array}{ccccc} W_1 & \xrightarrow{j_1} & W_1 \oplus W_2 & \xleftarrow{j_2} & W_2 \\ \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_1 \oplus f_2} & & \downarrow{\scriptstyle f_2} \\ V_1 & \xrightarrow{i_1} & V_1 \oplus V_2 & \xleftarrow{i_2} & V_2 \end{array}$$

commute. In terms of elements, we have

$$(f_1 \oplus f_2)(\boldsymbol{w}_1 \oplus \boldsymbol{w}_2) = f_1(\boldsymbol{w}_1) \oplus f_2(\boldsymbol{w}_2).$$

Moreover, if also $g_1 \colon V_1 \to U_1$ and $g_2 \colon V_2 \to U_2$ are k-linear maps, then

$$(g_1 \oplus g_2) \circ (f_1 \oplus f_2) = (g_1 \circ f_1) \oplus (g_2 \circ f_2).$$

In particular, we have a well-defined group homomorphism

$$\mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \xrightarrow{\oplus} \mathrm{GL}(V_1 \oplus V_2)$$

that to (f_1, f_2) assigns $f_1 \oplus f_2$.

Definition 4.7. Let (V_1, π_1) and (V_2, π_2) be k-linear representations of two groups G_1 and G_2, respectively. The k-linear representation $(V_1 \oplus V_2, \pi_1 \boxplus \pi_2)$ of the product group $G_1 \times G_2$, where $\pi_1 \boxplus \pi_2$ is the composite group homomorphism

$$G_1 \times G_2 \xrightarrow{\pi_1 \times \pi_2} \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \xrightarrow{\oplus} \mathrm{GL}(V_1 \oplus V_2),$$

is called the exterior sum of (V_1, π_1) and (V_2, π_2).

Spelling out the definition in terms of elements, we have

$$(\pi_1 \boxplus \pi_2)(g_1, g_2)(\boldsymbol{v}_1 \oplus \boldsymbol{v}_2) = \pi_1(g_1)(\boldsymbol{v}_1) \oplus \pi_2(g_2)(\boldsymbol{v}_2)$$

for $g_1 \in G_1$, $g_2 \in G_2$, $\boldsymbol{v}_1 \in V_1$ and $\boldsymbol{v}_2 \in V_2$.

For every group G, the diagonal map

$$G \xrightarrow{\Delta_G} G \times G$$

defined by $\Delta_G(g) = (g, g)$ is also a group homomorphism.

Definition 4.8. Let (V_1, π_1) and (V_2, π_2) be k-linear representations of the *same* group G. The k-linear representation $(V_1 \oplus V_2, \pi_1 \oplus \pi_2)$ of G, where $\pi_1 \oplus \pi_2$ is defined to be the composite group homomorphism

$$G \xrightarrow{\Delta_G} G \times G \xrightarrow{\pi_1 \boxplus \pi_2} \mathrm{GL}(V_1 \oplus V_2),$$

is called the sum of π_1 and π_2.[17]

Again, spelling out the definition in terms of elements, we have

$$(\pi_1 \oplus \pi_2)(g)(\boldsymbol{v}_1 \oplus \boldsymbol{v}_2) = \pi_1(g)(\boldsymbol{v}_1) \oplus \pi_2(g)(\boldsymbol{v}_2)$$

for $g \in G$, $\boldsymbol{v}_1 \in V_2$, and $\boldsymbol{v}_2 \in V_2$.

Remark 4.9. If (V_1, π_1) and (V_2, π_2) both are k-linear representations of the same group G, then it may seem as if there is not much difference between the representations $\pi_1 \boxplus \pi_2$ and $\pi_1 \oplus \pi_2$. However, there is a big difference, which is that the former is a representation of the group $G \times G$, while the latter is a representation of the much smaller group G.

Example 4.10. Let (V, π) be a k-linear representation of a group G. If $U_1, U_2 \subset V$ are π-invariant subspaces, then the canonical inclusion maps

$$U_1 \xrightarrow{j_1} V \xleftarrow{j_2} U_2$$

are intertwining between π_{U_i} and π, and hence, the induced map

$$U_1 \oplus U_2 \xrightarrow{j_1 + j_2} V$$

is intertwining between $\pi_{U_1} \oplus \pi_{U_2}$ and π. We recall that $j_1 + j_2$ is surjective if and only if $U_1 + U_2 = V$ and that $j_1 + j_2$ is injective if and only if $U_1 \cap U_2 = \{\mathbf{0}\}$. In particular, if $j_1 + j_2$ is bijective, then $\pi \simeq \pi_{U_1} \oplus \pi_{U_2}$.

We can now restate Theorems 2.12 and 2.13 as follows:

Theorem 4.11. *A finite dimensional k-linear representation (V, π) of a group G is completely reducible if and only if $\pi \simeq \pi_1 \oplus \cdots \oplus \pi_m$ with $\pi_1, \ldots, \pi_m$ irreducible.*

Theorem 4.12. *Let (V, π) be a k-linear representation of a group G, and suppose that $\pi \simeq \pi_1 \oplus \cdots \oplus \pi_m$ with $\pi_1, \ldots, \pi_m$ irreducible. If $U \subset V$ is π-invariant, then π_U is isomorphic to the sum of some of the π_i, and $\pi_{V/U}$ is isomorphic to the sum of the remaining π_i.*

Lemma 4.13. *Let (V, π) be a k-linear representation of a group G, and suppose that $V_1, \ldots, V_m \subset V$ are π-invariant subspaces such that the representations $\pi_{V_1}, \ldots, \pi_{V_m}$ are irreducible and pairwise non-isomorphic. In this case, the canonical map*

$$V_1 \oplus \cdots \oplus V_m \longrightarrow V$$

is injective, so $\pi_{V_1} \oplus \cdots \oplus \pi_{V_m}$ is a subrepresentation of π.

[17] The book writes $\pi_1 + \pi_2$ instead of $\pi_1 \oplus \pi_2$.

Proof. We argue by induction on $m \geq 0$, the case $m = 0$ being trivial. So we assume that the statement has been proved for $m < r$ and prove it for $m = r$. By the inductive hypothesis, the canonical map

$$V_1 \oplus \cdots \oplus V_{r-1} \longrightarrow V$$

is injective with image $V_1 + \cdots + V_{r-1}$, so the kernel of the canonical map

$$V_1 \oplus \cdots \oplus V_{r-1} \oplus V_r \longrightarrow V$$

is equal to $(V_1 + \cdots + V_{r-1}) \cap V_r$. Since π_{V_r} is irreducible, the kernel in question is nonzero if and only if $V_r \subset V_1 + \cdots + V_{r-1}$. However, by Theorem 4.12, this is not possible, because $\pi_{V_r} \not\simeq \pi_{V_i}$ for all $1 \leq i < r$. □

We can now prove the following analogue of unique prime factorization for semisimple representations.

Theorem 4.14. *Let $\pi_1, \ldots, \pi_m$ and $\rho_1, \ldots, \rho_n$ be irreducible k-linear representations of a group G, and suppose that $\pi_1 \oplus \cdots \oplus \pi_m \simeq \rho_1 \oplus \cdots \oplus \rho_n$. In this case, $m = n$ and, up to a reordering, $\pi_i \simeq \rho_i$ for all $1 \leq i \leq m$.*

Proof. The proof is by induction on $m \geq 0$, the case $m = 0$ being trivial. So we assume that the statement has been proved for $m < r$ and prove it for $m = r$. We choose any $0 < s < r$ and consider the two subrepresentations

$$\pi_1 \oplus \cdots \oplus \pi_s, \pi_{s+1} \oplus \cdots \oplus \pi_r \subset \pi_1 \oplus \cdots \oplus \pi_r \simeq \rho_1 \oplus \cdots \oplus \rho_n.$$

Theorem 4.12 shows that $\pi_1 \oplus \cdots \oplus \pi_s$ is a sum of some of the ρ_i, and that $\pi_{s+1} \oplus \cdots \oplus \pi_r$ is the sum of the remaining ρ_i. Since we $s < r$ and $r - s < r$, it follows from the inductive hypothesis that, up to a reordering, $\pi_i \simeq \rho_i$ for $1 \leq i \leq s$ and for $s + 1 \leq i \leq r$. This proves the induction step, and hence, the theorem. □

Tensor products of representations

We finally define tensor products of representations, and again there is both an exterior tensor product "⊠" and an interior tensor product "⊗." First, we recall that a tensor product of two k-vector spaces is a k-bilinear map

$$V_1 \times V_2 \xrightarrow{p_{V_1,V_2}} V_1 \otimes V_2$$

with the property that for every k-bilinear map

$$V_1 \times V_2 \xrightarrow{b} U$$

there is exists a unique k-linear map $\widehat{b}\colon V_1 \otimes V_2 \to U$ that makes the diagram

$$\begin{array}{ccc} V_1 \times V_2 & \xrightarrow{p_{V_1,V_2}} & V_1 \otimes V_2 \\ & {\scriptstyle b}\searrow \quad \swarrow{\scriptstyle \widehat{b}} & \\ & U & \end{array}$$

commute. We say that this property of the tensor product is its defining universal property. In particular, we conclude that if also $q_{V_1,V_2}\colon V_1 \times V_2 \to V_1 \mathbin{\widetilde{\otimes}} V_2$ is a tensor product of V_1 and V_2, then the unique k-linear maps

$$V_1 \otimes V_2 \underset{\widehat{p}}{\overset{\widehat{q}}{\rightleftarrows}} V_1 \mathbin{\widetilde{\otimes}} V_2$$

are each other's inverses. In this way, a tensor product of V_1 and V_2 is unique, up to unique isomorphism, so we often abuse language and call it *the* tensor product of V_1 and V_2. Given $\boldsymbol{v}_1 \in V_1$ and $\boldsymbol{v}_2 \in V_2$, we write[18]

$$\boldsymbol{v}_1 \otimes \boldsymbol{v}_2 = p_{V_1,V_2}(\boldsymbol{v}_1, \boldsymbol{v}_2) \in V_1 \otimes V_2.$$

We recall that if the families $(\boldsymbol{e}_i)_{i\in I}$ and $(\boldsymbol{f}_j)_{j\in J}$ are bases of V_1 and V_2, respectively, then the family $(\boldsymbol{e}_i \otimes \boldsymbol{f}_j)_{(i,j)\in I\times J}$ is a basis of $V_1 \otimes V_2$. In particular, we have

$$\dim_k(V_1 \otimes V_2) = \dim_k(V_1) \cdot \dim_k(V_2).$$

Suppose that $f_1\colon W_1 \to V_1$ and $f_2\colon W_2 \to V_2$ are two k-linear maps. It follows from the defining universal property of the tensor product, there is a unique k-linear map $f_1 \otimes f_2\colon W_1 \otimes W_2 \to V_1 \otimes V_2$ that makes the diagram

$$\begin{array}{ccc} W_1 \times W_2 & \xrightarrow{p_{W_1,W_2}} & W_1 \otimes W_2 \\ \downarrow{\scriptstyle f_1\times f_2} & & \downarrow{\scriptstyle f_1\otimes f_2} \\ V_1 \times V_2 & \xrightarrow{p_{V_1,V_2}} & V_1 \otimes V_2 \end{array}$$

commute. Indeed, the map $p_{V_1,V_2} \circ (f_1 \times f_2)$ is k-bilinear. By the uniqueness of this assignment, we conclude that there is a well-defined map

$$\mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \xrightarrow{\otimes} \mathrm{GL}(V_1 \otimes V_2)$$

that to (f_1, f_2) assigns $f_1 \otimes f_2$ and that this map is a group homomorphism.

Definition 4.15. Let (V_1, π_1) and (V_2, π_2) be k-linear representations of two groups G_1 and G_2, respectively. The k-linear representation $(V_1 \otimes V_2, \pi_1 \boxtimes \pi_2)$ of the product group $G_1 \times G_2$, where $\pi_1 \boxtimes \pi_2$ is the composite group homomorphism

$$G_1 \times G_2 \xrightarrow{\pi_1\times\pi_2} \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \xrightarrow{\otimes} \mathrm{GL}(V_1 \otimes V_2),$$

is called the exterior tensor product of (V_1, π_1) and (V_2, π_2).[19]

Spelling out the definition in terms of pure tensors, we have

$$(\pi_1 \boxtimes \pi_2)(g_1, g_2)(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) = \pi_1(g_1)(\boldsymbol{v}_1) \otimes \pi_2(g_2)(\boldsymbol{v}_2),$$

where $g_1 \in G_1$, $g_2 \in G_2$, $\boldsymbol{v}_1 \in V_1$, and $\boldsymbol{v}_2 \in V_2$.

---

[18] The tensors of the form $\boldsymbol{v}_1 \otimes \boldsymbol{v}_2$ are called pure tensors. They are the tensors that belong to the image of the map $p_{V_1,V_2}\colon V_1 \times V_2 \to V_1 \otimes V_2$. Every tensor can be written as a sum of pure tensors, but it is almost never useful to do so, since the sum is not unique. A tensor that is not a pure tensor is said to be entangled. This is the source of entanglement in quantum mechanics.

[19] Confusingly, the book writes $\pi_1 \otimes \pi_2$ instead of $\pi_1 \boxtimes \pi_2$.

Definition 4.16. Let (V_1, π_1) and (V_2, π_2) be k-linear representations of the *same* group G. The k-linear representation $(V_1 \otimes V_2, \pi_1 \otimes \pi_2)$, where $\pi_1 \otimes \pi_2$ is defined to be the composite group homomorphism

$$G \xrightarrow{\Delta_G} G \times G \xrightarrow{\pi_1 \boxtimes \pi_2} \mathrm{GL}(V_1 \otimes V_2).$$

is called the tensor product of π_1 and π_2.[20]

Again, spelling out the definition in terms of pure tensors, we have

$$(\pi_1 \otimes \pi_2)(g)(\boldsymbol{v}_1 \otimes \boldsymbol{v}_2) = \pi_1(g)(\boldsymbol{v}_1) \otimes \pi_2(g)(\boldsymbol{v}_2).$$

Example 4.17. 1) We recall that sum and tensor product satisfy a distributive law in the sense that the canonical map

$$(U \otimes V_1) \oplus (U \otimes V_2) \longrightarrow U \otimes (V_1 \oplus V_2)$$

is an isomorphism. So if $\tau \colon G \to \mathrm{GL}(k^n)$ is the trivial k-linear representation of G on $V = k^n$, and if $\pi \colon G \to \mathrm{GL}(U)$ is any k-linear representation, then

$$\pi \otimes \tau \simeq \pi \otimes (k \oplus \cdots \oplus k) \simeq (\pi \otimes k) \oplus \cdots \oplus (\pi \otimes k) \simeq \pi \oplus \cdots \oplus \pi,$$

where there are n summands.

2) If U and V are right k-vector spaces, then there is a natural k-linear map

$$V \otimes U^* \xrightarrow{\alpha_{U,V}} \mathrm{Hom}_k(U, V)$$

defined by $\alpha(\boldsymbol{v} \otimes f)(\boldsymbol{u}) = \boldsymbol{v} \cdot f(\boldsymbol{u})$. It is an isomorphism if at least one of U and V is finite dimensional. That the map $\alpha_{U,V}$ is natural means that if $\varphi \colon U_2 \to U_1$ and $\psi \colon V_1 \to V_2$ are k-linear maps, then the diagram

$$\begin{array}{ccc}
V_1 \otimes U_1^* & \xrightarrow{\alpha_{U_1,V_1}} & \mathrm{Hom}_k(U_1, V_1) \\
\downarrow{\scriptstyle \psi \otimes U_1^*} & & \downarrow{\scriptstyle \mathrm{Hom}(U_1,\psi)} \\
V_2 \otimes U_1^* & \xrightarrow{\alpha_{U_1,V_2}} & \mathrm{Hom}_k(U_1, V_2) \\
\downarrow{\scriptstyle V_2 \otimes \varphi^*} & & \downarrow{\scriptstyle \mathrm{Hom}(\varphi,V_2)} \\
V_2 \otimes U_2^* & \xrightarrow{\alpha_{U_2,V_2}} & \mathrm{Hom}_k(U_2, V_2)
\end{array}$$

commutes. In particular, if (V, π) is a k-linear represenstation of G, then

$$\begin{array}{ccc}
V \otimes V^* & \xrightarrow{\alpha_{V,V}} & \mathrm{Hom}_k(V, V) \\
\downarrow{\scriptstyle \pi(g) \otimes V^*} & & \downarrow{\scriptstyle \mathrm{Hom}(V,\pi(g))} \\
V \otimes V^* & \xrightarrow{\alpha_{V,V}} & \mathrm{Hom}_k(V, V) \\
\downarrow{\scriptstyle V \otimes \pi(g^{-1})^*} & & \downarrow{\scriptstyle \mathrm{Hom}(\pi(g^{-1}),V)} \\
V \otimes V^* & \xrightarrow{\alpha_{V,V}} & \mathrm{Hom}_k(V, V)
\end{array}$$

[20] The book writes $\pi_1\pi_2$ instead of $\pi_1 \otimes \pi_2$.

commutes for all $g \in G$. The outer diagram is

$$\begin{array}{ccc} V \otimes V^* & \xrightarrow{\alpha_{V,V}} & \mathrm{End}_k(V) \\ \big\downarrow{\scriptstyle (\pi\otimes\pi^*)(g)} & & \big\downarrow{\scriptstyle \mathrm{Ad}(\pi(g))} \\ V \otimes V^* & \xrightarrow{\alpha_{V,V}} & \mathrm{End}_k(V). \end{array}$$

So if V is finite dimensional, then $\alpha_{V,V}$ is isomorphism of k-linear representations

$$\pi \otimes \pi^* \simeq \mathrm{Ad} \circ \pi,$$

where the right-hand side is the k-linear representation of G on V given by the composite group homomorphism

$$G \xrightarrow{\pi} \mathrm{GL}(V) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(V).$$

In particular, if we take $G = \mathrm{GL}(V)$ and $\pi = \mathrm{id}$, then π and π^* are irreducible, but their tensor product $\pi \otimes \pi^*$ is not!

Remark 4.18. If π_1 and π_2 are irreducible representations, then it is an important problem called "scattering" to determine how $\pi_1 \otimes \pi_2$ decomposes as a sum

$$\pi_1 \otimes \pi_2 \simeq \rho_1 \oplus \cdots \oplus \rho_m$$

of irreducible representations. The name "scattering" comes from physics. Indeed, by colliding the elementary particles π_1 and π_2, one obtains the state $\pi_1 \otimes \pi_2$, which, in turn, decays to the collection of elementary particles $\rho_1, \ldots, \rho_m$.

Let $f\colon A \to B$ be a ring homomorphism. If $N = (N, +, \cdot)$ is a right B-module, then we define a right A-module

$$f_*(N) = (N, +, \star)$$

with the same underlying set and addition, but with right scalar multiplication by elements $a \in A$ on elements $\boldsymbol{y} \in N$ given by

$$\boldsymbol{y} \star a = \boldsymbol{y} \cdot f(a).$$

Moreover, if $h\colon N_1 \to N_2$ is a B-linear map between right B-modules N_1 and N_2, then the same map is an A-linear map

$$f_*(N_1) \xrightarrow{f_*(h)=h} f_*(N_2)$$

between the right A-modules $f_*(N_1)$ and $f_*(N_2)$.

Conversely, if $M = (M, +, \cdot)$ is a right A-module, then we define

$$f^*(M) = (M \otimes_A B, +, \cdot)$$

to be the right B-module, where for $\boldsymbol{x} \in M$ and $b_1, b_2 \in B$,

$$(\boldsymbol{x} \otimes b_1) \cdot b_2 = \boldsymbol{x} \otimes (b_1 b_2).$$

If $g\colon M_1 \to M_2$ is an A-linear map, then we define

$$f^*(M_1) \xrightarrow{f^*(g)} f^*(M_2)$$

to be the unique B-linear map such that for $\boldsymbol{x} \in M$ and $b \in B$,

$$f^*(g)(\boldsymbol{x} \otimes b) = g(\boldsymbol{x}) \otimes b.$$

It is well-defined, because g is A-linear. Indeed, if $\boldsymbol{x} \in M$, $a \in A$, and $b \in B$, then

$$f^*(g)(\boldsymbol{x}a \otimes b) = g(\boldsymbol{x}a) \otimes b = g(\boldsymbol{x})a \otimes b = g(\boldsymbol{x}) \otimes f(a)b = f^*(g)(\boldsymbol{x} \otimes f(a)b).$$

We say that f^* is the extension of scalars along f, and we say that f_* is the restriction of scalars along f. They are functors

$$\mathrm{Mod}_A \underset{f_*}{\overset{f^*}{\rightleftarrows}} \mathrm{Mod}_B$$

between the respective categories of right modules and linear maps. Indeed, it follows immediately from the definitions that, as required,

$$\begin{aligned} f^*(\mathrm{id}_M) &= \mathrm{id}_{f^*(M)} \\ f^*(g_1 \circ g_2) &= f^*(g_1) \circ f^*(g_2) \end{aligned}$$

and that

$$\begin{aligned} f_*(\mathrm{id}_N) &= \mathrm{id}_{f_*(N)} \\ f_*(h_1 \circ h_2) &= f_*(h_1) \circ f_*(h_2). \end{aligned}$$

Example 5.1. If $\sigma\colon \mathbb{C} \to \mathbb{C}$ is the ring homomorphism given by complex conjugation,

$$\sigma(a + ib) = a - ib,$$

and if $V = (V, +, \cdot)$ is a right $\mathbb{C}$-vector space, then

$$\sigma_*(V) = (V, +, \star)$$

is the complex conjugate right $\mathbb{C}$-vector space $\bar{V}$.

Returning to the general situation above, we define the unit map

$$M \xrightarrow{\eta_M} f_*f^*(M)$$

by $\eta_M(\boldsymbol{x}) = \boldsymbol{x} \otimes 1$ and the counit map

$$f^*f_*(N) \xrightarrow{\epsilon_N} N$$

by $\epsilon_N(\boldsymbol{y} \otimes b) = \boldsymbol{y}b$. They are both natural transformations of functors, which means that if $g\colon M_1 \to M_2$ and $h\colon N_1 \to N_2$ are an A-linear map and a B-linear map, respectively, then the following diagrams commute:

$$\begin{array}{ccc} M_1 & \xrightarrow{\eta_{M_1}} & f_*f^*(M_1) \\ \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f_*f^*(g)} \\ M_2 & \xrightarrow{\eta_{M_2}} & f_*f^*(M_2) \end{array} \qquad \begin{array}{ccc} f^*f_*(N_1) & \xrightarrow{\epsilon_{N_1}} & N_1 \\ \downarrow{\scriptstyle f^*f_*(h)} & & \downarrow{\scriptstyle h} \\ f^*f_*(N_2) & \xrightarrow{\epsilon_{N_2}} & N_2 \end{array}$$

Moreover, for every right A-module M and every right B-module N, the diagrams

$$\begin{array}{ccc} f^*(M) & \xrightarrow{f^*(\eta_M)} & f^*f_*f^*(M) \\ & \searrow{=} & \downarrow{\scriptstyle \epsilon_{f^*(M)}} \\ & & f^*(M) \end{array} \qquad \begin{array}{ccc} f_*(N) & \xrightarrow{\eta_{f_*(N)}} & f_*f^*f_*(N) \\ & \searrow{=} & \downarrow{\scriptstyle f_*(\epsilon_N)} \\ & & f_*(N) \end{array}$$

commute. We refer to this by saying that η and ϵ satisfy the triangle identities and that the quadruple $(f^*, f_*, \epsilon, \eta)$ is an adjunction from Mod_B to Mod_A.

Proposition 5.2. *In the above situation, the maps*

$$\mathrm{Hom}_B(f^*(M), N) \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \mathrm{Hom}_A(M, f_*(N))$$

defined by $\alpha(g) = f_*(g) \circ \eta_M$ *and* $\beta(h) = \epsilon_N \circ f^*(h)$ *are each other's inverses.*

Proof. By definition, the map $\alpha(g)$ is the composite map

$$M \xrightarrow{\eta_M} f_*f^*(M) \xrightarrow{f_*(g)} f_*(N)$$

so the map $(\beta \circ \alpha)(g) = \beta(\alpha(g))$ is the composition of the upper horizontal maps and right-hand vertical map in the following diagram:

$$\begin{array}{ccccc} f^*(M) & \xrightarrow{f^*(\eta_M)} & f^*f_*f^*(M) & \xrightarrow{f^*f_*(g)} & f^*f_*(N) \\ & \searrow{=} & \downarrow{\scriptstyle \epsilon_{f^*(M)}} & & \downarrow{\scriptstyle \epsilon_N} \\ & & f^*(M) & \xrightarrow{\quad g \quad} & N \end{array}$$

But the left-hand triangle commutes by the triangle identities, and the right-hand square commutes by the naturality of ϵ. So we conclude that $(\beta \circ \alpha)(g) = g$, as desired. Similarly, the map $\beta(h)$ is defined to be the composite map

$$f^*(M) \xrightarrow{f^*(h)} f^*f_*(N) \xrightarrow{\epsilon_N} N$$

so the map $(\alpha \circ \beta)(h) = \alpha(\beta(h))$ is the composition of the left-hand vertical map and lower horizontal maps in the following diagram:

$$\begin{array}{ccccc} M & \xrightarrow{\ h\ } & f_*(N) & = & \\ \downarrow{\scriptstyle \eta_M} & & \downarrow{\scriptstyle \eta_{f_*(N)}} & & \\ f_*f^*(M) & \xrightarrow{f_*f^*(h)} & f_*f^*f_*(N) & \xrightarrow{f_*(\epsilon_N)} & f_*(N), \end{array}$$

But the left-hand square commutes by the naturality of η, and the right-hand triangle commutes by the triangle identities, so we conclude that $(\alpha \circ \beta)(h) = h$. □

Let $f\colon k \to k'$ be an extension of fields. If V is a k-vector space, then the map

$$\mathrm{GL}(V) \xrightarrow{f^*} \mathrm{GL}(f^*(V))$$

is a group homomorphism, because f^* is a functor. Hence, if (V, π) is a k-linear representation of a group G, then we obtain a k'-linear representation of G given by the pair $(f^*(V), f^*\pi)$, where $f^*\pi$ is the composite group homomorphism

$$G \xrightarrow{\pi} \mathrm{GL}(V) \xrightarrow{f^*} \mathrm{GL}(f^*(V))$$

We call $f^*\pi$ the k'-linear representation obtained from the k-linear representation π by extension of scalars along f.

Similarly, if V' is a k'-vector space, then the map

$$\mathrm{GL}(V') \xrightarrow{f_*} \mathrm{GL}(f_*(V'))$$

is a group homomorphism, because f_* is a functor. Hence, if (V', π') is a k'-linear representation of G, then we obtain a k-linear representation of G given by the pair $(f_*(V'), f_*\pi')$, where $f_*\pi'$ is the composite group homomorphism

$$G \xrightarrow{\pi'} \mathrm{GL}(V') \xrightarrow{f_*} \mathrm{GL}(f_*(V')).$$

We call $f_*\pi'$ the k-linear representation obtained from the k'-linear representation π' by restriction of scalars along f.

Remark 5.3. If $f\colon k \to k'$ is a field extension, then

$$\begin{aligned} \dim_{k'}(f^*(V)) &= \dim_k(V) \\ \dim_k(f_*(V')) &= d \cdot \dim_{k'}(V'), \end{aligned}$$

where $d = [k' : k]$ is the degree of the extension.

Theorem 5.4. *Let $f\colon k \to k'$ be a field extension. Two finite-dimensional k-linear representations (V_1, π_1) and (V_2, π_2) of a group G are isomorphic if and only if the k'-linear representations $(f^*(V_1), f^*\pi_1)$ and $(f^*(V_2), f^*\pi_2)$ are so.*

Proof. If $h\colon V_1 \to V_2$ is a k-linear isomorphism that is intertwining between π_1 and π_2, then $f^*(h)\colon f^*(V_1) \to f^*(V_2)$ is a k'-linear isomorphism that is intertwining between $f^*\pi_1$ and $f^*\pi_2$. This proves the “only if” part of the statement.

To prove the "if" part of the statement, we assume that $f^*\pi_1 \simeq f^*\pi_2$ and prove that $\pi_1 \simeq \pi_2$. The proof that we give here uses that the field k is infinite. A different proof based on the Krull–Schmidt theorem works for all k. We first note that

$$\dim_k(V_1) = \dim_{k'}(f^*(V_1)) = \dim_{k'}(f^*(V_2)) = \dim_k(V_2),$$

where the middle equality holds by the assumption that $f^*\pi_1 \simeq f^*\pi_2$. So we may consider π_1 and π_2 to be matrix representations

$$G \xrightarrow{\pi_1,\pi_2} \mathrm{GL}_n(k).$$

Moreover, by viewing k as a subfield $k \subset k'$, namely, as the image of the extension $f\colon k \to k'$, we may consider $f^*\pi_1$ and $f^*\pi_2$ as the matrix representations

$$G \xrightarrow{\pi_1,\pi_2} \mathrm{GL}_n(k) \subset \mathrm{GL}_n(k').$$

Now, that $\pi_1 \simeq \pi_2$ means that there exists $C \in M_n(k)$ such that

(a) For all $g \in G$, $C \cdot \pi_1(g) = \pi_2(g) \cdot C$.
(b) The determinant $\det(C)$ is nonzero.

The requirement (a) is a system of linear equations of k in n^2 variables. By Gauss elimination, we know that a general solution has can be written uniquely as

$$C = t_1C_1 + \cdots + t_mC_m$$

with $(C_1, \dots, C_m)$ a linearly independent family of vectors in the k-vector space $M_n(k)$ and with $(t_1, \dots, t_m)$ a family of scalars in the field k. The requirement (b) is the statement that there exists a family $(t_1, \dots, t_m)$ of scalars in k such that the value of the polynomial

$$p(x_1, \dots, x_m) = \det(x_1C_1 + \cdots + x_mC_m) \in k[x_1, \dots, x_m]$$

at $(x_1, \dots, x_m) = (t_1, \dots, t_m)$ is nonzero. Similarly, that $f^*\pi_1 \simeq f^*\pi_2$ means that there exists $C' \in M_n(k')$ such that

(a') For all $g \in G$, $C' \cdot \pi_1(g) = \pi_2(g) \cdot C'$.
(b') The determinant $\det(C')$ is nonzero.

But (a') is the same system of linear equations as (a), so Gauss elimination tells us that a general solution $C' \in M_n(k')$ can be written uniquely as

$$C' = t'_1C_1 + \cdots + t'_mC_m$$

with $(C_1, \dots, C_m)$ as before and with $(t'_1, \dots, t'_m)$ a family of scalars in the field k'. And (b') is the requirement that there exists a family $(t'_1, \dots, t'_m)$ of scalars in k' such that the value of the polynomial

$$p(x_1, \dots, x_m) \in k[x_1, \dots, x_m] \subset k'[x_1, \dots, x'_m]$$

is nonzero. Since k, and hence, k' is infinite, the k'-linear map

$$k'[x_1, \dots, x_m] \xrightarrow{\mathrm{ev}} \mathrm{Map}((k')^m, k')$$

is injective, so our assumption that $f^*\pi_1 \simeq f^*\pi_2$ implies that the polynomial

$$p(x_1, \dots, x_m) \in k[x_1, \dots, x_m]$$

is nonzero. But, since k is infinite, the k-linear map

$$k[x_1, \dots, x_m] \xrightarrow{\mathrm{ev}} \mathrm{Map}(k^m, k)$$

is injective, so we exists $(t_1, \dots, t_m) \in k^m$ such that $p(t_1, \dots, t_m) \neq 0$. This shows that $\pi_1 \simeq \pi_2$, as desired. $\square$

Suppose that $f \colon k \to k'$ is a Galois extension with Galois group

$$\Gamma = \mathrm{Aut}_k(k').$$

If V is a k-vector space, then the group homomorphism

$$\Gamma \xrightarrow{\rho} \mathrm{GL}(f_* f^*(V))$$

given by the formula

$$\rho(\gamma)(\boldsymbol{x} \otimes b) = \boldsymbol{x} \otimes \gamma(b)$$

with $\boldsymbol{x} \in V$ and $b \in k'$ defines a k-linear representation of Γ on $f_* f^*(V)$. Suppose that we also have a k-linear representation

$$G \xrightarrow{\pi} \mathrm{GL}(V)$$

of some group G on V. In this situation, the group homomoprhism

$$G \xrightarrow{f_* f^* \pi} \mathrm{GL}(f_* f^*(V))$$

also defines a k-linear representation of the group G on $f_* f^*(V)$. We note that for all $g \in G$, the k-linear isomorphism $f_* f^* \pi(g)$ is intertwining with respect to γ. Similarly, for all $\gamma \in \Gamma$, the k-linear isomorphism $\rho(\gamma)$ is intertwining with respect to $f_* f^*(\pi)$. Indeed, for all $\gamma \in \Gamma$, $g \in G$, $\boldsymbol{x} \in V$, and $b \in k'$, we have

$$\begin{aligned} \rho(\gamma)(f_* f^* \pi(g)(\boldsymbol{x} \otimes b)) &= \rho(\gamma)(\pi(g)(\boldsymbol{x}) \otimes b) \\ &= \pi(g)(\boldsymbol{x}) \otimes \gamma(b) \\ &= f_* f^* \pi(g)(\boldsymbol{x} \otimes \gamma(b)) \\ &= f_* f^* \pi(g)(\rho(\gamma)(\boldsymbol{x} \otimes b)). \end{aligned}$$

Equivalently, the map

$$G \times \Gamma \xrightarrow{\tau} \mathrm{GL}(f_* f^*(V))$$

given by

$$\tau(g, \gamma)(\boldsymbol{x} \otimes b) = \pi(g)(\boldsymbol{x}) \otimes \gamma(b)$$

is a group homomorphism and defines a representation of the group $G \times \Gamma$ on the k-vector space $f_* f^*(V)$. It follows that the subspace

$$W = (f_* f^*(V))^{\Gamma} = \{\boldsymbol{y} \in f_* f^*(V) \mid \rho(\gamma)(\boldsymbol{y}) = \boldsymbol{y} \text{ for all } \gamma \in \Gamma\} \subset f_* f^*(V)$$

is $f_* f^* \pi$-invariant. Moreover, the unit map

$$V \xrightarrow{\eta} f_* f^*(V)$$

is intertwining between π and $f_* f^* \pi$ and induces a map

$$V \xrightarrow{\tilde{\eta}} W = (f_* f^*(V))^{\Gamma}$$

that is intertwining between π and $(f_* f^* \pi)_W$.

Theorem 5.5. *Let $f\colon k \to k'$ be a finite Galois extension with group $\Gamma = \mathrm{Gal}(k'/k)$. If (V, π) is a k-linear representation of a group G, then the map*

$$V \xrightarrow{\tilde{\eta}} W = (f_* f^*(V))^{\Gamma}$$

is an isomorphism between π and $(f_* f^* \pi)_W$.

Proof. By faithfully flat descent for modules, the diagram

$$V \xrightarrow{\eta_V} f_* f^*(V) \underset{\eta_{f_* f^*(V)}}{\overset{f_* f^*(\eta_V)}{\rightrightarrows}} f_* f^* f_* f^*(V)$$

is an equalizer. This only uses that $f\colon k \to k'$ is faithfully flat, which is true for every extension of fields. That the diagram is an equalizer means that the map η_V is injective and that its image is equal to the subspace

$$W' = \{\boldsymbol{y} \in f_* f^*(V) \mid f_* f^*(\eta_V)(\boldsymbol{y}) = \eta_{f_* f^*(V)}(\boldsymbol{y})\} \subset f_* f^*(V).$$

So we wish to show that $W = W'$ and write out the diagram above as

$$V \xrightarrow{\eta_V} V \otimes_k k' \underset{\eta_{f_* f^*(V)}}{\overset{f_* f^*(\eta_V)}{\rightrightarrows}} V \otimes_k k' \otimes_k k'$$

with $\eta_V(\boldsymbol{x}) = \boldsymbol{x} \otimes 1$, $f_* f^*(\eta_V)(\boldsymbol{x} \otimes b) = \boldsymbol{x} \otimes 1 \otimes b$, and $\eta_{f_* f^*(V)}(\boldsymbol{x} \otimes b) = \boldsymbol{x} \otimes b \otimes 1$. The assumption that $f\colon k \to k'$ is a finite Galois extension with group Γ implies that the ring homomorphism

$$k' \otimes_k k' \xrightarrow{h} \prod_{\gamma \in \Gamma} k'$$

with γth component $h_\gamma(b_1 \otimes b_2) = b_1 \gamma(b_2)$ is an isomorphism. Thus the subspace $W' \subset V \otimes_k k'$ is equal to the equalizer of the two composite maps

$$V \otimes_k k' \underset{\eta_{f_* f^*(V)}}{\overset{f_* f^*(\eta_V)}{\rightrightarrows}} V \otimes_k k' \otimes_k k' \xrightarrow{V \otimes h} \prod_{\gamma \in \Gamma} V \otimes_k k'.$$

Finally, the γth components of the two composite maps are given by

$$((V \otimes h_\gamma) \circ f_* f^*(\eta_V))(\boldsymbol{x} \otimes b) = \boldsymbol{x} \otimes \gamma(b)$$
$$((V \otimes h_\gamma) \circ \eta_{f_* f^*(V)})(\boldsymbol{x} \otimes b) = \boldsymbol{x} \otimes b,$$

which shows that $W = W'$ as desired. □

This was rather abstract! Let us now specialize to the case

$$k = \mathbb{R} \xrightarrow{f} k' = \mathbb{C}$$

which is Galois with group $\Gamma = \mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{\mathrm{id}, \sigma\}$, where $\sigma\colon \mathbb{C} \to \mathbb{C}$ is complex conjugation. If V is a real vector space, then it is common to write

$$V_{\mathbb{C}} = f^*(V)$$

and call it the complexification of V. If V' is a complex vector space, then it is also common to abuse of notation and write V' for the real vector space $f_*(V')$. This is very confusing, however, since V' is a complex vector space, whereas $f_*(V')$ is a real vector space.

If V is a real vector space, then so is $f_*(V_{\mathbb{C}})$, and we have the $\mathbb{R}$-linear map

$$f_*(V_{\mathbb{C}}) \xrightarrow{\rho(\sigma)} f_*(V_{\mathbb{C}})$$

where $\sigma \in \Gamma = \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ is complex conjugation. We will also refer to this map as complex conjugation, and given $\boldsymbol{y} \in f_*(V_{\mathbb{C}})$, we write

$$\bar{\boldsymbol{y}} = \rho(\sigma)(\boldsymbol{y}).$$

If we write $\boldsymbol{y} = \sum \boldsymbol{x}_i \otimes z_i$ with $\boldsymbol{x}_i \in V$ and $z_i \in \mathbb{C}$, then $\bar{\boldsymbol{y}} = \sum \boldsymbol{x}_i \otimes \bar{z}_i$.

If $W \subset V_{\mathbb{C}}$ is a complex subspace, then so is its image

$$\overline{W} = \rho(\sigma)(W) \subset V_{\mathbb{C}}$$

under complex conjugation. Indeed, if $\bar{\boldsymbol{y}} = \rho(\sigma)(\boldsymbol{y}) \in \overline{W}$ and $z \in \mathbb{C}$, then also

$$\bar{\boldsymbol{y}} \cdot z = \rho(\sigma)(\boldsymbol{y}) \cdot z = \rho(\sigma)(\boldsymbol{y} \cdot \bar{z}) \in \overline{W}.$$

Lemma 5.6. *Let V be a real vector space, and let $W \subset V_{\mathbb{C}}$ be a complex subspace of its complexification. The following are equivalent.*

(1) *The complex subspaces $W, \overline{W} \subset V_{\mathbb{C}}$ are equal.*
(2) *There exists a real subspace $U \subset V$ such that $W = U_{\mathbb{C}} \subset V_{\mathbb{C}}$.*

Proof. It is clear that (2) implies (1), so we assume (1) holds and prove (2). The unit map $\eta_V \colon V \to f_*(V_{\mathbb{C}})$ is $\mathbb{R}$-linear, and we define

$$U = \eta_V^{-1}(f_*(W)) \subset V.$$

By Proposition 5.2, the $\mathbb{R}$-linear map

$$U \xrightarrow{\eta_V|_U} f_*(W)$$

determines and is determined by the $\mathbb{C}$-linear map

$$U_{\mathbb{C}} = f^*(U) \xrightarrow{\beta(\eta_V|_U)} W,$$

and we claim that the latter map is an isomorphism. It is injective, because the diagram commutes and because the left-hand vertical map is injective.[21]

$$\begin{array}{ccc} f^*(U) & \xrightarrow{\beta(\eta_V|_U)} & W \\ \big\downarrow & & \big\downarrow \\ f^*(V) & =\!=\!= & V_{\mathbb{C}} \end{array}$$

To prove that it is also surjective, let $\boldsymbol{y} \in W$. We have $\bar{\boldsymbol{y}} \in \overline{W}$, so by (1), we also have $\bar{\boldsymbol{y}} \in W$. It follows that both $\boldsymbol{u} = \frac{1}{2}(\boldsymbol{y} + \bar{\boldsymbol{y}})$ and $\boldsymbol{v} = \frac{1}{2i}(\boldsymbol{y} - \bar{\boldsymbol{y}})$ belong to W. But $\bar{\boldsymbol{u}} = \boldsymbol{u}$ and $\bar{\boldsymbol{v}} = \boldsymbol{v}$, so by Theorem 5.5, we have

$$\boldsymbol{u}, \boldsymbol{v} \in \mathrm{im}(V \xrightarrow{\eta_V} f_*(V_{\mathbb{C}})).$$

and since also $\boldsymbol{u}, \boldsymbol{v} \in W$, we have

$$\boldsymbol{u}, \boldsymbol{v} \in \mathrm{im}(U \xrightarrow{\eta_V|_U} f_*(W)).$$

[21] Here we use that $f \colon \mathbb{R} \to \mathbb{C}$ is flat, as is any field extension. Indeed, extension of scalars along a ring homomorphism $f \colon A \to B$ preserves monomorphisms if and only if $f \colon A \to B$ is flat.

But this shows that

$$\boldsymbol{y} = \boldsymbol{u} + i\boldsymbol{v} \in \operatorname{im}(U_{\mathbb{C}} = f^*(U) \xrightarrow{\beta(\eta_V|_U)} W)$$

as desired. ☐

Example 5.7. We recall the real representation $(\mathbb{R}^2, \pi)$ of the additive group of real numbers $G = (\mathbb{R}, +)$ defined by

$$\pi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

It is an irreducible representation, but its complexification $\pi_{\mathbb{C}}$ is not. Indeed, in the basis $(\boldsymbol{e}_1 + i\boldsymbol{e}_2, \boldsymbol{e}_1 - i\boldsymbol{e}_2)$ of $(\mathbb{R}^2)_{\mathbb{C}} \simeq \mathbb{C}^2$, we have

$$\pi_{\mathbb{C}}(t) = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}.$$

Theorem 5.8. *Let $\pi \colon G \to \mathrm{GL}(V)$ be a real representation.*

(1) *If $\pi_{\mathbb{C}}$ is irreducible, then so is π.*
(2) *If π is irreducible, then either $\pi_{\mathbb{C}}$ is irreducible or a sum of two irreducible representations, which are each other's conjugate.*
(3) *The represention π is semisimple if and only if $\pi_{\mathbb{C}}$ is so.*

Proof. (1) If $U \subset V$ is π-invariant, then $U_{\mathbb{C}} \subset V_{\mathbb{C}}$ is $\pi_{\mathbb{C}}$-invariant. So $U_{\mathbb{C}}$ is equal to either $\{\mathbf{0}\}$ or $V_{\mathbb{C}}$, which shows that U is equal to either $\{\mathbf{0}\}$ or V as desired.
(2) Let $W \subset V_{\mathbb{C}}$ be a $\pi_{\mathbb{C}}$-invariant subspace with $\pi_{\mathbb{C},W}$ irreducible. In this situation,

$$W \cap \overline{W}, W + \overline{W} \subset V_{\mathbb{C}}$$

are both $\pi_{\mathbb{C}}$-invariant subspaces, and since

$$\overline{W \cap \overline{W}} = \overline{W} \cap \overline{\overline{W}} = \overline{W} \cap W$$
$$\overline{W + \overline{W}} = \overline{W} + \overline{\overline{W}} = \overline{W} + W,$$

it follows from Lemma 5.6 that both are complexifications of real subspaces of V. By the assumption that V is irreducible, the only possibilities are that

(i) $W \cap \overline{W} = W + \overline{W} = \{\mathbf{0}\}$,
(ii) $W \cap \overline{W} = \{\mathbf{0}\} \subset W + \overline{W} = V_{\mathbb{C}}$, or
(iii) $W \cap \overline{W} = W + \overline{W} = V_{\mathbb{C}}$.

In case (i), we have $W = \{\mathbf{0}\}$, in (ii), the map $W \oplus \overline{W} \to V_{\mathbb{C}}$ induced by the canonical inclusions is an isomorphism; and in case (iii), we have $W = V_{\mathbb{C}}$. This proves (2). Finally, (3) follows immediately from (1) and (2). ☐

Example 5.9. Let $\pi \colon \Sigma_3 \to \mathrm{GL}(\mathbb{R}^3)$ be the standard (permutation) representation of the symmetric group. The subspaces

$$V_1 = \{\boldsymbol{x} \in \mathbb{R}^3 \mid x_1 = x_2 = x_3\} \subset \mathbb{R}^3$$
$$V_2 = \{\boldsymbol{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\} \subset \mathbb{R}^3$$

are π-invariant, and moreover, the representations $\pi_1 = \pi_{V_1}$ and $\pi_2 = \pi_{V_2}$ are both irreducible and π_2 is faithful. We claim that $\pi_{2,\mathbb{C}}$ is irreducible. If not, then it is a sum of two irreducible representations, and since

$$\dim_{\mathbb{C}}(V_{2,\mathbb{C}}) = \dim_{\mathbb{R}}(V_2) = 2,$$

each of these two irreducible representations much be 1-dimensional. But $\pi_{2,\mathbb{C}}$ is again faithful,[22] so this would give an injective group homomorphism

$$\Sigma_3 \longrightarrow \mathrm{GL}_1(\mathbb{C}) \times \mathrm{GL}_1(\mathbb{C}),$$

which is impossible, since Σ_3 is non-abelian, while $\mathrm{GL}_1(\mathbb{C}) \times \mathrm{GL}_1(\mathbb{C})$ is abelian.

[22] This is true, because $f\colon \mathbb{R} \to \mathbb{C}$ is faithful

Let (V_1, π_1) and (V_2, π_2) be k-linear representations of a group G and recall that a k-linear map $f\colon V_1 \to V_2$ is intertwining between π_1 and π_2 if the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{f} & V_2 \\ \downarrow{\scriptstyle \pi_1(g)} & & \downarrow{\scriptstyle \pi_2(g)} \\ V_1 & \xrightarrow{f} & V_2 \end{array}$$

commutes for all $g \in G$. We will write

$$\mathrm{Hom}(\pi_1, \pi_2) \subset \mathrm{Hom}_k(V_1, V_2)$$

for the subspace of intertwining k-linear maps.

Theorem 6.1 (Schur's lemma)**.** *Let (V_1, π_1) and (V_2, π_2) be two irreducible k-linear representations of a group G. A k-linear map $f\colon V_1 \to V_2$ that is intertwining between π_1 and π_2 is either an isomorphism or the zero map.*

Proof. It follows immediately from the fact that $f\colon V_1 \to V_2$ is intertwining between π_1 and π_2, that $\ker(f) \subset V_1$ is π_1-invariant and that $\mathrm{im}(f) \subset V_2$ is π_2-invariant. Therefore, since π_1 was assumed to irreducible, either $\ker(f) = \{\mathbf{0}\}$ or $\ker(f) = V_1$, and since π_2 was assumed to be irreducible, either $\mathrm{im}(f) = \{\mathbf{0}\}$ or $\mathrm{im}(f) = V_2$. □

Theorem 6.2. *Suppose that $((V_s, \pi_s))_{s \in S}$ is a finite family of pairwise non-isomorphic irreducible k-linear representations of a group G, and that $U \subset \bigoplus_{s \in S} V_s$ is a $\bigoplus_{s \in S} \pi_s$-invariant subspace. There exists a (unique) subset $T \subset S$ such that*

$$U = \textstyle\bigoplus_{t \in T} V_t \ \subset\ \bigoplus_{s \in S} V_s.$$

Proof. Let $V = \bigoplus_{s \in S} V_s$, let $\pi = \bigoplus_{s \in S} \pi_s$, and let $i_s\colon V_s \to V$ be the canonical inclusion, which is intertwining between π_s and π. Theorem 4.12 shows that $U \subset V$ is the image of *some* injective k-linear map

$$\textstyle\bigoplus_{t \in T} V_t \xrightarrow{f} V$$

that is intertwining between $\bigoplus_{t \in T} \pi_t$ and π, and we wish to show that the map

$$\textstyle\bigoplus_{t \in T} V_t \xrightarrow{i = \sum_{t \in T} i_t} V$$

will do. Let $p_s\colon V \to V_s$ be the unique map such that $p_s \circ i_t\colon V_t \to V_s$ is the identity map of V_s if $s = t$ and the zero map if $s \neq t$. We consider the composite maps

$$\begin{array}{ccc} \bigoplus_{u \in T} V_u & \xrightarrow{f} & V \\ \uparrow{\scriptstyle i_t} & & \downarrow{\scriptstyle p_s} \\ V_t & \xrightarrow{f_{s,t}} & V_s \end{array}$$

for $s \in S$ and $t \in T$. Theorem 6.1 shows that $f_{s,t}$ is zero if $s \neq t$, so the diagram

$$\begin{array}{ccc} \bigoplus_{t \in T} V_t & \xrightarrow{\quad f \quad} & \\ \Big\downarrow {\scriptstyle \bigoplus_{t \in T} f_{t,t}} & & V \\ \bigoplus_{t \in T} V_t & \xrightarrow{\quad i \quad} & \end{array}$$

commutes. Moreover, the maps $f_{t,t}$ cannot be zero, since the top slanted map is injective, so Theorem 6.1 shows that the $f_{t,t}$ all are isomorphisms, and hence, the left-hand vertical map is an isomorphism. In particular,

$$U = \operatorname{im}(f) = \operatorname{im}(i),$$

as we wanted to prove. □

If (V, π) is an irreducible k-linear representation of G, then Schur's lemma shows, in particular, that the endomorphism ring

$$\operatorname{End}(\pi) \subset \operatorname{End}_k(V)$$

is a division algebra D over k. In general, every finite dimensional division algebra over k occurs as $\operatorname{End}(\pi)$ for a finite dimensional irreducible k-linear representation of some group G.[23] We now make the very simplifying assumption that

$$k = \bar{k}$$

is an algebraically closed field, so that, up to unique isomorphism, the only finite dimensional division algebra over k is $D = k$. In this case, Schur's lemma implies the following result, which is also known as Schur's lemma.

Theorem 6.3. *Let k be an algebraically closed field. If (V, π) is a finite dimensional irreducible k-linear representation of a group G, then the map*

$$k \xrightarrow{\ \eta\ } \operatorname{End}(\pi)$$

defined by $\eta(\lambda) = \lambda \cdot \mathrm{id}_V$ is a ring isomorphism.

Proof. The map η is injective, because V is nonzero, and to prove that it surjective, we let $f \colon V \to V$ be a k-linear map that is intertwining with respect to π. Since k is algebraically closed, the map f has an eigenvalue $\lambda \in k$, and since f is intertwining with respect to π, the eigenspace

$$\{\mathbf{0}\} \neq V_\lambda \subset V$$

is a π-invariant subspace. Since π is irreducible, we conclude that $V_\lambda = V$, or equivalently, that $f = \lambda \cdot \mathrm{id}_V$, which shows that η is surjective. □

Corollary 6.4. *Let (V_1, π_1) and (V_2, π_2) be isomorphic finite dimensional irreducible k-linear representations of a group G. If k is algebraically closed, then*

$$\dim_k \operatorname{Hom}(\pi_1, \pi_2) = 1.$$

[23] If D finite dimensional real division algebra, then $D \simeq \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$.

Proof. We choose any k-linear isomorphism $h\colon V_1 \to V_2$ that intertwines between π_1 and π_2. (Such an h exists by assumption.) Now, if $f\colon V_1 \to V_2$ is any k-linear map that intertwines between π_1 and π_2, then $f \circ h^{-1}\colon V_2 \to V_2$ is intertwining with respect to π_2. Hence, by Theorem 6.3, $f \circ h^{-1} = \lambda \cdot \mathrm{id}_{V_2}$ for a unique $\lambda \in k$, so we find that $f = \lambda \cdot h$. ☐

Remark 6.5. More precisely, Corollary 6.4 shows that the composition maps

$$\begin{aligned} \mathrm{Hom}(\pi_1, \pi_2) \times \mathrm{End}(\pi_1) &\xrightarrow{\circ} \mathrm{Hom}(\pi_1, \pi_2) \\ \mathrm{End}(\pi_2) \times \mathrm{Hom}(\pi_1, \pi_2) &\xrightarrow{\circ} \mathrm{Hom}(\pi_1, \pi_2) \end{aligned}$$

simultaneously make $\mathrm{Hom}(\pi_1, \pi_2)$ a free right $\mathrm{End}(\pi_1)$-module of rank 1 and a free left $\mathrm{End}(\pi_2)$-module of rank 1. However, neither module has a preferred generator: There is no preferred way to compare π_1 and π_2.

In Theorem 6.2, we considered a finite sum of pairwise non-isomorphic irreducible representations. In the next result, we will consider the opposite situation of a finite sum of irreducible representations, all of which are isomorphic.

Theorem 6.6. *Let k be an algebraically closed field, and let (U, τ) and (V, π) be finite dimensional k-linear representations of a group G such that τ is trivial and π is irreducible. Given a $\pi \otimes \tau$-invariant subspace $W \subset V \otimes U$ with $(\pi \otimes \tau)_W$ irreducible, there exists (a non-unique) vector $\boldsymbol{u} \in U$ such that the map*

$$V \xrightarrow{i_{\boldsymbol{u}}} V \otimes U$$

defined by $i_{\boldsymbol{u}}(\boldsymbol{v}) = \boldsymbol{v} \otimes \boldsymbol{u}$ is an isomorphism from V onto W and is intertwining between π and $(\pi \otimes \tau)_W$.

Proof. The map $i_{\boldsymbol{u}}$ is clearly k-linear and intertwining between π and $\pi \otimes \tau$, so we must show that $\boldsymbol{u} \in U$ can be chosen such that $W = i_{\boldsymbol{u}}(V) \subset V \otimes U$. We have seen earlier that every irreducible subrepresentations of $\pi \otimes \tau$ is isomorphic to π. In particular, we may choose a k-linear isomorphism $h\colon W \to V$ that intertwines between $(\pi \otimes \tau)_W$ and π. Now, for every $f \in U^*$, we let

$$V \otimes U \xrightarrow{c_f} V$$

to be the unique k-linear map such that $c_f(\boldsymbol{v} \otimes \boldsymbol{u}) = \boldsymbol{v} \cdot f(\boldsymbol{u})$. It intertwines between $\pi \otimes \tau$ and π, and since k is algebraically closed, Corollary 6.4 shows that

$$c_f|_W = \lambda(f) \cdot h$$

for a unique $\lambda(f) \in k$. The map

$$U^* \xrightarrow{\lambda} k$$

that to f assigns $\lambda(f)$ is k-linear, so $\lambda \in U^{**}$. But the map

$$U \xrightarrow{\eta} U^{**}$$

is an isomorphism, since U is finite dimensional, so $\lambda = \eta(\boldsymbol{u})$ for a unique $\boldsymbol{u} \in U$. We claim that for this $\boldsymbol{u} \in U$, we have $i_{\boldsymbol{u}}(V) = W \subset V \otimes U$. To prove this, we choose a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V and write $\boldsymbol{w} \in W$ as

$$\boldsymbol{w} = \textstyle\sum_{1 \leq i \leq n} \boldsymbol{v}_i \otimes \boldsymbol{u}_i$$

with $\boldsymbol{u}_i \in U$. (This way of writing $\boldsymbol{w}$ is unique, but we will not need this fact.) Now, for every $f \in U^*$, we have

$$\textstyle\sum_{1 \leq i \leq n} \boldsymbol{v}_i f(\boldsymbol{u}_i) = c_f(\boldsymbol{w}) = \lambda(f) \cdot h(\boldsymbol{w}) = f(\boldsymbol{u}) \cdot h(\boldsymbol{w}).$$

Therefore, if $f(\boldsymbol{u}) = 0$, then $f(\boldsymbol{u}_i) = 0$ for all $1 \leq i \leq n$, and hence,

$$\boldsymbol{u}_i \in \operatorname{span}_k(\boldsymbol{u}) \subset U.$$

This shows that $\{\boldsymbol{0}\} \neq W \subset i_{\boldsymbol{u}}(V) \subset V \otimes U$, and since $i_{\boldsymbol{u}}(V) \simeq V$ is irreducible, we conclude that $W = i_{\boldsymbol{u}}(V)$ as stated. Note that the vector $\boldsymbol{u} \in U$ depends on the choice of the isomorphism $h \colon W \to V$. □

Abelian groups

We will next show that if k is algebraically closed, then every finite dimensional k-linear irreducible representation of an abelian group is 1-dimensional. We have seen that the representation of the group $(\mathbb{R}, +)$ on $\mathbb{R}^2$ by rotation is irreducible, so the assumption that k be algebraically closed is necessary.

Theorem 6.7. *Let k be an algebraically closed field. If (V, π) is a finite dimensional irreducible k-linear representation of an abelian group A, then*

$$\dim_k(V) = 1.$$

Proof. Since A is abelian, we have

$$\pi(g) \circ \pi(h) = \pi(g \cdot h) = \pi(h \cdot g) = \pi(h) \circ \pi(g)$$

for all $g, h \in A$. Hence, for all $g \in A$, the k-linear map

$$V \xrightarrow{\pi(g)} V$$

is intertwining with respect to π, so by Schur's lemma, we have

$$\pi(g) = \lambda \cdot \mathrm{id}_V$$

for some $\lambda = \lambda(g) \in k$. But this implies that every subspace $W \subset V$ is π-invariant, and since V is irreducible, this shows that $\dim_k(V) = 1$. □

We recall that the abelianization of a group G is a group homomorphism

$$G \xrightarrow{p} G^{\mathrm{ab}}$$

with the property that for every group homomorphism $f \colon G \to A$ with A abelian, there exists a unique group homomorphism $f^{\mathrm{ab}} \colon G^{\mathrm{ab}} \to A$ such that $f = f^{\mathrm{ab}} \circ p$. This property characterizes the abelianization $p \colon G \to G^{\mathrm{ab}}$ uniquely, up to unique isomorphism under G. The group homomorphism p is surjective, and its kernel is the commutator subgroup $[G, G] \subset G$. In particular, any 1-dimensional k-linear representation $\pi \colon G \to \mathrm{GL}(V)$ of a group G determines and is determined by the 1-dimensional k-linear representation $\pi^{\mathrm{ab}} \colon G^{\mathrm{ab}} \to \mathrm{GL}(V)$ of G^{ab}. Moreover, if k is algebraically closed, then a finite dimensional k-linear representation of G^{ab} is 1-dimensional if and only if it is irreducible.

Example 6.8. The abelianization of the symmetric group Σ_n is the signature

$$\Sigma_n \xrightarrow{\text{sgn}} \{\pm 1\}.$$

Therefore, up to non-canonical isomorphism, the only 1-dimensional representations of Σ_n over an algebraically closed field k are the trivial representation and the sign representation.

Exterior tensor product

We have seen that tensor products of irreducible k-linear representations are typically not irreducible, even if k is algebraically closed. We now show that the exterior tensor product of irreducible k-linear representations is always irreducible.

Theorem 6.9. *Let k be an algebraically closed field. If both $\pi_1 \colon G_1 \to \mathrm{GL}(V_1)$ and $\pi_2 \colon G_2 \to \mathrm{GL}(V_2)$ are finite dimensional irreducible k-linear representations, then so is their exterior tensor product*

$$G_1 \times G_2 \xrightarrow{\pi_1 \boxtimes \pi_2} \mathrm{GL}(V_1 \otimes V_2).$$

Proof. We let $\{\mathbf{0}\} \neq W \subset V_1 \otimes V_2$ be a $\pi_1 \boxtimes \pi_2$-invariant subspace and must show that $W = V_1 \otimes V_2$. We note that

$$\begin{aligned} (\pi_1 \boxtimes \pi_2)(g_1, e) &= (\pi_1 \otimes \tau)(g_1) \\ (\pi_1 \boxtimes \pi_2)(e, g_2) &= (\tau \otimes \pi_2)(g_2), \end{aligned}$$

so $W \subset V_1 \otimes V_2$ is both a $\pi_1 \otimes \tau$-invariant subspace of the k-linear representation $\pi_1 \otimes \tau$ of $G_1 \times \{e\} \subset G_1 \times G_2$ and a $\tau \otimes \pi_2$-invariant subspace of the k-representation $\tau \otimes \pi_2$ of $\{e\} \times G_2 \subset G_1 \times G_2$. Hence, by Theorem 6.6, there exists $U_2 \subset V_2$ and $U_1 \subset V_1$ such that both the (injective) maps induced by the canonical inclusions

$$V_1 \otimes U_2 \longrightarrow V_1 \otimes V_2 \longleftarrow U_1 \otimes V_2$$

have image W. So the square diagram of inclusions

$$\begin{array}{ccc} U_1 \otimes U_2 & \longrightarrow & V_1 \otimes U_2 \\ \downarrow & & \downarrow \\ V_1 \otimes U_2 & \longrightarrow & W \end{array}$$

is cocartesian, and the right-hand vertical map and the lower horizontal are both isomorphisms. This implies (by the five-lemma) that

$$(V_1/U_1) \otimes U_2 \simeq \{\mathbf{0}\} \simeq U_1 \otimes (V_2/U_2),$$

which, in turn, implies that $U_1 = V_1$ and $U_2 = V_2$. So $W = V_1 \otimes V_2$ as desired. $\square$

Example 6.10. We show that, in Theorem 6.9, the assumption that k be algebraically closed is necessary. The representation $\pi \colon G = (\mathbb{R}, +) \to \mathrm{GL}(\mathbb{R}^2)$ defined by

$$\pi(t) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is irreducible, but we claim that $\pi \boxtimes \pi$ is not. Indeed, if $\pi \boxtimes \pi$ were irreducible, then also $(\pi \boxtimes \pi)_{\mathbb{C}}$ would either be irreducible or a sum of two irreducible representations. But this contradicts Theorem 6.7, since $G \times G$ is abelian and

$$\dim_{\mathbb{C}}((\pi \boxtimes \pi)_{\mathbb{C}}) = \dim_{\mathbb{R}}(\pi \boxtimes \pi) = 4.$$

Let k be a field, and let G be a group. We recall that $k[G]$ is k-vector space of all functions $f \colon G \to k$ and that

$$G \xrightarrow{L,R} \mathrm{GL}(k[G])$$

are the left and right regular k-linear representations defined by

$$L(g)(f)(h) = f(g^{-1}h)$$
$$R(g)(f)(h) = f(hg).$$

Since the maps $R(g_1)$ and $L(g_2)$ commute, we obtain a representation

$$G \times G \xrightarrow{\mathrm{Reg}} \mathrm{GL}(k[G])$$

defined by $\mathrm{Reg}(g_1, g_2) = L(g_2) \circ R(g_1) = R(g_1) \circ L(g_2)$. We call this representation the two-sided regular representation. Spelling out the definition, we have

$$\mathrm{Reg}(g_1, g_2)(f)(h) = f(g_2^{-1} h g_1).$$

Given any k-linear representation

$$G \xrightarrow{\pi} \mathrm{GL}(V),$$

the k-linear map

$$V \otimes V^* \xrightarrow{\mu} k[G]$$

defined by $\mu(\boldsymbol{x} \otimes f)(h) = f(\pi(h)(\boldsymbol{x}))$ is intertwining between $\pi \boxtimes \pi^*$ and Reg. We define the space of matrix coefficients (or matrix elements) of π to be its image

$$M(\pi) = \mu(V \otimes V^*) \subset k[G].$$

The reason for this name is as follows. Suppose that V is finite dimensional. If we let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V, let $(\boldsymbol{v}_1^*, \dots, \boldsymbol{v}_n^*)$ be the dual basis of V^*, and let

$$A(h) = (a_{ij}(h)) \in M_n(k)$$

be the matrix that represents $\pi(h)$ with respect to $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$, then

$$a_{ij}(h) = \mu(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)(h).$$

This shows that that space of matrix coefficients is the subspace

$$M(\pi) = \mathrm{span}_k(a_{ij} \mid 1 \leq i, j \leq n) \subset k[G]$$

spanned by the functions $a_{ij} \colon G \to k$, whence the name. The reason that we do not take this formula as our definition of $M(\pi)$ is that it is not a priori clear that the subspace $\mathrm{span}_k(a_{ij}) \subset k[G]$ is independent of the choice of basis.

Theorem 6.11. *Let k be an algebraically closed field. If (V, π) is a finite dimensional irreducible k-linear representation of a group G, then*

$$\pi \boxtimes \pi^* \xrightarrow{\mu} \mathrm{Reg}_{M(\pi)}$$

is an isomorphism of k-linear representations of $G \times G$. In particular, the k-linear representation $\mathrm{Reg}_{M(\pi)}$ is irreducible.

Proof. The map $\mu\colon V \otimes V^* \to M(\pi)$ is surjective, by definition, and it intertwines between $\pi \boxtimes \pi^*$ and $\mathrm{Reg}_{M(\pi)}$. Since π is finite dimensional and irreducible, the same is true for π^*, and since k is algebraically closed, Theorem 6.9 shows that $\pi \boxtimes \pi^*$ is irreducible. Hence, the kernel of μ is either zero or all of $V \otimes V^*$. But it is easy to see that $\mu\colon V \otimes V^* \to k[G]$ is not the zero map. Indeed, choosing a basis of V as above, we see that $a_{ii}(e) = 1$, so $0 \neq a_{ii} \in M(\pi) = \mathrm{im}(\mu)$. So μ is injective. $\square$

We list some consequences of Theorem 6.11:

(1) If k is algebraically closed and if π is an irreducible k-linear representation of finite dimension n, then
$$\dim_k(M(\pi)) = n^2.$$
(2) For π as in (1), we have $R_{M(\pi)} \simeq \pi \oplus \cdots \oplus \pi$ and $L_{M(\pi)} \simeq \pi^* \oplus \cdots \oplus \pi^*$, where there are n summands in both cases.
(3) If k is algebraically closed and if π_1 and π_2 are finite dimensional irreducible k-linear representations, then $\mathrm{Reg}_{M(\pi_1)} \simeq \mathrm{Reg}_{M(\pi_2)}$ implies that $\pi_1 \simeq \pi_2$.
(4) If k is algebraically closed and if $\pi_1, \dots, \pi_m$ are pairwise non-isomorphic finite dimensional irreducible k-linear representations of G, then
$$M(\pi_1) \oplus \cdots \oplus M(\pi_m) \longrightarrow k[G]$$
is injective.

Unitary representations

Our final application of Schur's lemma concerns unitary representations. A finite dimensional complex representation (V, π) of a group G is unitary if there exists a hermitian inner product $\langle -, - \rangle$ on V that is π-invariant in the sense that
$$\langle \pi(g)(\boldsymbol{x}), \pi(g)(\boldsymbol{y}) \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle,$$
for all $g \in G$ and $\boldsymbol{x}, \boldsymbol{y} \in V$, or equivalently, if the induced isomorphism
$$\bar{V} \xrightarrow{\;b\;} V^*$$
given by $b(\boldsymbol{x})(\boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is intertwining between $\bar{\pi}$ and π^*. We will now show that a π-invariant hermitian inner product is unique, up to scaling.

Theorem 6.12. *Suppose that $\pi\colon G \to \mathrm{GL}(V)$ is a finite dimensional irreducible unitary representation. If both $\langle -, - \rangle_1$ and $\langle -, - \rangle_2$ are π-invariant hermitian inner products on V, then there exists a real number $\lambda > 0$ such that for all $\boldsymbol{x}, \boldsymbol{y} \in V$,*
$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_2 = \lambda \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle_1.$$

Proof. We define $h\colon \bar{V} \to \bar{V}$ to be the composite isomorphism
$$\bar{V} \xrightarrow{\;b_1\;} V^* \xrightarrow{\;(b_2)^{-1}\;} \bar{V}.$$
By assumption, both b_1 and b_2 are intertwining with respect to $\bar{\pi}$ and π^*, so h is intertwining with respect to $\bar{\pi}$. Finally, since $\bar{\pi}$ is irreducible, Schur's lemma shows that $h = \lambda \cdot \mathrm{id}_{\bar{V}}$ for some $\lambda \in \mathbb{C}$, and $\lambda \neq 0$, because h is an isomorphism. Finally, for any $\boldsymbol{0} \neq \boldsymbol{x} \in V$, both $\langle \boldsymbol{x}, \boldsymbol{x} \rangle_1$ and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle_2$ are positive real numbers, so λ is necessarily real and positive. $\square$

Theorem 6.13. *Let (V, π) be a finite dimensional unitary representation of a group G, and suppose that $U_1, U_2 \subset V$ are π-invariant subspaces with the property that the representations $\pi_1 = \pi_{U_1}$ and $\pi_2 = \pi_{U_2}$ are non-isomorphic and irreducible. In this situation, the subspaces $U_1, U_2 \subset V$ are necessarily orthogonal with respect to any π-invariant hermitian inner product on V.*

Proof. We choose a π-invariant hermitian inner product $\langle -, - \rangle$ on V. Since $U_1 \subset V$ is π-invariant, so is its orthogonal complement $W_1 \subset V$ with respect to $\langle -, - \rangle$, and moreover, the composition of the canonical inclusion and the canonical projection

$$U_1 \xrightarrow{i_1} V \xrightarrow{q_1} V/W_1$$

is a complex linear isomorphism $h = q_1 \circ i_1$ that intertwines between π_{U_1} and π_{V/W_1}. The orthogonal projection $p\colon V \to U_1$ with respect to $\langle -, - \rangle$ is the composition

$$V \xrightarrow{q_1} V/W_1 \xrightarrow{h^{-1}} U_1,$$

so it is intertwining between π and π_1. Now, the composite map

$$U_2 \xrightarrow{i_2} V \xrightarrow{p_1} U_1$$

is intertwining between π_2 and π_1, and since these representations are assumed to be irreducible and non-isomorphic, it follows from Schur's lemma that the composite map is zero. This shows that the subspaces $U_1, U_2 \subset V$ are orthogonal with respect to $\langle -, - \rangle$, as stated. $\square$

Let us first show that, for every field k an irreducible k-linear representation of a finite group G is necessarily finite dimensional.

Lemma 7.1. *Let k be a field. If a k-linear representation (V, π) of a finite group G is irreducible, then the k-vector space V is finite dimensional.*

Proof. Let (V, π) be an irreducible k-linear representation of G. Since V is nonzero, there exists a nonzero vector $\boldsymbol{x} \in V$, so the subspace $W \subset V$ spanned by the family $(\pi(g)(\boldsymbol{x}))_{g \in G}$ is nonzero. But it is also π-invariant, so $W = V$, by the assumption that π is irreducible. Since G is finite, the family $(\pi(g)(\boldsymbol{x}))_{g \in G}$ is a finite family, so $W = V$ is a finite generated, and hence, finite dimensional k-vector space. □

So let G be a finite group. Because of Lemma 7.1, we will only consider finite dimensional k-linear representations of G. We will also assume that k is algebraically closed. Since G is finite, a basis of $k[G]$ is given by the family $(\delta_x)_{x \in G}$, where

$$\delta_x(y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$$

We let (V, π) be a finite dimensional k-linear representation of G and recall that the subspace of matrix coefficients

$$M(\pi) \subset k[G]$$

is defined to be the image of the map

$$V \otimes V^* \xrightarrow{\mu} k[G]$$

given by $\mu(\boldsymbol{x} \otimes \varphi)(h) = \varphi(\pi(h)(\boldsymbol{x}))$. It is a Reg-invariant subspace, where

$$G \times G \xrightarrow{\text{Reg}} \mathrm{GL}(k[G])$$

is the two-sided regular representation of $G \times G$ on V defined by

$$\mathrm{Reg}(g_1, g_2)(f)(h) = f(g_2^{-1} h g_1).$$

Since k is algebraically closed, Schur's lemma implies the following statements, which we proved last time.

(1) If π is irreducible, then $\mathrm{Reg}_{M(\pi)} \simeq \pi \boxtimes \pi^*$.
(2) If π_1 and π_2 are irreducible, then $\pi_1 \simeq \pi_2$ if and only if $M(\pi_1) = M(\pi_2)$.
(3) If $\pi_1, \dots, \pi_m$ are pairwise non-isomorphic and irreducible, then the map

$$M(\pi_1) \oplus \cdots \oplus M(\pi_m) \longrightarrow k[G]$$

induced by the canonical inclusions is injective.

Theorem 7.2. *If G is a finite group and if k is an algebraically closed field, then G has at most $|G|$ pairwise non-isomorphic irreducible k-linear representations.*

Proof. It follows from (3) that

$$q \leq \dim_k(M(\pi_1) \oplus \cdots \oplus M(\pi_q)) \leq \dim_k(k[G]) = |G|,$$

which proves the theorem. □

Theorem 7.3. *Let G be a finite group, and let k be an algebraically closed field of characteristic zero. If $\pi_1, \dots, \pi_q$ are representatives of the isomorphism classes of irreducible k-linear representations, then the map*

$$M(\pi_1) \oplus \cdots \oplus M(\pi_q) \longrightarrow k[G]$$

induced by the canonical inclusions is an isomorphism.

Proof. The map is injective by (3) above, so it remains to prove that it is also surjective. Let $R\colon G \to \mathrm{GL}(k[G])$ be the right regular representation, which, we recall, is defined by $R(g)(f)(h) = f(hg)$. We claim that

$$M(R) = k[G].$$

Indeed, let $\epsilon\colon k[G] \to k$ be the k-linear map defined by $\epsilon(f) = f(e)$. So $\epsilon \in k[G]^*$ and for all $f \in k[G]$, the calculation

$$\mu(f \otimes \epsilon)(h) = \epsilon(R(h)(f)) = f(e \cdot h) = f(h)$$

shows that $f = \mu(f \otimes \epsilon) \in M(R)$. Now, since G is a finite group, whose order $|G|$ is not divisible by the characteristic of k, it follows from Maschke's theorem that every finite dimensional k-linear representation of G is semisimple. So

$$R \simeq \pi_1^{n_1} \oplus \cdots \oplus \pi_q^{n_q}.$$

But if ρ and τ are any finite dimensional k-linear representations of G, then

$$M(\rho \oplus \tau) = M(\rho) + M(\tau) \subset k[G].$$

Therefore, we conclude that

$$k[G] = M(R) \subset M(\pi_1) + \cdots + M(\pi_q) \subset k[G],$$

which shows the surjectivity of the map in the statement. □

Addendum 7.4. *Let G be a finite group, and let k be an algebraically closed field of characteristic zero. If $(V_1, \pi_1), \dots, (V_q, \pi_q)$ are representatives of the isomorphism classes of irreducible k-linear representations, then*

$$|G| = n_1^2 + \cdots + n_q^2,$$

where $n_i = \dim_k(V_i)$.

Proof. By Theorem 7.3, we have

$$|G| = \dim_k(k[G]) = \dim_k(M(\pi_1)) + \cdots + \dim_k(M(\pi_q)),$$

and $\dim_k(M(\pi_i)) = n_i^2$. Indeed, since k is algebraically closed and π_i irreducible, the map $\mu_{\pi_i}\colon V_i \otimes V_i^* \to M(\pi_i)$ is an isomorphism, and $\dim_k(V_i^*) = \dim_k(V_i) = n_i$. □

Example 7.5. Let k be algebraically closed of characteristic zero.

1) A finite abelian group A has precisely $|A|$ pairwise non-isomorphic irreducible k-linear representations, all of which 1-dimensional.

2) Let $G = \Sigma_3$. We have found three pairwise non-isomorphic irreducible k-linear representations of G, namely,

(i) the 1-dimensional trivial representation τ,
(ii) the 1-dimensional sign representation sgn, and

(iii) the 2-dimensional representation π of G on

$$V = \{\boldsymbol{x} \in k^3 \mid x_1 + x_2 + x_3 = 0\} \subset k^3$$

defined by

$$\pi(\sigma)(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = \begin{pmatrix} x_{\sigma(1)} \\ x_{\sigma(2)} \\ x_{\sigma(3)} \end{pmatrix}.$$

Since we have

$$|G| = 6 = 1^2 + 1^2 + 2^2 = \dim_k(\tau)^2 + \dim_k(\mathrm{sgn})^2 + \dim_k(\pi)^2,$$

we conclude from Addendum 7.4 that, up to non-canonical isomorphism, we have found all irreducible representations of G.

We next prove a lemma concerning matrix coefficients.

Lemma 7.6. *Let G be a group, let k be a field, and let (V, π) be a finite dimensional k-linear representation of G. There is a commutative diagram*

$$\begin{array}{ccc} \mathrm{End}_k(V) & \xrightarrow{\ \mu'\ } & k[G] \\ & {\scriptstyle\alpha}\nwarrow \quad \nearrow{\scriptstyle\mu} & \\ & V \otimes V^* & \end{array}$$

with $\mu'(f)(h) = \mathrm{tr}(\pi(h) \circ f)$ and $\alpha(\boldsymbol{x} \otimes \varphi)(\boldsymbol{y}) = \boldsymbol{x} \cdot \varphi(\boldsymbol{y})$, and moreover, the map α is an isomorphism. Accordingly, the subspace of matrix coefficients

$$M(\pi) \subset k[G]$$

is equal to the common image of μ and μ'.

Proof. Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V, and let $(\boldsymbol{v}_1^*, \dots, \boldsymbol{v}_n^*)$ be the dual basis of the dual k-vector space V^*. For $g \in G$, let

$$A(g) = (a_{ij}(g)) \in M_n(k)$$

be the matrix that represents $\pi(g)$ with respect to the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V. As we have calculated before, we have

$$\mu(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)(g) = a_{ij}(g).$$

Now, we have $\alpha(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*) = \boldsymbol{v}_j \cdot \boldsymbol{v}_i^*$, and

$$(\pi(g) \circ \boldsymbol{v}_j \cdot \boldsymbol{v}_i^*)(\boldsymbol{v}_k) = \begin{cases} \pi(g)(\boldsymbol{v}_j) & \text{if } i = k, \\ \boldsymbol{0} & \text{if } i \neq k, \end{cases}$$

and therefore, we find that also

$$(\mu' \circ \alpha)(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)(g) = \mathrm{tr}(\pi(g) \circ \boldsymbol{v}_j \cdot \boldsymbol{v}_i^*) = a_{ij}(g),$$

which shows that indeed $\mu = \mu' \circ \alpha$ as stated. Finally, the map α is an isomorphism, since it maps the basis $(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)_{1 \le i,j \le n}$ of the k-vector space $V \otimes V^*$ to the basis $(\boldsymbol{v}_j \cdot \boldsymbol{v}_i^*)_{1 \le i,j \le n}$ of the k-vector space $\mathrm{End}_k(V)$. □

Remark 7.7. In the situation of Lemma 7.6, the maps $\mu = \mu_\pi$ and $\mu' = \mu'_\pi$ depend on the k-linear representation (V, π), where as the map $\alpha = \alpha_V$ only depends on the k-vector space V.

Definition 7.8. Let k be a field, and let G be a group. If (V, π) is a finite dimensional k-linear representation of G, then its character

$$\chi_\pi \in k[G]$$

is the function defined by $\chi_\pi(g) = \mathrm{tr}(\pi(g))$.

We note that χ_π belongs to the subspace of matrix coefficients. More precisely,

$$\chi_\pi = \mu'_\pi(\mathrm{id}_V) \in M(\pi) \subset k[G].$$

The main result of this lecture is that for G finite and k an algebraically closed field of characteristic zero, the character χ_π determines π, up to non-canonical isomorphism. We first record some properties of the character.

Proposition 7.9. *Let k be a field and let G be a group. The character of finite dimensional k-linear representations of G has the following properties.*

(1) *If $\pi_1 \simeq \pi_2$, then $\chi_{\pi_1} = \chi_{\pi_2}$.*
(2) *The character of the dual of a representation is given by*

$$\chi_{\pi^*}(g) = \chi_\pi(g^{-1}).$$

(3) *The character of a sum of representations is given by*

$$\chi_{\pi_1 \oplus \pi_2} = \chi_{\pi_1} + \chi_{\pi_2}.$$

(4) *The character of a tensor product of representations is given by*

$$\chi_{\pi_1 \otimes \pi_2} = \chi_{\pi_1} \cdot \chi_{\pi_2}.$$

(5) *For all $g, h \in G$, $\chi_\pi(ghg^{-1}) = \chi_\pi(h)$.*

Proof. (1) That $\pi_1 \simeq \pi_2$ means that there exists a k-linear isomorphism $h \colon V_1 \to V_2$ such that $\pi_2(g) = h \circ \pi_1(g) \circ h^{-1}$ for all $g \in G$. But then

$$\chi_{\pi_2}(g) = \mathrm{tr}(\pi_2(g)) = \mathrm{tr}(h \circ \pi_1(g) \circ h^{-1}) = \mathrm{tr}(\pi_1(g)) = \chi_{\pi_1}(g).$$

(2) By definition, $\pi^*(g) = \pi(g^{-1})^*$, so

$$\chi_{\pi^*}(g) = \mathrm{tr}(\pi^*(g)) = \mathrm{tr}(\pi(g^{-1})^*) = \mathrm{tr}(\pi(g^{-1})) = \chi_\pi(g^{-1}).$$

(3) This follows immediately from the fact that $\mathrm{tr}(f_1 \oplus f_2) = \mathrm{tr}(f_1) + \mathrm{tr}(f_2)$.
(4) Since $\mathrm{tr}(f_1 \otimes f_2) = \mathrm{tr}(f_1) \cdot \mathrm{tr}(f_2)$, we have, more generally, that

$$\chi_{\pi_1 \boxtimes \pi_2}(g_1, g_2) = \chi_{\pi_1}(g_1) \cdot \chi_{\pi_2}(g_2),$$

so restricting along $\Delta \colon G \to G \times G$, the stated formula follows.
(5) This follows from the fact that $\mathrm{tr}(f_1 \circ f_2) = \mathrm{tr}(f_2 \circ f_1)$. □

Definition 7.10. Let k be a field, and let G be a finite group. A function $f \colon G \to k$ is central if $f(ghg^{-1}) = f(h)$ for all $g, h \in G$.

It is clear that the subset of $k[G]$ that consists of the central functions is a k-linear subspace. We denote this subspace by[24]

$$Z(k[G]) \subset k[G].$$

The explanation for this notation is as follows. The k-vector space $k[G]$ becomes a k-algebra under the convolution product $*$ defined by

$$(f_1 * f_2)(g) = \textstyle\sum_{h_1 h_2 = g} f_1(h_1) f_2(h_2),$$

[24] The book write $k[G]^\#$ for this subspace.

and moreover, the map $\mu'_\pi\colon \operatorname{End}_k(V) \to k[G]$ is a k-algebra homomorphism with respect to the composition product on $\operatorname{End}_k(V)$ and the convolution product on $k[G]$ in the sense that μ'_π is k-linear and

$$\mu'_\pi(f_1 \circ f_2) = \mu'_\pi(f_1) * \mu'_\pi(f_2)$$

for all $f_1, f_2 \in \operatorname{End}_k(V)$. Now, the subspace $Z(k[G]) \subset k[G]$ is the center of the k-algebra $(k[G], +, *)$ in the sense that $f \in Z(k[G])$ if and only if

$$f * h = h * f$$

for all $h \in k[G]$.

Lemma 7.11. *Let G be a finite group, and let k be an algebraically closed field of characteristic zero. If π is an irreducible k-linear representation of G, then*

$$Z(k[G]) \cap M(\pi) = \operatorname{span}_k(\chi_\pi).$$

Proof. We use that by Lemma 7.6, we have

$$M(\pi) = \{\mu'_\pi(f) \mid f \in \operatorname{End}_k(V)\} \subset k[G],$$

where $\mu'_\pi(f)(g) = \operatorname{tr}(\pi(g) \circ f)$. Since π is irreducible, the map μ_π, and hence, also the map μ'_π is injective, so the induced map

$$\operatorname{End}_k(V) \xrightarrow{\mu'_\pi} M(\pi)$$

is an isomorphism. Now, we calculate that

$$\begin{aligned}\mu'_\pi(f)(ghg^{-1}) &= \operatorname{tr}(\pi(ghg^{-1}) \circ f)\\ &= \operatorname{tr}(\pi(g) \circ \pi(h) \circ \pi(g)^{-1} \circ f)\\ &= \operatorname{tr}(\pi(h) \circ \pi(g)^{-1} \circ f \circ \pi(g))\\ &= \mu'_\pi(\pi(g)^{-1} \circ f \circ \pi(g))(h),\end{aligned}$$

which shows that the function $\mu'_\pi(f)$ is central if and only if for all $g \in G$, we have

$$f = \pi(g)^{-1} \circ f \circ \pi(g).$$

So $\mu'_\pi(f)$ is central if and only if $f\colon V \to V$ is π-invariant. By Schur's lemma, $f\colon V \to V$ is π-invariant if and only if $f = c \cdot \operatorname{id}_V$ for some $c \in k$. But

$$\mu'_\pi(c \cdot \operatorname{id}_V) = c \cdot \mu'_\pi(\operatorname{id}_V) = c \cdot \chi_\pi,$$

which proves the lemma. □

Theorem 7.12. *Let k be an algebraically closed field of characteristic zero, let G be a finite group, and let $\pi_1, \dots, \pi_q$ be representatives of the isomorphism classes of irreducible k-linear representations of G. In this situation, the family*

$$(\chi_{\pi_1}, \dots, \chi_{\pi_q})$$

of their characters is a basis of the k-vector space $Z(k[G])$.

Proof. This follows immediately form Theorem 7.3 and Lemma 7.11. □

The promised main result concerning characters is the following corollary.

Corollary 7.13. *In the situation of Theorem 7.12, the following hold:*

(1) *The dimension of the k-vector space $Z(k[G])$ is equal to the number of isomorphism classes of irreducible k-linear representations of G, which, in turn, is equal to the number of conjugacy classes of elements of G.*
(2) *The isomorphism class of any finite dimensional k-linear representation π of G (not necessarily irreducible) is determined by its character χ_π.*

Proof. (1) The fact that the dimension of the k-vector space $Z(k[G])$ is equal to the number of isomorphism classes of irreducible k-linear representations follows immediately from Theorem 7.12. But $Z(k[G])$ is defined to be the k-vector spaces of central functions $f\colon G \to k$, and a function $f\colon G \to k$ is central if and only if it factors through the canonical projection $p\colon G \to G\backslash G^{\mathrm{ad}}$ onto the set of orbits for action by G on itself by conjugation. Hence, the dimension of $Z(k[G])$ is also equal to the cardinality of $G\backslash G^{\mathrm{ad}}$.
(2) Since π is semisimple, we have $\pi \simeq \pi_1^{m_1} \oplus \cdots \oplus \pi_q^{m_q}$, so by Proposition 7.9,

$$\chi_\pi = m_1\chi_{\pi_1} + \cdots + m_q\chi_{\pi_q}.$$

But k has characteristic zero, so the unique ring homomorphism $\mathbb{Z} \to k$ is injective, and therefore, this identity in $k[G]$ determines the integers $m_1, \ldots, m_q$.[25] □

Let us use this result to determine the isomorphism classes of irreducible complex representations of the symmetric group $G = \Sigma_4$, which has order $|G| = 24$. We recall that the cycle-type of a permutation of n letters is the partition of n obtained from counting the number of elements in cycles.

Lemma 7.14. *The map that to a permutation $\sigma \in \Sigma_n$ assigns its cycle type induces a bijection of the set of conjugacy classes of elements in Σ_n onto the set of cycle-types of permutations of n letters.*

Proof. Indeed, a conjugation of a permutation corresponds to a relabelling of the elements in $\{1, 2, \ldots, n\}$. □

For $n = 4$, there are five cycle-types, namely

$$1+1+1+1,\ 2+1+1,\ 2+2,\ 3+1, \text{ and } 4,$$

and the permutations

$$e,\ (12),\ (12)(34),\ (123), \text{ and } (1234)$$

represent the corresponding conjugacy classes of elements in $G = \Sigma_4$. Hence,

$$\dim_{\mathbb{C}}(Z(\mathbb{C}[G])) = 5,$$

and there are five isomorphism classes of irreducible complex representations of G. We know three of these already, namely,

(i) the 1-dimensional trivial representation $\pi_1 = \tau$,
(ii) the 1-dimensional sign representation $\pi_2 = \mathrm{sgn}$, and

[25] If instead the characteristic of k were a prime number ℓ, then the identity in $k[G]$ would only determine the congruence classes of the integers $m_1, \ldots, m_q$ modulo ℓ.

(iii) the 3-dimensional representation π_3 of G on

$$V_3 = \{\boldsymbol{x} \in \mathbb{C}^4 \mid z_1 + z_2 + z_3 + z_4 = 0\} \subset \mathbb{C}^4$$

defined by

$$\pi_3(\sigma)(\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}) = \begin{pmatrix} z_{\sigma(1)} \\ z_{\sigma(2)} \\ z_{\sigma(3)} \\ z_{\sigma(4)} \end{pmatrix}.$$

By Addendum 7.4, the dimensions n_4 and n_5 of the remaining two irreducible complex representations π_4 and π_5 satisfy

$$24 = 1^2 + 1^2 + 3^2 + n_4^2 + n_5^2,$$

which implies that $n_4 = 3$ and $n_5 = 2$. We claim that

$$\pi_4 \simeq \pi_2 \otimes \pi_3.$$

To prove this, we must show that $\pi_2 \otimes \pi_3$ is irreducible and not isomorphic to π_3. Now, the representation $\pi_2 \otimes \pi_3$ is irreducible, because

$$\operatorname{sgn} \otimes \pi_2 \otimes \pi_3 = \operatorname{sgn} \otimes \operatorname{sgn} \otimes \pi_3 \simeq \pi_3,$$

and because π_3 is irreducible, and to show that $\pi_2 \otimes \pi_3$ is not isomorphic to π_3, it suffices by Corollary 7.13 to show that

$$\chi_{\pi_2 \otimes \pi_3} = \chi_{\pi_2} \cdot \chi_{\pi_3} = \operatorname{sgn} \cdot \chi_{\pi_3} \neq \chi_{\pi_3}.$$

To prove this, it will suffice to find $\sigma \in G$ such that $\operatorname{sgn}(\sigma) = -1$ and $\chi_{\pi_3}(\sigma) \neq 0$. To this end, we consider $\pi = \pi_1 \oplus \pi_3$, which has

$$\chi_\pi = \chi_{\pi_1} + \chi_{\pi_3} = 1 + \chi_{\pi_3}.$$

But π is isomorphic to the standard permutation representation of G on $\mathbb{C}^4$, so the matrix that represents $\pi((12))$ with respect to the standard basis of $\mathbb{C}^4$ is

$$A((12)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

so $\chi_\pi((12)) = 2$, which shows that $\chi_{\pi_3}((12)) = 1 \neq 0$ as desired. This proves the claim that $\pi_4 \simeq \pi_2 \otimes \pi_3$.

What about the remaining 2-dimensional irreducible complex representation π_5? In general, for every set X, there is a group homomorphism

$$\operatorname{Aut}(X) \xrightarrow{\iota_X} \operatorname{Aut}(\mathcal{P}(X))$$

from the group of permutations of the set X to the group of permutation of its power set $\mathcal{P}(X)$ defined by

$$\iota_X(\sigma)(U) = \{\sigma(x) \in X \mid x \in U\} \subset X.$$

Hence, given a (left) action $\rho\colon G \to \operatorname{Aut}(X)$ by a group G on a set X, we get the induced action $\iota_X \circ \rho\colon G \to \operatorname{Aut}(\mathcal{P}(X))$ of G on $\mathcal{P}(X)$. We let $X = \{1,2,3,4\}$, and let $\rho\colon G \to \operatorname{Aut}(X)$ be the identity map and consider the action

$$G \xrightarrow{\iota_{\mathcal{P}(X)} \circ \iota_X \circ \rho} \operatorname{Aut}(\mathcal{P}(\mathcal{P}(X)))$$

on the iterated power set $\mathcal{P}(\mathcal{P}(X))$. It leaves the subset

$$Y = \{\{\{1,2\},\{3,4\}\},\{\{1,3\},\{2,4\}\},\{\{1,4\},\{2,3\}\}\} \subset \mathcal{P}(\mathcal{P}(X))$$

with three elements invariant, so we obtain a group homomorphism

$$G \xrightarrow{p} \operatorname{Aut}(Y) \simeq \Sigma_3.$$

Clearly, the kernel of p is the (necessarily normal) subgroup

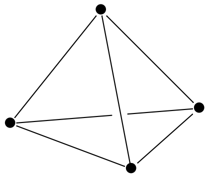
$$N = \{e, (12)(34), (13)(24), (14)(23)\} \subset G,$$

so comparing orders, we conclude that p is surjective. Hence, the 2-dimensional irreducible complex representation (V, π) of Σ_3 defines the 2-dimensional irreducible complex representation $(V, \pi \circ p)$ of G, and this is π_5.

Remark 7.15. Geometrically, we can picture the group homomorphism

$$G \xrightarrow{p} \operatorname{Aut}(Y)$$

above as follows. We may view G as the group of symmetries of the tetrahedron



It has 6 edges, and hence, 3 pairs of an edge and its opposing edge. Now the map p takes a permutation of the 4 vertices to the induced permutation of these 3 pairs of opposing edges.

8. Transitive group actions. Schur orthogonality

Before we get to Schur orthogonality, we will finish some left-over business from last week. We recall that a left G-set is defined to be a pair (X, ρ) of a set X and a group homomorphism $\rho\colon G \to \mathrm{Aut}(X)$, and we say that ρ is a left action by G on X. As is common, we will abbreviate and write $g \cdot x$ or simply gx for $\rho(g)(x)$, where $g \in G$ and $x \in X$. We define the isotropy subgroup (or stabilizer) at $x \in X$ for the left action by G on X to be the subgroup

$$G_x = \{g \in G \mid gx = x\} \subset G,$$

and we define the orbit through $x \in X$ of the left action by G on X to be the subset

$$G \cdot x = \{gx \in X \mid g \in G\} \subset X.$$

Moreover, there is a well-defined bijection

$$G/G_x \xrightarrow{p_x} G \cdot x$$

from the set of left cosets of the isotropy subgroup $G_x \subset G$, which is typically not normal, and onto the orbit $G \cdot x \subset X$ defined by $p_x(hG_x) = hx$. The map p_x is equivariant with respect to the action of G on G/G_x by left multiplication and by the action of G on $G \cdot x \subset X$ obtained by restriction of the action by G on X. Indeed, given $hG_x \in G/G_x$ and $g \in G$, we find that

$$p_x(g \cdot hG_x) = p_x(ghG_x) = gh \cdot x = g \cdot hx = g \cdot p_x(hG_x)$$

as required. The orbits of the action by G on X are the equivalence classes of the equivalence relation $R \subset X \times X$ defined by the image of the map

$$G \times X \xrightarrow{(\mu, p)} X \times X$$

where $\mu\colon G \times X \to X$ is given by $\mu(g, x) = \rho(g)(x)$, and where $p\colon G \times X \to X$ is the canonical projection. We write

$$G\backslash X = \{G \cdot x \in \mathcal{P}(X) \mid x \in X\}$$

for the set of orbits. If there is only one orbit, in which case $G\backslash X = \{X\}$, then we say that the action by G on X is transitive. Equivalently, the action by G on X is transitive if for all $x, y \in X$, there exists $g \in G$ such that $y = gx$.

If two elements $x, y \in X$ belong to the same orbit for the left action by G on X, then their isotropy subgroups $G_x, G_y \subset G$ are conjugate, albeit not canonically so. Indeed, if we choose $g \in G$ such that $y = gx$, then the map

$$G_x \xrightarrow{c_g} G_y$$

defined by $c_g(h) = ghg^{-1}$ is a group isomorphism. We remark that this isomorphism depends on the choice of $g \in G$ with $y = gx$.

If $H \subset G$ is a subgroup, then we define the subset of H-fixed points for the left action by G on X to be the subset

$$X^H = \{x \in X \mid hx = x \text{ for all } h \in H\} \subset X.$$

It is generally not a G-invariant subset, so the action by G on X does generally not restrict to an action by G on X^H. However, we claim that the action by G on X

restricts to an action by the normalizer subgroup

$$N_G(H) = \{g \in G \mid gHg^{-1} \subset H\} \subset G$$

on X^H. Indeed, if $g \in N_G(H)$, then for all $h \in H$, there exists some $h' \in H$ such that $hg = gh'$, and therefore, if $x \in X^H$, then $hgx = gh'x = gx$, which shows that also $gx \in X^H$. Let $\rho_H \colon N_G(H) \to \mathrm{Aut}(X^H)$ denote this action. By definition, this group homomorphism maps every element of $H \subset N_G(H)$ to the identity map of X^H, so it factors (uniquely) as the composition

$$\begin{array}{ccc} N_G(H) & \xrightarrow{\rho_H} & \mathrm{Aut}(X^H) \\ & {\scriptstyle p_H} \searrow \quad \nearrow {\scriptstyle \bar{\rho}_H} & \\ & W_G(H) & \end{array}$$

of the canonical projection p_H of $N_G(H)$ onto the quotient

$$W_G(H) = N_G(H)/H$$

and a left action $\bar{\rho}_H$ of $W_G(H)$ on X^H. The group $W_G(H)$ is called the Weyl group of H in G.

Example 8.1. 1) The group $G = O(3)$ of orthogonal 3×3-matrices acts on

$$S^2 = \{\boldsymbol{x} \in \mathbb{R}^3 \mid \|\boldsymbol{x}\| = 1\} \subset \mathbb{R}^3$$

by left multiplication. The action is transitive, and the "North Pole"

$$\boldsymbol{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in S^2$$

has isotropy subgroup

$$G_{\boldsymbol{x}} = \left\{ \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \in O(3) \mid Q \in O(2) \right\} \subset G.$$

Hence, writing $G_{\boldsymbol{x}} = O(2)$, by abuse of notation, we have a canonical bijection

$$O(3)/O(2) \xrightarrow{p_{\boldsymbol{x}}} S^2.$$

This bijection is in fact a homeomorphism.

(2) Let k be a field. The action by $\mathrm{GL}_m(k)$ on $M_{m,n}(k)$ be left multiplication is not transitive, except in trivial cases. The theorem in linear algebra, which we call Gauss elimination, states that the map

$$\{A \in M_{m,n}(k) \mid A \text{ is on reduced echelon form}\} \longrightarrow \mathrm{GL}_m(k)\backslash M_{m,n}(k)$$

that to A assigns the orbit $\mathrm{GL}_m(k) \cdot A$ is a bijection. Indeed, we learn in linear algebra that two matrices $B, C \in M_{m,n}(k)$ belong to the same orbit for the action by $\mathrm{GL}_m(k)$ on $M_{m,n}(k)$ if and only if B can be transformed to C by means of row operations, and that every orbit for the action by $\mathrm{GL}_m(k)$ on $M_{m,n}(k)$ contains exactly one matrix on reduced echelon form.[26]

[26] Challenge problem: Let $A \in M_{m,n}(k)$ be a matrix on reduced echelon form. Determine the isotropy subgroup $\mathrm{GL}_m(k)_A \subset \mathrm{GL}_m(k)$.

Let G be a group. If $H \subset G$ is a subgroup, then the action by G on G/H by left multiplication is transitive, and conversely, if a left action by G on a set X is transitive, then we have the G-equivariant bijection $p_x \colon G/G_x \to X$, once we choose an element $x \in X$. So every transitive left G-set is non-canonically isomorphic to G/H for some subgroup $H \subset G$. Let k be a field. By analogy with the two-sided regular representation of $G \times G$ on $k[G]$, we have the k-linear representation

$$W_G(H) \times G \xrightarrow{\bar{\rho}} \mathrm{GL}(k[G/H])$$

defined by

$$\bar{\rho}(g_1H, g_2)(\bar{f})(gH) = \bar{f}(g_2^{-1}gg_1H).$$

Moreover, from the general discussion above, we find that the two-sided regular representation of $G \times G$ on $k[G]$ restricts to a k-linear representation

$$W_G(H) \times G \xrightarrow{\rho} \mathrm{GL}(k[G]^{H\times\{e\}}),$$

where we use that the canonical projection

$$W_{G\times G}(H \times \{e\}) \longrightarrow W_G(H) \times G$$

is an isomorphism of groups. This representation is given by

$$\rho(g_1H, g_2)(f)(g) = f(g_2^{-1}gg_1).$$

In this situation, we have the following result.

Lemma 8.2. *Let $p \colon G \to G/H$ be the canonical projection. The map*

$$k[G/H] \xrightarrow{p^*} k[G]^{H\times\{e\}}$$

defined by $p^*(f) = f \circ p$ is a k-linear isomorphism that is intertwining with respect to $\bar{\rho}$ and ρ.

Proof. The right-hand side is the set consisting of the functions $f \colon G \to k$ such that $f(gh) = f(g)$ for all $g \in G$ and $h \in H$. But for every such function, there exists a unique function $\bar{f} \colon G/H \to k$ such that $f = \bar{f} \circ p$. So the map p^* is a bijection, and it is clear that it is k-linear and intertwining with respect to $\bar{\rho}$ and ρ. □

If $H \subset G$ is a subgroup of a group G, and if (V, π) is a k-linear representation of G, then we write (V^H, π^H) for the k-linear representation of $W_G(H)$ on V^H, where

$$W_G(H) \xrightarrow{\pi^H} \mathrm{GL}(V^H)$$

is given by $\pi^H(gH)(\boldsymbol{x}) = \pi(g)(\boldsymbol{x})$.

Theorem 8.3. *Let G be a finite group, let $H \subset G$ be a subgroup, and let k be an algebraically closed field of characteristic zero. Let $(V_1, \pi_1), \dots, (V_q, \pi_q)$ be representatives of the isomorphism classes of the irreducible k-linear representations of G. In this situation, the isomorphism*

$$\pi_1 \boxtimes \pi_1^* \oplus \cdots \oplus \pi_q \boxtimes \pi_q^* \xrightarrow{\mu} \mathrm{Reg}$$

of k-linear representations of $G \times G$ restricts to an isomorphism

$$\pi_1^H \boxtimes \pi_1^* \oplus \cdots \oplus \pi_q^H \boxtimes \pi_q^* \xrightarrow{\mu^{H\times\{e\}}} \mathrm{Reg}^{H\times\{e\}}$$

of k-linear representations of $W_G(H) \times G$.

Proof. In general, if Γ is a group and $K \subset \Gamma$ is a subgroup, then an isomorphism of k-linear representations of Γ induces an isomorphism of the k-linear representations of $W_\Gamma(K)$ obtained by taking K-fixed points. We apply this to $\Gamma = G \times G$ and $K = H \times \{e\}$. □

Let G be a finite group, and let (X, ρ) be a transitive left G-set. We will use Theorem 8.3 to determine the structure of the left regular representation

$$G \xrightarrow{L} k[X]$$

which, we recall, is given by $L(g)(f)(x) = f(\rho(g)^{-1}(x))$.

Corollary 8.4. *Let G be a finite group, let (X, ρ) be a transitive left G-set, and let $H = G_x \subset G$ be the isotropy subgroup of an element $x \in X$. Let k be an algebraically closed field of characteristic zero, let $(V_1, \pi_1), \dots, (V_q, \pi_q)$ be representatives of the isomorphism classes of irreducible k-linear representations of G, and let $m_i = \dim_k(V_i^H)$. In this situation, there exists a non-canonical isomorphism*

$$L \simeq \bigoplus_{i=1}^q \pi_i^{m_i}$$

with $L \colon G \to \mathrm{GL}(k[X])$ the left regular representation of G on $k[X]$.

Proof. The map $p_x \colon G/H \to X$ defined by $p_x(gH) = gx$ is an isomorphism of left G-sets. Moreover, Lemma 8.2 and Theorem 8.3 give k-linear isomorphisms

$$k[G/H] \xrightarrow{p^*} k[G]^{H \times \{e\}} \xleftarrow{\mu} \bigoplus_{i=1}^q V_i^H \times V_i^*,$$

which are intertwining with respect to the respective representations of the group $W_G(H) \times G$ on these k-vector spaces. In particular, they are also intertwining with respect to the subgroup $G = \{H\} \times G \subset W_G(H) \times G$. Therefore, we conclude that the representation π_i^* appears with multiplicity $\dim_k(\pi_i^H)$ in L. But the dual representations $\pi_1^*, \dots, \pi_q^*$ also form a set of representatives of the isomorphism classes of irreducible k-linear representations of G, so we may equivalently conclude that the representation π_i appears with multiplicity $\dim_k((\pi_i^*)^H)$ in L. Thus, it remains to prove that $\dim_k((\pi_i^*)^H)$ and $\dim_k(\pi_i^H)$ are equal.

More generally, for every finite dimensional k-linear representation (V, π) of G, we will prove that $\dim_k((V^*)^H) = \dim_k(V^H)$. The composition

$$H \longrightarrow G \xrightarrow{\pi} \mathrm{GL}(V)$$

of the canonical inclusion and the representation π is a finite dimensional k-linear representation of H, and hence, it decomposes as a sum

$$\pi \simeq \rho_1^{m_1} \oplus \cdots \oplus \rho_r^{m_r}$$

of irreducible k-linear representations of H. It follows that

$$\pi^* \simeq (\rho_1^*)^{m_1} \oplus \cdots \oplus (\rho_r^*)^{m_r}.$$

Exactly one of $\rho_1, \dots, \rho_r$ is a trivial (1-dimensional) representation of H, and exactly one of $\rho_1^*, \dots, \rho_r^*$ is a trivial (1-dimensional) representation of H. Moreover, ρ_i is trivial if and only if ρ_i^* is trivial. Reordering, if necessary, we can assume that ρ_1 and ρ_1^* are trivial. But then

$$\pi^H \simeq \pi_1^{m_1} \simeq (\pi_1^*)_1^m \simeq (\pi^*)^H,$$

so their dimensions agree, as we wanted to show. ☐

Remark 8.5. In addition to the choice of an element $x \in X$, the isomorphism in Corollary 8.4 depends on a choice of basis of V_i^H for all $1 \leq i \leq q$, and therefore, it is non-canonical.

Schur orthogonality

We now let $k = \mathbb{C}$ be the complex numbers and continue to let G be a finite group. Given a finite dimensional complex representation of G, we have defined the associated subspace of matrix coefficients

$$M(\pi) \subset \mathbb{C}[G]$$

to be the common image of the maps μ_π and μ'_π in the diagram

$$\begin{array}{ccc} \operatorname{End}_{\mathbb{C}}(V) & \xrightarrow{\mu'_\pi} & \mathbb{C}[G] \\ {\scriptstyle \alpha_V}\nwarrow{\scriptstyle \sim} & & \nearrow{\scriptstyle \mu_\pi} \\ & V \otimes V^* & \end{array}$$

which are defined by $\mu_\pi(\boldsymbol{x} \otimes \varphi) = \varphi(\pi(g)(\boldsymbol{x}))$ and $\mu'_\pi(h)(g) = \operatorname{tr}(\pi(g) \circ h)$. The map α_V in the diagram is defined by $\alpha_V(\boldsymbol{x} \otimes \varphi)(\boldsymbol{y}) = \boldsymbol{x} \cdot \varphi(\boldsymbol{y})$ and is an isomorphism. We also saw that if $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is a basis of V and $(\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*)$ is the dual basis of V^*, then the family $(\mu_\pi(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*))_{1 \leq i,j \leq n}$ always generates $M(\pi)$, and if π is irreducible, then the family is a basis of $M(\pi)$.

Let $\langle -, - \rangle$ be a hermitian inner product on V. I will use the convention that

$$\langle \boldsymbol{x} \cdot z, \boldsymbol{y} \cdot w \rangle = \bar{z} \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot w,$$

which is the opposite of the convention used in the book. It determines and is determined by the $\mathbb{C}$-linear isomorphism

$$\overline{V} \xrightarrow{b} V^*$$

defined by $b(\boldsymbol{x})(\boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$. So we may also identify $M(\pi) \subset \mathbb{C}[G]$ with the image of the composite map

$$V \otimes \overline{V} \xrightarrow{V \otimes b} V \otimes V^* \xrightarrow{\mu_\pi} \mathbb{C}[G],$$

which is given by

$$(\mu \circ (V \otimes b))(\boldsymbol{x} \otimes \boldsymbol{y})(g) = b(\boldsymbol{y})(\pi(g)(\boldsymbol{x})) = \langle \boldsymbol{y}, \pi(g)(\boldsymbol{x}) \rangle.$$

So if $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is a basis of V that is orthonormal with respect to $\langle -, - \rangle$, then the matrix $A(g) = (a_{ij}(g)) \in M_{n,n}(\mathbb{C})$ that represents $\pi(g) \colon V \to V$ with respect to this basis has entries given by

$$a_{ij}(g) = \langle \boldsymbol{v}_i, \pi(g)(\boldsymbol{v}_j) \rangle.$$

The hermitian inner product $\langle -, - \rangle$ on V gives rise to a hermitian inner product $\langle -, - \rangle_{\mathrm{Frob}}$ on $\operatorname{End}_{\mathbb{C}}(V)$ called the Frobenius inner product. To define it, we recall that given a complex linear map $h\colon V \to V$, its adjoint with respect to $\langle -, - \rangle$ is the unique complex linear map $h^*\colon V \to V$ such that

$$\langle h^*(\boldsymbol{x}), \boldsymbol{y} \rangle = \langle \boldsymbol{x}, h(\boldsymbol{y}) \rangle$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$. Equivalently, the adjoint with respect to $\langle -, - \rangle$ is the unique complex linear map $h^*\colon V \to V$ that makes the diagram

$$\begin{array}{ccc} \overline{V} & \xrightarrow{\;b\;} & V^* \\ \big\downarrow{\scriptstyle h^*} & & \big\downarrow{\scriptstyle h^*} \\ \overline{V} & \xrightarrow{\;b\;} & V^* \end{array}$$

commute. Here, the right-hand vertical map is given by $h^*(\varphi)(\boldsymbol{x}) = \varphi(h(\boldsymbol{x}))$. Now, for $h_1, h_2 \in \mathrm{End}_{\mathbb{C}}(V)$, the Frobenius inner product is defined by

$$\langle h_1, h_2 \rangle_{\mathrm{Frob}} = \mathrm{tr}(h_1^* \circ h_2).$$

It is a hermitian inner product, which, we stress, depends on the choice of the hermitian inner product $\langle -, - \rangle$ on V.

Definition 8.6. Let G be a finite group. The Schur inner product on $\mathbb{C}[G]$ is the hermitian inner product given by

$$\langle f_1, f_2 \rangle_{\mathrm{Sch}} = \frac{1}{|G|} \sum_{g \in G} \overline{f_1(g)}\, f_2(g).$$

It is straightforward to verify that the Schur inner product is invariant with respect to the two-sided regular representation of $G \times G$ on $\mathbb{C}[G]$ in the sense that

$$\langle \mathrm{Reg}(g_1, g_2)(f_1), \mathrm{Reg}(g_1, g_2)(f_2) \rangle_{\mathrm{Sch}} = \langle f_1, f_2 \rangle_{\mathrm{Sch}}$$

for all $(g_1, g_2) \in G \times G$ and $f_1, f_2 \in \mathbb{C}[G]$.

Theorem 8.7 (Schur orthogonality)**.** *Let G be a finite group.*

(a) *If π_1 and π_2 are non-isomorphic irreducible complex representations of G, then their subspaces of matrix coefficients*

$$M(\pi_1), M(\pi_2) \subset \mathbb{C}[G]$$

are orthogonal with respect to the Schur inner product.

(b) *If (V, π) is an irreducible complex representation of G that is unitary with respect to an hermitian inner product $\langle -, - \rangle$ on V, then*

$$\langle \mu'_\pi(h_1), \mu'_\pi(h_2) \rangle_{\mathrm{Sch}} = \frac{1}{n} \cdot \langle h_1, h_2 \rangle_{\mathrm{Frob}}$$

for all $h_1, h_2 \in \mathrm{End}_{\mathbb{C}}(V)$, where $n = \dim_{\mathbb{C}}(V)$.

Proof. (a) We wish to prove that the composition

$$M(\pi_1) \xrightarrow{\;i\;} \mathbb{C}[G] \xrightarrow{\;p\;} M(\pi_2)$$

of the canonical inclusion and the orthogonal projection with respect to the Schur inner product $\langle -, - \rangle_{\mathrm{Sch}}$ is the zero map. But the map is intertwining with respect to $\mathrm{Reg}_{M(\pi_1)}$ and $\mathrm{Reg}_{M(\pi_2)}$, and we have proved before that, as complex representations of $G \times G$, $\mathrm{Reg}_{M(\pi_1)}$ and $\mathrm{Reg}_{M(\pi_2)}$ are irreducible and non-isomorphic. So Schur's lemma proves that the map is zero, as desired.

(b) The representation $\pi\colon G \to \mathrm{GL}(V)$ gives rise to a representation

$$G \times G \xrightarrow{\rho} \mathrm{GL}(\mathrm{End}_{\mathbb{C}}(V))$$

defined by $\rho(g_1, g_2)(h) = \pi(g_1) \circ h \circ \pi(g_2)^{-1}$, and we claim that the map

$$\mathrm{End}_{\mathbb{C}}(V) \xrightarrow{\mu'_\pi} \mathbb{C}[G]$$

is intertwining between ρ and Reg. Indeed, we have

$$\begin{aligned}\mu'_\pi(\rho(g_1, g_2)(h))(g) &= \mathrm{tr}(\pi(g) \circ \pi(g_1) \circ h \circ \pi(g_2)^{-1}) \\ &= \mathrm{tr}(\pi(g_2)^{-1} \circ \pi(g) \circ \pi(g_1) \circ h) \\ &= \mathrm{tr}(\pi(g_2^{-1} g g_1) \circ h) \\ &= \mathrm{Reg}(g_1, g_2)(\mu'_\pi(h))(g).\end{aligned}$$

Since π is irreducible, the map μ'_π is injective, and hence, defines an isomorphism

$$\mathrm{End}_{\mathbb{C}}(V) \xrightarrow{\mu'_\pi} M(\pi)$$

that is intertwining between ρ and $\mathrm{Reg}_{M(\pi)}$. Now, we have two hermitian inner products on $\mathrm{End}_{\mathbb{C}}(V)$, namely, the Frobenius inner product $\langle -, - \rangle_{\mathrm{Frob}}$ and, via the isomorphism μ'_π, the Schur inner product $\langle -, - \rangle'_{\mathrm{Sch}}$ defined by

$$\langle h_1, h_2 \rangle'_{\mathrm{Sch}} = \langle \mu'_\pi(h_1), \mu'_\pi(h_2) \rangle,$$

and both are ρ-invariant. But $\rho \simeq \pi \boxtimes \pi^*$ is irreducible, so Theorem 6.12 shows that there exists a positive real number c such that

$$\langle h_1, h_2 \rangle'_{\mathrm{Sch}} = c \cdot \langle h_1, h_2 \rangle_{\mathrm{Frob}}$$

for all $h_1, h_2 \in \mathrm{End}_{\mathbb{C}}(V)$. It remains to determine the constant c.

We choose a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V that is orthonormal with respect to the given hermitian inner product. Since π is unitary with respect to $\langle -, - \rangle$, the matrix

$$A(g) = (a_{ij}(g)) \in M_n(\mathbb{C})$$

that represents $\pi(g)\colon V \to V$ with respect to the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is a unitary matrix. Therefore, we find that

$$\langle a_{ij}, a_{kl} \rangle_{\mathrm{Sch}} = \frac{1}{|G|} \sum_{g \in G} \overline{a_{ij}(g)}\, a_{kl}(g) = \frac{1}{|G|} \sum_{g \in G} a_{ji}(g^{-1})\, a_{kl}(g),$$

where the second identity holds, because the matrix $A(g)$ is unitary. This formula gives us the idea to consider the sum

$$\begin{aligned}&\sum_{1 \le i \le n} \langle \alpha_V(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*), \alpha_V(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*) \rangle'_{\mathrm{Sch}} = \sum_{1 \le i \le n} \langle \mu_\pi(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*), \mu_\pi(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*) \rangle_{\mathrm{Sch}} \\ &= \frac{1}{|G|} \sum_{1 \le i \le n} \sum_{g \in G} a_{ji}(g^{-1}) a_{il}(g) = \frac{1}{|G|} \sum_{g \in G} \sum_{1 \le i \le n} a_{ji}(g^{-1}) a_{il}(g) = \delta_{jl},\end{aligned}$$

where the last identity holds, because $A(g^{-1}) = A(g)^{-1}$. By comparison,

$$\begin{aligned}&\sum_{1 \le i \le n} \langle \alpha_V(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*), \alpha_V(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*) \rangle_{\mathrm{Frob}} = \sum_{1 \le i \le n} \mathrm{tr}(\alpha_V(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)^* \circ \alpha_V(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*)) \\ &= \sum_{1 \le i \le n} \mathrm{tr}(\alpha_V(\boldsymbol{v}_i \otimes \boldsymbol{v}_j^*) \circ \alpha_V(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*)) = \sum_{1 \le i \le n} \delta_{jl} = n \cdot \delta_{jl},\end{aligned}$$

so we find that $c = \frac{1}{n}$, as stated. Let us explain the second and third identity in this calculation. The matrix B that represents $\alpha_V(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)\colon V \to V$ with respect to the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ has only one nonzero entry, namely, $b_{ij} = 1$. Since the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is orthonormal with respect to $\langle -, - \rangle$, the matrix that represents the adjoint map $\alpha_V(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)^*\colon V \to V$ with respect to this basis is the adjoint matrix $C = B^*$, whose only nonzero entry is $c_{ji} = 1$. Similarly, the only nonzero entry in the matrix D that represents $\alpha_V(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*)$ with respect to $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is $d_{il} = 1$. Finally, the matrix that represents $\alpha_V(\boldsymbol{v}_j \otimes \boldsymbol{v}_i^*)^* \circ \alpha_V(\boldsymbol{v}_l \otimes \boldsymbol{v}_i^*)\colon V \to V$ with respect to the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is $F = C \cdot D$, whose only nonzero entry is $f_{jl} = 1$, so its trace is indeed δ_{jl} as stated. □

Corollary 8.8. *Let G be a finite group, and let $Z(\mathbb{C}[G]) \subset \mathbb{C}[G]$ be the sub-$\mathbb{C}$-vector space consisting of the central functions. If $\pi_1, \dots, \pi_q$ are representatives of the irreducible complex representations of G, then the basis of $Z(\mathbb{C}[G])$ given by the family of their characters $(\chi_{\pi_1}, \dots, \chi_{\pi_q})$ is orthonormal with respect to the Schur inner product.*

Proof. Theorem 8.7 (a) is precisely the statement that $(\chi_{\pi_1}, \dots, \chi_{\pi_q})$ is orthogonal with respect to the Schur inner product. Moreover, we have,

$$\langle \chi_{\pi_i}, \chi_{\pi_i} \rangle_{\mathrm{Sch}} = \langle \mu'_{\pi_i}(\mathrm{id}_{V_i}), \mu'_{\pi_i}(\mathrm{id}_{V_i}) \rangle'_{\mathrm{Sch}} = \frac{1}{n_i} \langle \mathrm{id}_{V_i}, \mathrm{id}_{V_i} \rangle_{\mathrm{Frob}},$$

where the second identity is Theorem 8.7 (b), and by definition

$$\langle \mathrm{id}_{V_i}, \mathrm{id}_{V_i} \rangle_{\mathrm{Frob}} = \mathrm{tr}(\mathrm{id}_{V_i}^* \circ \mathrm{id}_{V_i}) = \mathrm{tr}(\mathrm{id}_{V_i} \circ \mathrm{id}_{V_i}) = \mathrm{tr}(\mathrm{id}_{V_i}) = n_i,$$

which shows that $\langle \chi_{\pi_i}, \chi_{\pi_i} \rangle_{\mathrm{Sch}} = 1$, as desired. □

Corollary 8.9. *Let G be a finite group, and let $\pi_1, \dots, \pi_q$ be representatives of the irreducible complex representations of G. If (V, π) is any finite dimensional complex representation of G, then there is a non-canonical isomorphism*

$$\pi \simeq \pi_1^{m_1} \oplus \dots \oplus \pi_q^{m_q},$$

where $m_i = \langle \chi_\pi, \chi_{\pi_i} \rangle_{\mathrm{Sch}}$.

Proof. By Corollary 7.13, it suffices to show that

$$\chi_\pi = m_1 \chi_{\pi_1} + \dots + m_q \chi_{\pi_q}$$

with m_i as stated. But this follows immediately from Corollary 8.8. □

9. Six-functor formalism for $\operatorname{QCoh}([G\backslash X])$

Let (V, π) be a k-linear representation representation of a group G. If $f\colon G' \to G$ is a group homomorphism, then the composite group homomorphism

$$G' \xrightarrow{\ f\ } G \xrightarrow{\ \pi\ } \operatorname{GL}(V)$$

defines a k-linear representation of G' that we write $f^*(\pi)$ and call the *restriction* of π along f. We will show that, given a k-linear representation (V', π') of G', there are two ways to produce a k-linear representation of G. We write $f_!(\pi')$ and $f_*(\pi')$ for these two k-linear representation of G and call them *compact induction* of π' along f and *induction* of π' along f, respectively. However, to define and understand these, it is better to first generalize our notion of k-linear representation. So, in this lecture, I will assume some familiarity with categories, functors, natural transformations, and adjunctions. We have already encountered these in Lecture 5, when we discussed extension/restriction of scalars.

If G is a group, then we define a category BG, whose set of objects is the singleton set $1 = \{0\}$, and whose set of morphisms is

$$\operatorname{Map}(0,0) = G.$$

We define the composition of morphisms in the category BG to be the product of these as elements of the group G, that is,

$$g \circ h = gh,$$

and we define the identity morphism of the unique object 0 in the category BG to be the identity element in the group G, that is,

$$\operatorname{id}_0 = e.$$

Let k be a field, and let Vect_k be the category, whose set of objects is the set of (small, right) k-vector spaces, and whose set of morphisms is the set of k-linear maps between such k-vector spaces. Composition of morphisms is defined to be composition of maps, and the identity morphism of V is defined to be the identity map id_V. Now, a k-linear representation (V, π) of G determines a functor

$$BG \xrightarrow{\ \pi\ } \operatorname{Vect}_k$$

that to the unique object 0 assigns the k-vector space $\pi(0) = V$ and that to the morphism $g\colon 0 \to 0$ assigns the k-linear map $\pi(g)\colon V \to V$. Indeed, it is a functor, since for all morphisms $g, h\colon 0 \to 0$ in BG, we have

$$\pi(g \circ h) = \pi(gh) = \pi(g) \circ \pi(h),$$

and for the unique object 0 in BG, we have

$$\pi(\operatorname{id}_0) = \pi(e) = \operatorname{id}_V = \operatorname{id}_{\pi(0)}.$$

Conversely, a functor $\pi\colon BG \to \operatorname{Vect}_k$ determines a k-linear representation

$$G \xrightarrow{\ \pi\ } \operatorname{GL}(V),$$

where $V = \pi(0)$, and where $\pi(g)\colon V \to V$ is the k-linear map $\pi(g)\colon \pi(0) \to \pi(0)$. This map is invertible. Indeed, every morphism $g\colon 0 \to 0$ in BG is an isomorphism, and every functor takes isomorphisms to isomorphisms, but let us give the proof.

Let $g\colon 0 \to 0$ be a morphism in BG. That $h\colon 0 \to 0$ is an inverse of g means that $g \circ h = \mathrm{id}_0$ and $h \circ g = \mathrm{id}_0$. Since $\pi\colon BG \to \mathrm{Vect}_k$ is a functor, we have

$$\pi(g) \circ \pi(h) = \pi(g \circ h) = \pi(\mathrm{id}_0) = \mathrm{id}_{\pi(0)},$$
$$\pi(h) \circ \pi(g) = \pi(h \circ g) = \pi(\mathrm{id}_0) = \mathrm{id}_{\pi(0)},$$

which shows that $\pi(g) \in \mathrm{GL}(V)$, as claimed.

We generalize this as follows. Let G be a group and recall that a left G-set is defined to be a pair (X, ρ) of a set X and a group homomorphism

$$G \xrightarrow{\ \rho\ } \mathrm{Aut}(X).$$

As we explained in the Lecture 8, we also write $g \cdot x$ or gx instead of $\rho(g)(x)$ and we say that G acts from the left on the set X. Given a left G-set (X, ρ), we define a category called the translation groupoid of (X, ρ) and denoted

$$[G \backslash X]$$

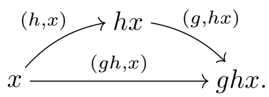
as follows. The set of objects is the set $[G \backslash X]_0 = X$, and the set of morphisms is the set $[G \backslash X]_1 = G \times X$. The source and target maps

$$[G \backslash X]_1 \underset{t}{\overset{s}{\rightrightarrows}} [G \backslash X]_0$$

are given by $s(g, x) = x$ and $t(g, x) = gx$, respectively, and the identity map

$$[G \backslash X]_0 \xrightarrow{\ e\ } [G \backslash X]_1$$

is given by $e(x) = (e, x)$. So, in other words, we view the pair (g, x) as a morphism from x to gx, and we define the identity morphism of x to be the pair (e, x). The composition of $(g, hx)\colon hx \to ghx$ and $(h, x)\colon x \to hx$ is $(gh, x)\colon x \to ghx$:



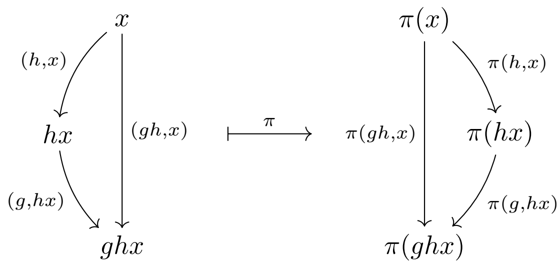
In the case of the trivial action of G on the set $1 = \{0\}$, we recover the category

$$BG = [G \backslash 1].$$

We now define a k-linear representation of $[G \backslash X]$ to be a functor

$$[G \backslash X] \xrightarrow{\ \pi\ } \mathrm{Vect}_k\,.$$

Such a functor assigns k-vector spaces and k-linear maps as indicated below.



The category $[G\backslash X]$ is a simple example of what is called a stack, and a functor $\pi\colon [G\backslash X] \to \mathrm{Vect}_k$ is also called a quasi-coherent sheaf on this stack. We write

$$\mathrm{QCoh}([G\backslash X]) = \mathrm{Fun}([G\backslash X], \mathrm{Vect}_k)$$

for the category, whose objects are the functors $\pi\colon [G\backslash X] \to \mathrm{Vect}_k$, and whose morphisms are natural transformations between such functors. So a morphism

$$\pi \xrightarrow{h} \pi'$$

is a family $(h_x)_{x\in X}$ of k-linear maps

$$\pi(x) \xrightarrow{h_x} \pi'(x)$$

such that for every $(g,x) \in G \times X$, the diagram

$$\begin{array}{ccc} \pi(x) & \xrightarrow{h_x} & \pi'(x) \\ \downarrow{\scriptstyle \pi(g,x)} & & \downarrow{\scriptstyle \pi'(g,x)} \\ \pi(gx) & \xrightarrow{h_{gx}} & \pi'(gx) \end{array}$$

commutes. In particular, the category

$$\mathrm{Rep}_k(G) = \mathrm{QCoh}(BG) = \mathrm{QCoh}([G\backslash 1])$$

is the category of k-linear representations and intertwining k-linear maps.

It happens rarely that categories are equal or even that they are isomorphic. Being equal or being isomorphic are not good notions for categories. (In fact, they are so-called "evil" notions, because they involve equality.) Instead, the notion of equivalence is a good notion. A functor

$$\mathcal{D} \xrightarrow{F} \mathcal{C}$$

is defined to be an equivalence, if there exists a functor

$$\mathcal{C} \xrightarrow{H} \mathcal{D}$$

in the opposite direction and natural transformations

$$F \circ H \xrightarrow{\epsilon} \mathrm{id}_{\mathcal{C}}$$

$$\mathrm{id}_{\mathcal{D}} \xrightarrow{\eta} H \circ F$$

such that for all $c \in \mathcal{C}$ and $d \in \mathcal{D}$, the morphisms

$$(F \circ H)(c) \xrightarrow{\epsilon_c} \mathrm{id}_{\mathcal{C}}(c) = c$$

$$d = \mathrm{id}_{\mathcal{D}}(d) \xrightarrow{\eta_d} (H \circ F)(d)$$

in $\mathcal{C}$ and $\mathcal{D}$, respectively, are isomorphisms. In this situation, we say that ϵ and η are natural isomorphisms, and that H is a quasi-inverse of F. We note, however, that H is *not* uniquely determined by F.

Remark 9.1. If $F\colon \mathcal{D} \to \mathcal{C}$ is an equivalence (of categories), then it is always possible to choose $H\colon \mathcal{C} \to \mathcal{D}$ and $\epsilon\colon F\circ H \to \mathrm{id}_{\mathcal{C}}$ and $\eta\colon \mathrm{id}_{\mathcal{D}} \to H\circ F$ such that the following diagrams of natural transformations commute.

$$
\begin{tikzcd}
 & F\circ H\circ F \arrow[dr, "\epsilon\circ F"] & \\
F \arrow[ur, "F\circ\eta"] \arrow[rr, "\mathrm{id}_F"] & & F
\end{tikzcd}
\qquad
\begin{tikzcd}
 & H\circ F\circ H \arrow[dr, "H\circ\epsilon"] & \\
H \arrow[ur, "\eta\circ H"] \arrow[rr, "\mathrm{id}_H"] & & H
\end{tikzcd}
$$

In this situation, we say that ϵ and η satisfy the triangle identities and that the quadruple (F, G, ϵ, η) is an adjoint equivalence from $\mathcal{D}$ to $\mathcal{C}$.

Proposition 9.2. *Let G be a group, and let (X, ρ) be a transitive left G-set. Let $x \in X$, and let $G_x \subset G$ be the isotropy subgroup. The canonical inclusion functor*

$$BG_x = [G_x\backslash\{x\}] \xrightarrow{\;i\;} [G\backslash X]$$

is an equivalence.

Proof. To produce a quasi-inverse, we choose for all $y \in X$, an element $h_y \in G$ such that $y = h_y x$, and define

$$[G\backslash X] \xrightarrow{\;H\;} [G_x\backslash\{x\}],$$

to be the functor given on objects and morphisms by

$$
\begin{tikzcd}
y \arrow[d, "{(g,y)}"'] \\ gy
\end{tikzcd}
\quad \overset{H}{\longmapsto} \quad
\begin{tikzcd}
x \arrow[d, "{(h_{gy}^{-1} g h_y, x)}"] \\ x
\end{tikzcd}
$$

We further define $\epsilon\colon i\circ H \to \mathrm{id}_{[G\backslash X]}$ and $\eta\colon H\circ i \to \mathrm{id}_{[G_x\backslash\{x\}]}$ by

$$
\begin{tikzcd}
(i\circ H)(y) \arrow[r, "\epsilon_y"] \arrow[d, equal] & y \arrow[d, equal] \\
x \arrow[r, "{(h_y,x)}"] & y
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, "\eta_x"] \arrow[d, equal] & (H\circ i)(x) \arrow[d, equal] \\
x \arrow[r, "{(h_x^{-1},x)}"] & x.
\end{tikzcd}
$$

respectively. The family $\epsilon = (\epsilon_y)_{y\in X}$ is a natural transformation, since the diagram

$$
\begin{tikzcd}
x \arrow[r, "{(h_y,x)}"] \arrow[d, "h_{gy}^{-1} g h_y"] & y \arrow[d, "{(g,y)}"] \\
x \arrow[r, "{(h_{gy},x)}"] & gy
\end{tikzcd}
$$

commutes for all $(g, y) \in G\times X$, and similarly, the family $\eta = (\eta_x)_{x\in\{x\}}$ is a natural transformation since the diagram

$$
\begin{tikzcd}
x \arrow[r, "{(h_x^{-1},x)}"] \arrow[d, "{(g,x)}"] & x \arrow[d, "{(h_x^{-1} g h_x,x)}"] \\
x \arrow[r, "{(h_x^{-1},x)}"] & x
\end{tikzcd}
$$

commutes for all $g \in G_x$. Both ϵ and η are automatically natural isomorphisms, since all morphisms in $[G\backslash X]$ and $[G_x\backslash\{x\}]$ are isomorphisms. $\square$

A category $\mathcal{G}$ is defined to be a groupoid if all morphisms in $\mathcal{G}$ are isomorphisms. The translation groupoid $[G\backslash X]$ of any left G-set (X,ρ) is indeed a groupoid.

Corollary 9.3. *In the situation of Proposition 9.2, the restriction along i,*

$$\mathrm{QCoh}([G\backslash X]) \xrightarrow{i^*} \mathrm{QCoh}([G_x\backslash\{x\}]) = \mathrm{Rep}_k(G_x),$$

which to π assigns $\pi \circ i$, is an equivalence.

Proof. If H is a quasi-inverse of i, then H^* is a quasi-inverse of i^*. □

Remark 9.4. Something better is true, namely, that, as opposed to the equivalence i, the equivalence i^* has a canonical quasi-inverse $i_! \simeq i_*$ given by the left or right Kan extension along i. Explicitly, the functors $i_!$ and i_* are given by

$$\begin{aligned} i_!(\pi)(y) &\simeq \varinjlim(\pi \mid BG_x \times_{[G\backslash X]} [G\backslash X]_{/y}) \\ &\simeq (\textstyle\bigoplus_{(h,x)\colon x\to y} \pi(x))/G_x \\ i_*(\pi)(y) &\simeq \varprojlim(\pi \mid BG_x \times_{[G\backslash X]} [G\backslash X]_{y/}) \\ &\simeq (\textstyle\prod_{(h,y)\colon y\to x} \pi(x))^{G_x}. \end{aligned}$$

It is the possibility of forming sums and products in Vect_k, which we cannot do in $[G\backslash X]$, that makes it possibly to define these functors.

Example 9.5. Let G be a group, and let $H \subset G$ be a subgroup. The pair (X,ρ) consisting of the set $X = G/H$ of left cosets of H in G and the group homomorphism $\rho\colon G \to \mathrm{Aut}(X)$ defined by $\rho(g)(g'H) = gg'H$, is a transitive left G-set. If we use Corollary 9.3 with $x = H = eH \in G/H$, then we find that

$$\mathrm{QCoh}([G\backslash(G/H)]) \xrightarrow{i^*} \mathrm{Rep}_k(H)$$

is an equivalence.

Let (X,ρ) be a left G-set, and let

$$X \xrightarrow{p} G\backslash X$$

be the canonical projection onto the set of orbits. (We remark that

$$G\backslash X \simeq \pi_0([G\backslash X])$$

is the set of isomorphism classes of objects in $[G\backslash X]$.) If we choose an element

$$x = s(\bar{x}) \in \bar{x} = G\cdot x \in G\backslash X$$

in each orbit, then we obtain an isomorphism of left G-sets

$$\textstyle\coprod_{\bar{x}\in G\backslash X} G/G_x \longrightarrow X$$

that to gG_x assigns $g\cdot x$. We note that this isomorphism is highly non-canonical, since it depends on the choice a section $s\colon G\backslash X \to X$ of $p\colon X \to G\backslash X$. Be that as it may, given this choice, we obtain equivalences

$$\textstyle\coprod_{\bar{x}\in G\backslash X} BG_x \longrightarrow \coprod_{\bar{x}\in G\backslash X}[G\backslash(G/G_x)] \longrightarrow [G\backslash X].$$

Finally, taking functors into Vect_k, we obtain the following result.

Proposition 9.6. *Let G be a group, and let (X, ρ) be a left G-set. A choice of representative $x \in \bar{x} \in G \backslash X$ of each orbit determines an equivalence*

$$\mathrm{QCoh}([G \backslash X]) \longrightarrow \textstyle\prod_{\bar{x} \in G \backslash X} \mathrm{Rep}_k(G_x).$$

In Proposition 9.6, the big advantage of the left-hand side is that it only depends on the left G-set (X, ρ), whereas the right-hand side also depends on a choice[27] of section $s\colon G \backslash X \to X$ of the canonical projection $p\colon X \to G \backslash X$. We are now ready to define the compact induction and induction functors.

So let G be a group, and let $f\colon Y \to X$ be a G-equivariant map between left G-sets X and Y. We do not assume that G, X, or Y is finite. It induces a functor

$$[G \backslash Y] \xrightarrow{\;f\;} [G \backslash X],$$

which, by abuse of notation, we again denote by f, and that maps

$$\begin{array}{ccc} y & & f(y) \\ {\scriptstyle (g,y)}\downarrow & \overset{f}{\longmapsto} & \downarrow{\scriptstyle (g,f(y))} \\ gy & & f(gy) \end{array}$$

Since the map $f\colon Y \to X$ is G-equivariant, we have $f(gy) = gf(y)$, so this functor is well-defined. The functor f induces a functor

$$\mathrm{QCoh}([G \backslash X]) \xrightarrow{\;f^*\;} \mathrm{QCoh}([G \backslash Y])$$

that to π assigns $\pi \circ f$ and that we call the restriction along f. It admits both a left adjoint functor $f_!$ and a right adjoint functor f_* given by the left Kan extension along f and the right Kan extension along f, respectively. We call the functor $f_!$ compact induction along f, and we call the functor f_* induction along f. We now spell these out two functions out in detail. First, the functor

$$\mathrm{QCoh}([G \backslash Y]) \xrightarrow{\;f_!\;} \mathrm{QCoh}([G \backslash X])$$

is given by

$$\begin{array}{ccc} f_!(\tau)(x) & \xrightarrow{\;f_!(\tau)(g,x)\;} & f_!(\tau)(gx) \\ \| & & \| \\ \bigoplus_{f(y)=x} \tau(y) & \xrightarrow{\;\bigoplus \tau(g,y)\;} & \bigoplus_{f(y)=x} \tau(gy), \end{array}$$

where the two sums are indexed by

$$f^{-1}(x) = \{y \in Y \mid f(y) = x\},$$

and where we use that, since $f\colon Y \to X$ is G-equivariant, we have

$$\textstyle\bigoplus_{f(y')=gx} \tau(y') = \bigoplus_{f(y)=x} \tau(gy).$$

We define natural transformations $\epsilon = (\epsilon_\pi)$ and $\eta = (\eta_\tau)$ with[28]

$$f_! f^*(\pi) \xrightarrow{\;\epsilon_\pi\;} \pi \qquad\qquad \tau \xrightarrow{\;\eta_\tau\;} f^* f_!(\tau)$$

[27] In general, we need the axiom of choice to even know that it is possible to make this choice!

[28] We abbreviate and write $f_! f^*$ instead of $f_! \circ f^*$, etc.

as follows. The k-linear map

$$\begin{array}{ccc} f_! f^*(\pi)(x) & \xrightarrow{\epsilon_{\pi,x}} & \pi(x) \\ \Vert & & \Vert \\ \bigoplus_{f(y)=x} \pi(f(y)) & \xrightarrow{\nabla} & \pi(x) \end{array}$$

is the fold map (or co-diagonal), whose restriction to each summand is the identity map of $\pi(x) = \pi(f(y))$, and the k-linear map

$$\begin{array}{ccc} \tau(y) & \xrightarrow{\eta_{\tau,y}} & f^* f_!(\tau)(y) \\ \Vert & & \Vert \\ \tau(y) & \xrightarrow{i_y} & \bigoplus_{f(y')=f(y)} \tau(y') \end{array}$$

is the inclusion of the summand indexed by y. One verifies that ϵ and η are indeed well-defined natural transformations and that the triangle identities

$$\begin{array}{ccccc} & & f_! f^* f_! & & \\ & \overset{f_!\eta}{\nearrow} & & \overset{\epsilon f_!}{\searrow} & \\ f_! & & \xrightarrow{\mathrm{id}_{f_!}} & & f_! \end{array} \qquad \begin{array}{ccccc} & & f^* f_! f^* & & \\ & \overset{\eta f^*}{\nearrow} & & \overset{f^*\epsilon}{\searrow} & \\ f^* & & \xrightarrow{\mathrm{id}_{f^*}} & & f^* \end{array}$$

hold. As explained in Lecture 5, this immediately implies:

Theorem 9.7 (Frobenius reciprocity I)**.** *In the situation above, the maps*

$$\mathrm{Map}(f_!(\tau), \pi) \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \mathrm{Map}(\tau, f^*(\pi))$$

defined by $\alpha(h) = f^*(h) \circ \eta_\tau$ and $\beta(k) = \epsilon_\pi \circ f_!(k)$ are each other's inverses.

Similarly, the functor

$$\mathrm{QCoh}([G \backslash Y]) \xrightarrow{f_*} \mathrm{QCoh}([G \backslash X])$$

is given by

$$\begin{array}{ccc} f_*(\tau)(x) & \xrightarrow{f_*(\tau)(g,x)} & f_*(\tau)(gx) \\ \Vert & & \Vert \\ \prod_{f(y)=x} \tau(y) & \xrightarrow{\prod \tau(g,y)} & \prod_{f(y)=x} \tau(gy), \end{array}$$

where the products are indexed by $f^{-1}(x)$ as before. The natural transformations $\epsilon = (\epsilon_\tau)$ and $\eta = (\eta_\pi)$ with

$$f^* f_*(\tau) \xrightarrow{\epsilon_\tau} \tau \qquad\qquad \pi \xrightarrow{\eta_\pi} f_* f^*(\pi)$$

as follows. The k-linear map

$$\begin{array}{ccc} f^* f_*(\tau)(y) & \xrightarrow{\epsilon_{\tau,y}} & \tau(y) \\ \Vert & & \Vert \\ \prod_{f(y')=f(y)} \tau(y') & \xrightarrow{p_y} & \tau(y) \end{array}$$

is the projection on the factor indexed by y, and the k-linear map

$$\begin{array}{ccc} \pi(x) & \xrightarrow{\eta_{\pi,x}} & f_*f^*(\pi)(x) \\ \| & & \| \\ \pi(x) & \xrightarrow{\Delta} & \prod_{f(y)=x} \pi(f(y)) \end{array}$$

is given by the diagonal map. One verifies that ϵ and η are well-defined natural transformations and that they satisfy the triangle identities:

$$f^* \xrightarrow{f^*\eta} f^*f_*f^* \xrightarrow{\epsilon f^*} f^*, \quad \text{composite } = \mathrm{id}_{f^*} \qquad\qquad f_* \xrightarrow{\eta f_*} f_*f^*f_* \xrightarrow{f_*\epsilon} f_*, \quad \text{composite } = \mathrm{id}_{f_*}$$

This gives the following result:

Theorem 9.8 (Frobenius reciprocity II)**.** *In the situation above, the maps*

$$\mathrm{Map}(f^*(\pi),\tau) \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \mathrm{Map}(\pi, f_*(\tau))$$

defined by $\alpha(h) = f_*(h) \circ \eta_\pi$ *and* $\beta(k) = \epsilon_\tau \circ f^*(k)$ *are each other's inverses.*

There is a canonical natural transformation called the norm map

$$f_! \xrightarrow{\mathrm{Nm}_f} f_*.$$

In our description of $f_!$ and f_*, it is given by the canonical inclusion

$$\begin{array}{ccc} f_!(\tau)(x) & \xrightarrow{\mathrm{Nm}_{f,\tau,x}} & f_*(\tau)(x) \\ \| & & \| \\ \bigoplus_{f(y)=x} \tau(y) & \longrightarrow & \prod_{f(y)=x} \tau(y) \end{array}$$

of the sum in the product. We will say that a map $f\colon Y \to X$ is proper, if for all $x \in X$, the inverse image $f^{-1}(x) \subset Y$ is finite.

Theorem 9.9. *If $f\colon Y \to X$ is proper, then the norm map*

$$f_! \xrightarrow{\mathrm{Nm}_f} f_*.$$

is a natural isomorphism.

Proof. Indeed, finite sums and finite products of k-vector spaces agree. $\square$

Finally, we will prove an important theorem called the base-change theorem. A commutative diagram of left G-sets and G-equivariant maps

(9.10)
$$\begin{array}{ccc} Y' & \xrightarrow{h'} & Y \\ \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ X' & \xrightarrow{h} & X \end{array}$$

induces a diagram of categories and functors

$$\begin{array}{ccc} \mathrm{QCoh}([G\backslash Y']) & \xleftarrow{h'^*} & \mathrm{QCoh}([G\backslash Y]) \\ \uparrow{\scriptstyle f'^*} & & \uparrow{\scriptstyle f^*} \\ \mathrm{QCoh}([G\backslash X']) & \xleftarrow{h^*} & \mathrm{QCoh}([G\backslash X]) \end{array}$$

that commutes, up to unique natural isomorphism. The diagram (9.10) is defined to be cartesian, if the map

$$Y' \xrightarrow{(f',h')} X' \times_X Y = \{(x',y) \in X' \times Y \mid h(x') = f(y)\} \subset X' \times Y$$

is a bijection. In this case, also the diagrams

$$\begin{array}{ccc} \mathrm{QCoh}([G\backslash Y']) & \xleftarrow{h'^*} & \mathrm{QCoh}([G\backslash Y]) \\ \downarrow{\scriptstyle f'_* \text{ (resp. } f'_!)} & & \downarrow{\scriptstyle f_* \text{ (resp. } f_!)} \\ \mathrm{QCoh}([G\backslash X']) & \xleftarrow{h^*} & \mathrm{QCoh}([G\backslash X]) \end{array}$$

commute, up to specified natural isomorphisms. Here is a precise statement:

Theorem 9.11 (Base-change)**.** *If a diagram of G-sets and G-equivariant maps as in (9.10) is cartesian, then the following hold.*

(1) *The composite natural transformation*

$$h^*f_* \xrightarrow{\eta h^* f_*} f'_*\, f'^* h^* f_* \simeq f'_*\, h'^* f^* f_* \xrightarrow{f'_* h'^* \epsilon} f'_*\, h'^*$$

is a natural isomorphism.

(2) *The composite natural transformation*

$$f'_!\, h'^* \xrightarrow{f'_! h'^* \eta} f'_!\, h'^* f^* f_! \simeq f'_!\, f'^* h^* f_! \xrightarrow{\epsilon h^* f_!} h^* f_!$$

is a natural isomorphimsm.

Proof. We first remark that (1) and (2) are in fact equivalent statements. Indeed, the natural transformation $h^*f_* \to f'_*h'^*$ in (1), determines and is determined by a natural transformation $h'_!f'^* \to h_!f^*$, which, up to interchanging the role of f and h, precisely is the natural transformation in (2). So it suffices to prove (1). To this end, let $\tau \in \mathrm{QCoh}([G\backslash Y])$, and let $x' \in X'$. On the one hand, we have

$$h^*f_*(\tau)(x') = f_*(\tau)(h(x')) = \textstyle\prod_{f(y)=h(x')} \tau(y),$$

and, on the other hand, we have

$$f'_*\, h'^*(\tau)(x') = \textstyle\prod_{f'(y')=x'} h'^*(\tau)(y') = \prod_{f'(y')=x'} \tau(h'(y')),$$

and since the diagram (9.10) is cartesian, the two products agree. Finally, one checks that the composite map in the statement takes the factor indexed by (y, x') with $f(y) = h'(x')$ to the factor indexed by the unique $y' \in Y'$ such that $f'(y') = x'$ and $h'(y') = y$ by the identity map

$$\tau(y) \xrightarrow{\mathrm{id}} \tau(h'(y')).$$

So it is an isomorphism, which proves (1). □

In the next lecture, we consider the special case, where G is a finite group, where $H, K \subset G$ are two subgroups, and where (9.10) is the cartesian diagram

$$\begin{array}{ccc} G/H \times G/K & \xrightarrow{h'} & G/H \\ \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ G/K & \xrightarrow{h} & G/G. \end{array}$$

Here, f and h are the unique maps (note that $G/G = \{G\}$ only has one element), and h' and f' are the canonical projections. The left G-sets G/H and G/K are both transitive, but $G/H \times G/K$ is not, unless either $H = G$ or $K = G$ or both. Proposition 9.6 gives a product decomposition of $\mathrm{QCoh}([G\backslash(G/H \times G/K)])$, once we fix a choice of representatives of the G-orbits in $G/H \times G/K$. As we will see, this turns out to be rather complicated!

Remark 9.12. The formulas for the left Kan extension $f_!$ and right Kan extension f_* that we have given above are based on the fact that the diagram of anima[29]

$$\begin{array}{ccc} Y & \longrightarrow & [G\backslash Y] \\ \downarrow & & \downarrow \\ X & \longrightarrow & [G\backslash X] \end{array}$$

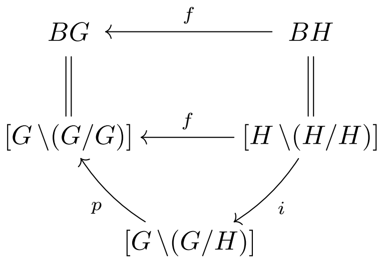
is cartesian. (This follows from Theorem 6.1.0.6 in Lurie's Higher Topos Theory, since anima form an ∞-topos.)

[29] In Lurie's Higher Topos Theory, anima are called "spaces." However, since these are nothing like topological spaces and are in fact discrete in nature, Clausen and Scholze have proposed to use the name anima or animated sets instead.

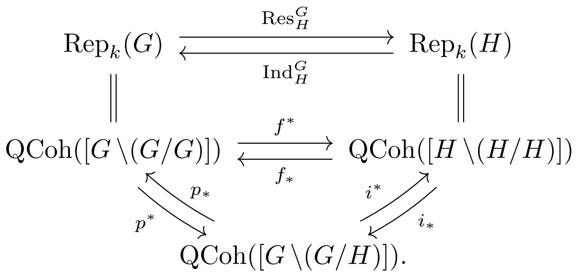
This time, we will apply the general theory that we developed last time to the particular map of left G-sets given by the unique map

$$G/H \xrightarrow{p} G/G = \{G\},$$

where $H \subset G$ is a subgroup.[30] We have defined functors



with i an equivalence of categories, and adjoint pairs of functors



We call $\mathrm{Res}_H^G = f^*$ the restriction from G to H and its right adjoint $\mathrm{Ind}_H^G = f_*$ the induction from H to G. (We also defined a left adjoint $f_!$ of f^*, which we call compact induction from H to G. It is sometimes written ind_H^G in all lower-case.) Since composition of functors is (strictly) associative, we have

$$f^* = (p \circ i)^* = i^* \circ p^*,$$

but it is *not* true that

$$f_* = (p \circ i)_* = p_* \circ i_*.$$

What is true, however, is that the two composite natural transformations

$$f_* \longrightarrow p_*p^*f_* \longrightarrow p_*i_*i^*p^*f_* = p_*i_*f^*f_* \longrightarrow p_*i_*$$

$$p_*i_* \longrightarrow f_*f^*p_*i_* = f_*i^*p^*p_*i_* \longrightarrow f_*i^*i_* \longrightarrow f_*$$

defined using the counits and units of the three adjunctions are each other's inverses. In this way, the two adjoints f_* and p_*i_* of f^* are uniquely naturally isomorphic. This is a general fact:

Proposition 10.1. *Let $(f^*, f_*, \epsilon, \eta)$ and $(f^*, \bar{f}_*, \bar{\epsilon}, \bar{\eta})$ be two adjunctions with the same left adjoint functor f^*. In this situation, the composite natural transformation*

$$f_* \xrightarrow{\bar{\eta} \circ f_*} \bar{f}_* \circ f^* \circ f_* \xrightarrow{\bar{f}_* \circ \epsilon} \bar{f}_*$$

[30] We do *not* assume that $H \subset G$ is normal.

is the unique natural transformation $\sigma\colon f_* \to \bar{f}_*$ that makes the diagrams

$$\begin{array}{ccc} f^* \circ f_* & \xrightarrow{f^*\circ\sigma} & f^* \circ \bar{f}_* \\ & {\scriptstyle\epsilon}\searrow \quad \swarrow{\scriptstyle\bar{\epsilon}} & \\ & \mathrm{id} & \end{array} \qquad \begin{array}{ccc} f_* \circ f^* & \xrightarrow{\sigma\circ f^*} & \bar{f}_* \circ f^* \\ & {\scriptstyle\eta}\nwarrow \quad \nearrow{\scriptstyle\bar{\eta}} & \\ & \mathrm{id} & \end{array}$$

commute. In particular, it is a natural isomorphism with inverse

$$\bar{f}_* \xrightarrow{\eta\circ\bar{f}_*} f_* \circ f^* \circ \bar{f}_* \xrightarrow{f_*\circ\bar{\epsilon}} f_*.$$

Proof. This is not so easy to show. See for example Saunders MacLane, Categories for the Working Mathematician, Chapter IV, Section 7, Theorem 2. □

Here is an application:

Corollary 10.2. *The adjunction*

$$\mathrm{QCoh}([G\backslash(G/H)]) \underset{i_*}{\overset{i^*}{\rightleftarrows}} \mathrm{Rep}_k(H)$$

is an adjoint equivalence.

Proof. In the adjunction $(i^*, i_*, \epsilon, \eta)$, the functors i^* and i_* are given by restriction and right Kan extension along the canonical inclusion

$$BH = [H\backslash(H/H)] \xrightarrow{i} [G\backslash(G/H)],$$

and we wish to prove that ϵ and η are natural isomorphisms. We have proved last time that i is an equivalence of categories. So if h is a quasi-inverse of i, then h^* is a quasi-inverse of i^*, and we can choose natural isomorphisms $\bar{\epsilon}\colon i^* \circ h^* \to \mathrm{id}$ and $\bar{\eta}\colon \mathrm{id} \to h^* \circ i^*$ such that $(i^*, h^*, \bar{\epsilon}, \bar{\eta})$ is an adjunction. By Proposition 10.1, the natural transformation $\sigma\colon i_* \to h^*$ defined as the composition

$$i_* \xrightarrow{\bar{\eta}\circ i_*} i_* i^* h^* \xrightarrow{\epsilon\circ h^*} h^*$$

is an isomorphism and is unique with the property that the diagrams

$$\begin{array}{ccc} i^* \circ i_* & \xrightarrow{i^*\circ\sigma} & i^* \circ h^* \\ & {\scriptstyle\epsilon}\searrow \quad \swarrow{\scriptstyle\bar{\epsilon}} & \\ & \mathrm{id} & \end{array} \qquad \begin{array}{ccc} i_* \circ i^* & \xrightarrow{\sigma\circ i^*} & h^* \circ i^* \\ & {\scriptstyle\eta}\nwarrow \quad \nearrow{\scriptstyle\bar{\eta}} & \\ & \mathrm{id} & \end{array}$$

commute. In particular, we conclude that ϵ and η are natural isomorphisms. □

Proposition 10.1 also implies that to "calculate" the induction functor

$$\mathrm{Rep}_k(H) \xrightarrow{\mathrm{Ind}_H^G} \mathrm{Rep}_k(G),$$

it suffices to produce an adjunction $(\mathrm{Res}_H^G, \mathrm{Ind}_H^G, \epsilon, \eta)$ with $\mathrm{Res}_H^G = f^*$. For in this situation, the proposition will give a unique natural isomorphism $\sigma\colon \mathrm{Ind}_H^G \to f_*$ to any other right adjoint functor f_* of f^*, say, to the right Kan extension along the functor $f\colon BH \to BG$.

Now, given a k-linear representation of H,

$$BH \xrightarrow{\tau} \mathrm{Vect}_k$$

with $W = \tau(0)$, we define the induced k-linear representation

$$BG \xrightarrow{\pi = \mathrm{Ind}_H^G(\tau)} \mathrm{Vect}_k$$

as follows. The k-vector space $\pi(0) = V = \mathrm{Map}_H(G, W)$ is given by the set of all maps $f \colon G \to W$ such that for all $h \in H$ and $x \in G$,

$$f(h \cdot x) = h \cdot f(x) = \tau(h)(f(x)),$$

with vector sum and scalar multiplication by $a \in k$ defined pointwise by

$$\begin{aligned}(f + f')(x) &= f(x) + f'(x) \\ (f \cdot a)(x) &= f(x) \cdot a,\end{aligned}$$

and for $g \in G$, the k-linear map $\pi(g) \colon V \to V$ is given by

$$\pi(g)(f)(x) = f(xg).$$

We define the counit $\epsilon_\tau \colon (\mathrm{Res}_H^G \circ \mathrm{Ind}_H^G)(\tau) \to \tau$ to be the k-linear map

$$\mathrm{Map}_H(G, W) \longrightarrow W$$

that to $f \colon G \to W$ assigns $\epsilon_\tau(f) = f(e)$, and the calculation

$$\epsilon_\tau(h \cdot f) = (h \cdot f)(e) = f(e \cdot h) = f(h \cdot e) = h \cdot f(e) = h \cdot \epsilon_\tau(f)$$

shows that it intertwines between the two k-linear representations of H in question. Finally, given a k-linear representation $\pi \colon BG \to \mathrm{Vect}_k$ of G with $V = \pi(0)$, we define the unit $\eta_\pi \colon \pi \to (\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\pi)$ to be the k-linear map

$$V \longrightarrow \mathrm{Map}_H(G, V)$$

given by $\eta_\pi(\boldsymbol{v})(x) = \pi(x)(\boldsymbol{v})$. The calculation

$$\eta_\pi(\boldsymbol{v})(h \cdot x) = \pi(h \cdot x)(\boldsymbol{v}) = (\pi(h) \circ \pi(x))(\boldsymbol{v}) = \pi(h)(\eta_\pi(\boldsymbol{v}))$$

shows that $\eta_\pi \in \mathrm{Map}_H(G, V)$, so the map is well-defined. And the calculation

$$\eta_\pi(g \cdot \boldsymbol{v})(x) = \pi(x)(\pi(g)(\boldsymbol{v})) = \pi(x \cdot g)(\boldsymbol{v}) = \eta_\pi(\boldsymbol{v})(x \cdot g) = (g \cdot \eta_\pi(\boldsymbol{v}))(x)$$

shows that it intertwines between the two representations in question. Thus, we obtain the following special case of Frobenius reciprocity II, which we proved in Theorem 9.8.

Theorem 10.3 (Frobenius reciprocity II)**.** *In the situation above, the maps*

$$\mathrm{Hom}(\mathrm{Res}_H^G(\pi), \tau) \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \mathrm{Hom}(\pi, \mathrm{Ind}_H^G(\tau))$$

defined by $\alpha(h) = \mathrm{Ind}_H^G(h) \circ \eta_\pi$ *and* $\beta(k) = \epsilon_\tau \circ \mathrm{Res}_H^G(k)$ *are each other's inverses.*

Example 10.4. Let $G = \Sigma_4$ be the group of permutations of the set $\{1, 2, 3, 4\}$, and let $H \subset G$ be the subgroup of permutations σ such that $\sigma(4) = 4$. We identify H with the group Σ_3 of permutations $\{1, 2, 3\}$ via the group isomorphism $\rho \colon H \to \Sigma_3$ defined by $\rho(\sigma) = \sigma|_{\{1,2,3\}}$. We let $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$ be the irreducible complex representations of G defined in Lecture 7, and let τ_1, τ_2, τ_3 be the irreducible complex representations of H defined in Lecture 1. So π_1 and τ_1 are the 1-dimensional trivial representations, π_2 and τ_2 are the 1-dimensional sign representations, π_3 and τ_3 are the standard representations of dimension 3 and 2, respectively, $\pi_4 \simeq \pi_2 \otimes \pi_3$ is 3-dimensional, and π_5 is 2-dimensional. We wish to understand

$$\pi = \operatorname{Ind}_H^G(\tau_1),$$

which has $\dim_{\mathbb{C}}(\pi) = [G : H] \cdot \dim_{\mathbb{C}}(\tau_1) = 4$. We have the canonical isomorphism

$$\bigoplus_{1 \leq i \leq 5} \operatorname{Hom}(\pi_i, \pi) \otimes \pi_i \longrightarrow \pi$$

that to $f_i \otimes \boldsymbol{x}_i$ assigns $f_i(\boldsymbol{x}_i)$, and by Frobenius reciprocity,

$$\operatorname{Hom}(\pi_i, \pi) = \operatorname{Hom}(\pi_i, \operatorname{Ind}_H^G(\tau_1)) \simeq \operatorname{Hom}(\operatorname{Res}_H^G(\pi_i), \tau_1).$$

We see immediately from the definitions that

$$\begin{aligned} \operatorname{Res}_H^G(\pi_1) &\simeq \tau_1 \\ \operatorname{Res}_H^G(\pi_2) &\simeq \tau_2 \\ \operatorname{Res}_H^G(\pi_3) &\simeq \tau_1 \oplus \tau_3, \end{aligned}$$

so by Schur's lemma, we conclude that the canonical map

$$\operatorname{Hom}(\pi_1, \pi) \otimes \pi_1 \oplus \operatorname{Hom}(\pi_3, \pi) \otimes \pi_3 \longrightarrow \pi$$

is an isomorphism. Hence, less canonically, we have an isomorphism

$$\operatorname{Ind}_H^G(\tau_1) \simeq \pi_1 \oplus \pi_3.$$

Let us finish the calculation of $\operatorname{Res}_H^G(\pi_i)$. Using that $\pi_4 = \pi_2 \otimes \pi_3$, we get

$$\begin{aligned} \operatorname{Res}_H^G(\pi_4) = \operatorname{Res}_H^G(\pi_2 \otimes \pi_3) &\simeq \operatorname{Res}_H^G(\pi_2) \otimes \operatorname{Res}_H^G(\pi_3) \\ &\simeq \tau_2 \otimes (\tau_1 \oplus \tau_3) \simeq \tau_2 \oplus \tau_3, \end{aligned}$$

where the second identification uses the "symmetric monoidal" structure on Res_H^G. Finally, we consider the diagram of groups

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N & \longrightarrow & G & \xrightarrow{q} & H & \longrightarrow & 1, \\ & & & & \uparrow{\scriptstyle f} & \mathrel{\rotatebox{45}{$=$}} & & & \\ & & & & H & & & & \end{array}$$

where $N = \{e, (12)(34), (13)(24), (14)(23)\}$, and where q maps $g \in G$ to the unique element $q(g) \in H \cap gN$. In Lecture 7, we defined $\pi_5 = q^*(\tau_3)$, so we find that

$$\operatorname{Res}_H^G(\pi_5) = (f^* \circ q^*)(\tau_3) = (q \circ f)^*(\tau_3) = \tau_3.$$

Remark 10.5. As Example 10.4 shows, if π is irreducible, then $\operatorname{Res}_H^G(\pi)$ may well not be so. (Physicists call this "symmetry breaking.") The example also shows that if τ is irreducible, then $\operatorname{Ind}_H^G(\tau)$ may also not be irreducible.

Suppose $H \subset G$ is a subgroup of finite index $[G : H] = n$. In this case, the map

$$G/H \xrightarrow{p} G/G$$

is proper, so by Theorem 9.9, the norm map $\operatorname{Nm}_p \colon p_! \to p_*$ is a natural isomorphism. This means that, under this assumption, the functor Ind_H^G is also left adjoint to Res_H^G. Let us spell out the adjunction

$$(\operatorname{Ind}_H^G, \operatorname{Res}_H^G, \epsilon', \eta').$$

We choose a family $(g_1, \dots, g_n)$ of representatives of the right cosets $Hg \in H\backslash G$. If (V, π) is a k-linear representation of G, then we define the counit

$$(\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\pi) \xrightarrow{\epsilon'_\pi} \pi$$

to be the k-linear map $\epsilon'_\pi \colon \mathrm{Map}_H(G, V) \to V$ given by $\epsilon'_\pi(f) = \sum_{1 \le i \le n} f(g_i)$, and if (W, τ) is a k-linear representation of H, then we define the unit

$$\tau \xrightarrow{\eta'_\tau} (\mathrm{Res}_H^G \circ \mathrm{Ind}_H^G)(\tau)$$

to be the k-linear map $\eta'_\tau \colon W \to \mathrm{Map}_H(G, W)$ given by

$$\eta'_\tau(\boldsymbol{w})(x) = \begin{cases} \tau(x)(\boldsymbol{w}) & \text{if } x \in H, \\ 0 & \text{if } x \notin H. \end{cases}$$

Therefore, by invoking Proposition 10.1, Frobenius reciprodicity I, which we proved in Theorem 9.7, specializes to the following result.

Theorem 10.6. *Let G be a group, and let $H \subset G$ be a subgroup of finite index. Given k-linear representations π and τ of G and H, respectively, the maps*

$$\mathrm{Hom}(\mathrm{Ind}_H^G(\tau), \pi) \underset{\beta'}{\overset{\alpha'}{\rightleftarrows}} \mathrm{Hom}(\tau, \mathrm{Res}_H^G(\pi))$$

defined by $\alpha'(h) = \mathrm{Res}_H^G(h) \circ \eta'_\tau$ and $\beta'(k) = \epsilon'_\pi \circ \mathrm{Ind}_H^G(k)$ are each other's inverses.

Remark 10.7. The restriction $\mathrm{Res}_H^G = f^*$ always has the left adjoint $\mathrm{ind}_H^K = f_!$, but the norm map $\mathrm{Nm}_f \colon f_! \to f_*$ is a natural isomorphism only if $[G : H] < \infty$.

Let $H, K \subset G$ be two subgroups, and let σ and τ be k-linear representations of H and K, respectively. Frobenius reciprocity gives us the canonical isomorphism

$$\mathrm{Hom}(\mathrm{Ind}_H^G(\sigma), \mathrm{Ind}_K^G(\tau)) \xrightarrow{\beta} \mathrm{Hom}((\mathrm{Res}_K^G \circ \mathrm{Ind}_H^G)(\sigma), \tau),$$

so we would like to understand the functor $\mathrm{Res}_K^G \circ \mathrm{Ind}_H^G$, and this is exactly what the base-change theorem allows us to do. We first determine the set

$$\mathrm{Map}_G(G/H, G/K)$$

of G-equivariant maps $f \colon G/H \to G/K$. Given such a map, we have $f(H) = aK$, for some $a \in G$, and hence, by the G-equivariance of f, we have

$$f(gH) = gaK$$

for all $g \in G$. In particular, we have $haK = aK$ for all $h \in H$, or equivalently,

$$a^{-1}Ha \subset K.$$

Conversely, given $a \in G$ such that $a^{-1}Ha \subset K$, the map $f_a \colon G/H \to G/K$ defined by $f_a(gH) = gaK$ is G-equivariant. Moreover, we observe that $f_a = f_b$ if and only if $aK = bK$, or equivalently, if and only if

$$a^{-1}b \in K.$$

If $a^{-1}Ha = K$, then $f_a = r_a$ is the G-equivariant map

$$G/H \xrightarrow{r_a} G/a^{-1}Ka$$

given by right multiplication by a. Indeed,

$$f_a(gH) = gaa^{-1}Ha = gHa = r_a(gH).$$

In general, if $a^{-1}Ha \subset K$, then f_a factors in two ways

$$\begin{array}{ccc} G/H & \xrightarrow{p_H^{aKa^{-1}}} & G/aKa^{-1} \\ \downarrow{\scriptstyle r_a} & \searrow^{f_a} & \downarrow{\scriptstyle r_a} \\ G/a^{-1}Ha & \xrightarrow{p_{a^{-1}Ha}^{K}} & G/K \end{array}$$

as the composition of r_a and the canonical projections.

We now assume that G is finite and consider the cartesian square of left G-sets

$$\begin{array}{ccc} X & \xrightarrow{p_1} & G/H \\ \downarrow{\scriptstyle p_2} & & \downarrow{\scriptstyle p_H^G} \\ G/K & \xrightarrow{p_K^G} & G/G, \end{array}$$

where $H, K \subset G$ are subgroups and $X = G/H \times G/K$. The base-change theorem, Theorem 9.11, gives a canonical natural isomorphism

$$(p_K^G)^* \circ (p_H^G)_* \longrightarrow p_{2*} \circ p_1^*,$$

so we wish to understand the left G-set X. The map $s\colon G/K \to X$ defined by $s(aK) = (H, aK)$ is not G-equivariant, unless $H = G$, but it induces a surjection

$$G/K \xrightarrow{\bar{s}} G\backslash X = \pi_0([G\backslash X])$$

that maps aK to the G-orbit $\bar{s}(aK) = G \cdot (H, aK)$ through $s(aK) = (H, aK)$, and moreover, (H, aK) and (H, bK) are in the same G-orbit if and only if $ab^{-1} \in H$. This shows that we have a bijection

$$H\backslash G/K \longrightarrow G\backslash X$$

that to HaK assigns the G-orbit $G \cdot (H, aK)$. Moreover, the isotropy subgroup at (H, aK) for the left action by G on X is equal to

$$G_{(H,aK)} = H \cap aKa^{-1},$$

since $(H, aK) = (gH, gaK)$ if and only if $g \in H$ and $g \in aKa^{-1}$. We now choose a map $a\colon \{1, 2, \dots, m\} \to G$, whose composition with the canonical projection

$$\{1, 2, \dots, m\} \xrightarrow{a} G \xrightarrow{q} H\backslash G/K$$

is a bijection. We write $a_s = a(s)$ and say that $(a_1, a_2, \dots, a_m)$ is a family of double coset representatives. With this choice, we obtain a G-equivariant bijection

$$\textstyle\coprod_{1 \le s \le m} G/(H \cap a_sKa_s^{-1}) \xrightarrow{u} X$$

that to $g(H \cap a_sKa_s^{-1})$ assigns (gH, ga_sK). Moreover, we have

$$p_1 \circ u = \textstyle\sum_{1 \le s \le m} p^H_{H \cap a_sKa_s^{-1}}$$

$$p_2 \circ u = \textstyle\sum_{1 \le s \le m} r_{a_s} \circ p^{a_sKa_s^{-1}}_{H \cap a_sKa_s^{-1}},$$

where "Σ" is notation for the map from the disjoint union that on the sth summand is given by the indicated map. Finally, we note that the diagram

$$\begin{array}{ccc} B(aKa^{-1}) & \xrightarrow{i_{aKa^{-1}}} & [G\backslash(G/aKa^{-1})] \\ \downarrow{\scriptstyle c_a} & & \downarrow{\scriptstyle r_a} \\ BK & \xrightarrow{i_K} & [G\backslash(G/K)], \end{array}$$

where $c_a\colon aKa^{-1} \to K$ maps aka^{-1} to k, commutes, up to the natural isomorphism

$$i_K \circ c_a \longrightarrow r_a \circ i_{aKa^{-1}}$$

defined by the isomorphism

$$\begin{array}{ccc} (i_K \circ c_a)(0) & \longrightarrow & (r_a \circ i_{aKa^{-1}})(0) \\ \| & & \| \\ K & \xrightarrow{(a,K)} & aK \end{array}$$

in the category $[G\backslash(G/K)]$.

With all these choices made, the base-change theorem gives rise to the following result known as the double coset formula.

Theorem 10.8. *In the situation above, there is a natural isomorphism*

$$\bigoplus\nolimits_{1\leq s\leq m} c_{a_s*} \circ \mathrm{Ind}_{H\cap a_sKa_s^{-1}}^{a_sKa_s^{-1}} \circ \mathrm{Res}_{H\cap a_sKa_s^{-1}}^{H} \longrightarrow \mathrm{Res}_K^G \circ \mathrm{Ind}_H^G$$

that depends on the various choices made.

Proof. By the base-change theorem, the diagram

$$\begin{array}{ccc} \mathrm{QCoh}([G\backslash X]) & \xleftarrow{p_1^*} & \mathrm{QCoh}([G\backslash(G/H)]) \\ \downarrow{\scriptstyle p_{2*}} & & \downarrow{\scriptstyle (p_H^G)_*} \\ \mathrm{QCoh}([G\backslash(G/K)]) & \xleftarrow{(p_K^G)^*} & \mathrm{QCoh}([G\backslash(G/G)]) \end{array}$$

commutes, up to canonical natural isomorphism. Moreover, using the (non-canonical) G-equivariant bijection

$$\coprod\nolimits_{1\leq s\leq m} G/(H\cap a_sKa_s^{-1}) \xrightarrow{u} X = G/H \times G/K,$$

this translates into a diagram

$$
\begin{array}{ccc}
\prod_{1 \le s \le m} \mathrm{Rep}_k(H \cap a_s K a_s^{-1}) & \xleftarrow{(\mathrm{Res}^H_{H \cap a_s K a_s^{-1}})} & \mathrm{Rep}_k(H) \\
\big\downarrow {\scriptstyle \prod \mathrm{Ind}^{a_s K a_s^{-1}}_{H \cap a_s K a_s^{-1}}} & & \\
\prod_{1 \le s \le m} \mathrm{Rep}_k(a_s K a_s^{-1}) & & \\
\big\downarrow {\scriptstyle \prod c_{a_s *}} & & \big\downarrow {\scriptstyle \mathrm{Ind}^G_H} \\
\prod_{1 \le s \le m} \mathrm{Rep}_k(K) & & \\
\big\downarrow {\scriptstyle \oplus} & & \\
\mathrm{Rep}_k(K) & \xleftarrow{\mathrm{Res}^G_K} & \mathrm{Rep}_k(G),
\end{array}
$$

which commutes, up to a natural isomorphism that depends on the (many) choices made. The translation uses the fact, which we stated as Proposition 10.1, that adjoints of functors, if they exist, are unique, up to unique natural isomorphism. □

We will use these results to prove a theorem called the intertwining number theorem. So we let G be a finite group, and let $H, K \subset G$ be subgroup. Let (V, σ) and (W, τ) be k-linear representations of H and K, respectively. By Frobenius reciprodicity I+II and the double coset formula, we obtain isomorphisms

$$
\begin{aligned}
&\mathrm{Hom}(\mathrm{Ind}^G_H(\sigma), \mathrm{Ind}^G_K(\tau)) \simeq \mathrm{Hom}((\mathrm{Res}^G_K \circ \mathrm{Ind}^G_H)(\sigma), \tau) \\
&\simeq \bigoplus\nolimits_{1 \le s \le m} \mathrm{Hom}((c_{a_s *} \circ \mathrm{Ind}^{a_s K a_s^{-1}}_{H \cap a_s K a_s^{-1}} \circ \mathrm{Res}^H_{H \cap a_s K a_s^{-1}})(\sigma), \tau) \\
&\simeq \bigoplus\nolimits_{1 \le s \le m} \mathrm{Hom}(\mathrm{Res}^H_{H \cap a_s K a_s^{-1}}(\sigma), (\mathrm{Res}^{a_s K a_s^{-1}}_{H \cap a_s K a_s^{-1}} \circ c^*_{a_s})(\tau)).
\end{aligned}
$$

We note that for $a \in G$, the k-vector space

$$
\mathrm{Hom}(\mathrm{Res}^H_{H \cap a K a^{-1}}(\sigma), (\mathrm{Res}^{a K a^{-1}}_{H \cap a K a^{-1}} \circ c^*_a)(\tau))
$$

consists of the k-linear maps $f \colon V \to W$ such that

$$
f(\sigma(h)(\boldsymbol{v})) = \tau(a^{-1} h a)(f(\boldsymbol{v}))
$$

for all $h \in G$ and $\boldsymbol{v} \in V$, or equivalently, such that

$$
f \circ \sigma(h) = \tau(k) \circ f
$$

for all $(h, k) \in H \times K$ with $ha = ak$. Let us write $d(\sigma, \tau; s)$ for the dimension of this k-vector space for $a = a_s$. To see that it only depends on σ, τ, and s, and not on the choice of $a_s \in H a_s K \in H \backslash G / K$, we rewrite the calculation of

$$
\mathrm{Hom}(\mathrm{Ind}^G_H(\sigma), \mathrm{Ind}^G_K(\tau))
$$

in a way that does not involve any choices. If we let

$$
X_s \xrightarrow{i_s} X = G/H \times G/K
$$

be the inclusion of the sth orbit, then the calculation becomes

$$
\begin{aligned}
&\mathrm{Hom}((p^G_H)_*(\sigma), (p^G_K)_*(\tau)) \simeq \mathrm{Hom}(((p^G_K)^* \circ (p^G_H)_*)(\sigma), \tau) \\
&\simeq \mathrm{Hom}(p_{2*}\, p_1^*(\sigma), \tau) \simeq \oplus_{1 \le s \le m} \mathrm{Hom}(p_{2*}\, i_{s*}\, i_s^*\, p_1^*(\sigma), \tau) \\
&\simeq \oplus_{1 \le s \le m} \mathrm{Hom}(p_{2!}\, i_{s!}\, i_s^*\, p_1^*(\sigma), \tau) \simeq \oplus_{1 \le s \le m} \mathrm{Hom}(i_s^*\, p_1^*(\sigma), i_s^*\, p_2^*(\tau)),
\end{aligned}
$$

which, in turn, gives the formula

$$d(\sigma, \tau; s) = \dim_k \operatorname{Hom}(i_s^* \, p_1^*(\sigma), i_s^* \, p_2^*(\tau)).$$

So this number manifestly only depends on σ, τ, and s. Finally, by taking dimensions everywhere, we obtain the following theorem due to Mackey.

Theorem 10.9 (Intertwining number theorem)**.** *In the situation above,*

$$\dim_k \operatorname{Hom}(\operatorname{Ind}_H^G(\sigma), \operatorname{Ind}_K^G(\tau)) = \textstyle\sum_{1 \le s \le m} d(\sigma, \tau; s).$$

11. Representations of the symmetric groups

Let X be a finite set with n elements, and let $G = \mathrm{Aut}(X)$ be its group of automorphisms. We proceed to construct representatives for all isomorphism classes of irreducible complex representations of G. Since the set of isomorphism classes of irreducible complex representations is bijective to the set $C(G)$ of conjugacy classes of elements in G, we first introduce some language to understand this set.

We recall that every permutation $g \in G$ can be written as a product

$$g = g_1 \dots g_m$$

of disjoint cycles and that this product decomposition is unique, up to a reordering of the factors. The cycle type of $g \in G$ is the sequence $(\lambda_1, \dots, \lambda_m)$ of lengths of the cycles $g_1, \dots, g_m$, listed in non-increasing order, and it is a basic fact that two permutations $g, h \in G$ are conjugate if and only if they have the same cycle type.

Definition 11.1. Let $n \geq 0$ be an integer. A partition of n is a non-increasing sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ of non-negative integers such that $\sum_{i \geq 1} \lambda_i = n$.

Let $\mathrm{Part}(n)$ be the set of partitions of n. The map that to a permutation $g \in G$ assigns its cycle type $\lambda(g) \in \mathrm{Part}(n)$ induces a bijection

$$C(G) \longrightarrow \mathrm{Part}(n).$$

Example 11.2. Let $n = 7$, and let $g \in G$ be the permutation given by

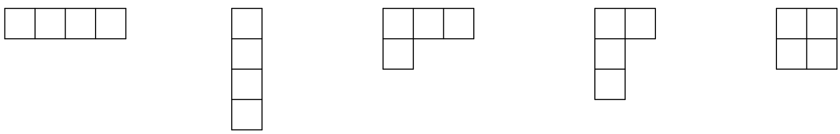
i	1	2	3	4	5	6	7
$g(i)$	5	1	6	3	2	4	7

We have $g = (152)(643)(7)$, so g has cycle type $\lambda(g) = (3, 3, 1)$.

Let us write $\mathbb{Z}_{>0}$ for the set of positive integers.

Definition 11.3. A Young diagram is a finite subset $S \subset \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ with the property that for all $(i, j) \in \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$, if either $(i + 1, j) \in S$ or $(i, j + 1) \in S$ or both, then $(i, j) \in S$. The cardinality of the set S is called the size of the Young diagram.

Example 11.4. We picture a Young diagram as a collection of boxes arranged as the entries in a matrix. For instance, there are five Young diagrams of size $n = 4$:



We will see that these correspond to the five isomorphism classes of irreducible complex representations $G = \Sigma_4$.

Given a Young diagram S of size n, we define its row partition $\lambda(S)$ by

$$\lambda(S)_i = \mathrm{card}(\{j \in \mathbb{Z}_{>0} \mid (i, j) \in S\}),$$

and we define its column partition $\mu(S)$ by

$$\mu(S)_j = \mathrm{card}(\{i \in \mathbb{Z}_{>0} \mid (i,j) \in S\}).$$

We write $\mathrm{Young}(n)$ for the set of Young diagrams of size n.

Proposition 11.5. *The maps* $\lambda, \mu\colon \mathrm{Young}(n) \to \mathrm{Part}(n)$ *that to a Young diagram assign its row partition* $\lambda(S)$ *and column partition* $\mu(S)$ *are bijections.*

Proof. This is clear from the definitions. □

Example 11.6. In Example 11.2, the row partition (resp. column partition) of the first (resp. third) Young diagram is equal to the column partition (resp. row partition) of the second (resp. fourth) Young diagram. The row and column partitions of the fifth Young diagram are equal.

Definition 11.7. Let X be a finite set. A Young tableau on X is an injective map

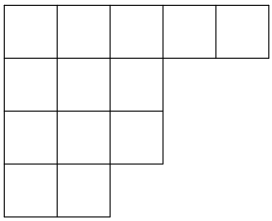
$$X \xrightarrow{\;u\;} \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$$

whose image $S = u(X)$ is a Young diagram.

Given a Young tableau, the map $u\colon X \to S = u(X)$ is a bijection. So to specify a Young tableau $u\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ with a given Young diagram S as its image amounts to assigning an element of X to each "box" in S.

Example 11.8. The figures

8	12	4	9	1
11	3	10		
5	6	13		
7	2			



illustrate a Young tableau on $X = \{1, 2, \dots, 13\}$ and its underlying Young diagram.

Let $\mathrm{Tabl}(X)$ be the set of Young tableaux on X. The group homomorphism

$$G^{\mathrm{op}} = \mathrm{Aut}(X)^{\mathrm{op}} \xrightarrow{\;\rho\;} \mathrm{Aut}(\mathrm{Tabl}(X))$$

given by $\rho(g)(u) = u \circ g$ defines a right action by the group G on $\mathrm{Tabl}(X)$. It is free action. Indeed, if $u \circ g = u$, then $g = e$, since u is injective.

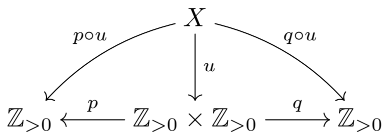
Proposition 11.9. *The map that to a Young tableau assigns its image induces a bijection*

$$\mathrm{Tabl}(X)/G \longrightarrow \mathrm{Young}(n)$$

from the set of orbits of the right action by G *on* $\mathrm{Tabl}(X)$ *onto the set of Young diagrams of size* $n = \mathrm{card}(X)$.

Proof. Indeed, the map is surjective, by the definition of a Young tableau, and it is injective, since two Young tableaux $u, v\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ have the same image if and only if there exists a bijection $g\colon X \to X$ such that $v = u \circ g$. □

Given a Young tableau u, we consider its composition with the projections



given by $p(i,j) = i$ and $q(i,j) = j$, respectively.

Definition 11.10. The row stabilizer $H \subset G$ and the column stabilizer $K \subset G$ of a Young tableau $u\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ are the subgroups

$$H = \{g \in G \mid p \circ u \circ g = p \circ u\} \subset G,$$
$$K = \{g \in G \mid q \circ u \circ g = q \circ u\} \subset G.$$

More informally, the row stabilizer consists of the permutations, which permute the elements within the rows of a Young tableau, but which do not permute elements that belong to separate rows. Similarly for the column stabilizer.

Lemma 11.11. *Let $u\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ be a Young tableau. If $H, K \subset G = \operatorname{Aut}(X)$ are its row and column stabilizers, then $H \cap K = \{e\}$.*

Proof. Indeed, if $p \circ u \circ g = p \circ u$ and $q \circ u \circ g = q \circ u$, then $u \circ g = u$, and, as we have already noticed, this implies that $g = e$, since u is injective. □

We give the set $\operatorname{Part}(n)$ of partitions of n the lexicographic order, where $\lambda > \mu$ if there exists an $m \geq 1$ such that $\lambda_m > \mu_m$ and $\lambda_i = \mu_i$ for $1 \leq i < m$. It is a total order in the sense that if $\lambda \neq \mu$, then either $\lambda > \mu$ or $\mu > \lambda$.

Lemma 11.12. *Let $u, v\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ be Young tableaux, let H be the row stabilizer of u, and let K be the column stabilizer of v. Let $S = u(X)$ and $T = v(X)$ be the underlying Young diagrams, and suppose that $\lambda(S) \geq \lambda(T)$. If, in addition, every row in u and every column in v have at most one element in common, then $S = T$ and there exists $h \in H$ and $k \in K$ such that $u \circ h = v \circ k$.*

Proof. We prove the statement by induction on $n = \operatorname{card}(X)$, the case $n = 1$ being trivial. So we let $n = m$ and assume that the statement has been proved for $n < m$. Let $X_i = (p \circ u)^{-1}(i) \subset X$ be the set of elements in the ith row of u, and let $Y_j = (q \circ v)^{-1}(j) \subset X$ for the set of elements in the jth column of v. By assumption, the intersection $X_i \cap Y_j$ has at most one element for all (i,j). Also by assumption, $\lambda(S) \geq \lambda(T)$, so in particular that $\lambda(S)_1 \geq \lambda(T)_1$. But since there are $\lambda(S)_1$ elements in X_1, and since at most one of them belongs to each of the columns $Y_1, \dots, Y_{\lambda(T)_1}$, we also have $\lambda(S)_1 \leq \lambda(T)_1$, so $\lambda(S)_1 = \lambda(T)_1$. We can now choose $h \in H$ such that for all $x \in X_1$,

$$(q \circ u \circ h)(x) = (q \circ v)(x),$$

and we can further choose $k \in K$ such that for $x \in X_1$,

$$(p \circ v \circ k)(x) = (p \circ u)(x) = 1.$$

It follows that for all $x \in X_1$, we have

$$(p \circ u \circ h)(x) = (p \circ u)(x) = (p \circ v \circ k)(x),$$
$$(q \circ u \circ h)(x) = (q \circ v)(x) = (q \circ v \circ k)(x),$$

which, in turn, implies that for all $x \in X_1$, we have

$$(u \circ h)(x) = (v \circ k)(x).$$

We can now bring ourselves in a position to invoke the inductive hypothesis. Indeed, we let $X' = X \smallsetminus X_1$, and define $u', v' \colon X' \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ by

$$\begin{aligned}
(p \circ u')(x) &= (p \circ u \circ h)(x) - 1, \\
(q \circ u')(x) &= (q \circ u \circ h)(x), \\
(p \circ v')(x) &= (p \circ v \circ k)(x) - 1, \\
(q \circ v')(x) &= (q \circ v \circ k)(x).
\end{aligned}$$

It is clear that $S' = u'(X')$ and $T' = v'(X')$ again are Young diagrams, so that u' and v' are Young tableaux; that $\lambda(S') \geq \lambda(T')$; and that every row in u' and every column in v' at most have one element of X' in common. Let $G' = \mathrm{Aut}(X')$, and let $H', K' \subset G'$ be the row stabilizer of u' and the column stabilizer of v'. Since $\mathrm{card}(X') < \mathrm{card}(X)$, we conclude from the inductive hypothesis that $S' = T'$ and that there exist $h' \in H'$ and $k' \in K'$ such that $u' \circ h' = v' \circ k'$. We conclude that $S = T$. Moreover, since the group homomorphism $\rho \colon G' \to G$ defined by

$$\rho(g')(x) = \begin{cases} x & \text{if } x \in X_1, \\ g'(x) & \text{if } x \in X', \end{cases}$$

maps H' and K' into H and K, respectively, we further conclude that

$$u \circ h \circ \rho(h') = v \circ k \circ \rho(k').$$

This completes the proof. □

We can now reap the benefits of the work that we did in the last two lectures together with the lemmas above and classify all irreducible complex representations of $G = \mathrm{Aut}(X)$, up to non-canonical isomorphism. Let S be a Young diagram, let $u \colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ be a Young tableau with $u(X) = S$, and let $H, K \subset G$ be its row and column stabilizers. We define

$$\begin{aligned}
\pi_S^+ &= (\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\tau) \\
\pi_S^- &= (\mathrm{Ind}_K^G \circ \mathrm{Res}_K^G)(\sigma),
\end{aligned}$$

where τ is the 1-dimensional trivial representation of G and σ is the 1-dimensional sign representation of G.

Theorem 11.13. *Let X be a finite set with n elements, and let $G = \mathrm{Aut}(X)$.*

(1) *If S is a Young diagram of size n, then, up to non-canonical isomorphism, there is a unique irreducible complex representation π_S of G, which occurs in the decompositions of both π_S^+ and π_S^-.*
(2) *If S and T are distinct Young diagrams of size n, then the representations π_S and π_T are non-isomorphic.*
(3) *If π is an irreducible complex representation of G, then $\pi \simeq \pi_S$, for some Young diagram S of size n.*

Proof. To prove (1), it suffices to show that

$$\dim_{\mathbb{C}} \mathrm{Hom}(\pi_S^+, \pi_S^-) = 1,$$

and to do so, we will use the results on induced representations that we proved in the last two lectures. We consider the cartesian diagram of left G-sets

$$\begin{array}{ccc} G/H \times G/K & \xrightarrow{p_1} & G/H \\ \downarrow{\scriptstyle p_2} & \searrow^{f} & \downarrow{\scriptstyle p_H^G} \\ G/K & \xrightarrow{p_K^G} & G/G \end{array}$$

where we have included $f = p_H^G \circ p_1 = p_K^G \circ p_2$. We have canonical isomorphisms

$$\begin{aligned} \operatorname{Hom}(\pi_S^+, \pi_S^-) &= \operatorname{Hom}((p_H^G)_*(p_H^G)^*\tau, (p_K^G)_*(p_K^G)^*\sigma) \\ &\simeq \operatorname{Hom}((p_K^G)^*(p_H^G)_*(p_H^G)^*\tau, (p_K^G)^*\sigma) \\ &\simeq \operatorname{Hom}((p_K^G)_!\,(p_K^G)^*(p_H^G)_*(p_H^G)^*\tau, \sigma) \\ &\simeq \operatorname{Hom}((p_K^G)_!\,(p_K^G)^*(p_H^G)_!\,(p_H^G)^*\tau, \sigma) \\ &\simeq \operatorname{Hom}((p_K^G)_!\,p_{2!}\,p_1^*(p_H^G)^*\tau, \sigma) \\ &\simeq \operatorname{Hom}(f_!\,f^*\tau, \sigma) \\ &\simeq \operatorname{Hom}(f^*\tau, f^*\sigma). \end{aligned}$$

Moreover, we defined a non-canonical isomorphism of left G-sets

$$\textstyle\coprod_{1 \le s \le m} G/H \cap a_s K a_s^{-1} \longrightarrow G/H \times G/K,$$

which depends on a choice of a family $(a_1, \dots, a_m)$ of representatives of the double cosets $H \backslash G/K$, among other things. So we conclude that

$$\operatorname{Hom}(\pi_S^+, \pi_S^-) \simeq \textstyle\prod_{1 \le s \le m} \operatorname{Hom}(\operatorname{Res}^G_{H \cap a_s K a_s^{-1}}(\tau), \operatorname{Res}^G_{H \cap a_s K a_s^{-1}}(\sigma)).$$

Since both $\operatorname{Res}^G_{H \cap aKa^{-1}}(\tau)$ and $\operatorname{Res}^G_{H \cap aKa^{-1}}(\sigma)$ are 1-dimensional representations of $H \cap aKa^{-1}$, we find that

$$\begin{aligned} &\dim_{\mathbb{C}} \operatorname{Hom}(\operatorname{Res}^G_{H \cap aKa^{-1}}(\tau), \operatorname{Res}^G_{H \cap aKa^{-1}}(\sigma)) \\ &\qquad = \begin{cases} 1 & \text{if } \operatorname{sgn}(g) = 1 \text{ for all } g \in H \cap aKa^{-1}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

For the double coset $HaK = HK$, we have

$$\dim_{\mathbb{C}} \operatorname{Hom}(\operatorname{Res}^G_{H \cap K}(\tau), \operatorname{Res}^G_{H \cap K}(\sigma)) = 1,$$

since $H \cap K = \{e\}$ by Lemma 11.11. Hence, we must show that if $a \notin HK$, then there exists $g \in H \cap aKa^{-1}$ such that $\operatorname{sgn}(g) = -1$. To this end, we consider, in addition to the tableau $u \colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$, the tableau $v = u \circ a^{-1} \colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$, whose column stabilizer is aKa^{-1}. We claim that there exists a row in u and a column in v, which have at least two elements in common. Granting this claim, the transposition g that interchanges these two elements belongs to $H \cap aKa^{-1}$ and has $\operatorname{sgn}(g) = -1$, which proves (1). To prove claim, we assume that every row in u and every column in v have at most one element in common. In this case, Lemma 11.12 shows that there exists $h \in H$ and $aka^{-1} \in aKa^{-1}$ such that

$$u \circ h = v \circ aka^{-1} = u \circ ka^{-1}.$$

But then $h = ka^{-1}$, so $a = h^{-1}k \in HK$, which is a contradicts that $a \notin HK$.

To prove (2), it suffices to show that

$$\dim_{\mathbb{C}} \mathrm{Hom}(\pi_S^+, \pi_T^-) = 0.$$

Arguing as in the proof of (1), we see that it further suffices to show that for all Young tableaux $u, v\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ with $u(X) = S$ and $v(X) = T$, there exists a row in u and a column in v that have at least two elements in common. But this follows immediately from Lemma 11.12. Indeed, since the lexicographic order on $\mathrm{Part}(n)$ is a total order, we can assume without loss of generality that $\lambda(S) \geq \lambda(T)$.

Finally, we prove (3). We have constructed the family

$$(\pi_S)_{S \in \mathrm{Young}(n)}$$

of pairwise non-isomorphic irreducible complex representations of G. But the set of Young diagrams of size n and the set of conjugacy classes of elements in G are bijective, so we have found all irreducible complex representations of G, up to non-canonical isomorphism. □

The representation π_S is called the Specht representation associated with the Young diagram S. Its isomorphism class is independent of the choice of Young tableau u that we made in its definition.

Remark 11.14. We defined $\pi_S^+ = (\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\tau)$ and $\pi_S^- = (\mathrm{Ind}_K^G \mathrm{Res}_K^G)(\sigma)$, but we could of course just as well have switched τ and σ in this definition.

If the subset $S \subset \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ is a Young diagram of size n, then so is the subset

$$S' = \{(i,j) \in \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \mid (j,i) \in S\},$$

which we call the conjugate Young diagram of S.

Lemma 11.15. *If S is a Young diagram of size n, and if S' is its conjugate Young diagram, then the associated Specht representations are related by*

$$\pi_{S'} \simeq \pi_S \otimes \sigma.$$

Proof. Let $u\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ be a Young tableau with $u(X) = S$, and let H and K be its row stabilizer and column stabilizer. Let $u'\colon X \to \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ be the unique map with $p \circ u' = q \circ u$ and $q \circ u' = p \circ u$. Then $u'(X) = S'$ and u' has row stabilizer K and column stabilizer H. Thus,

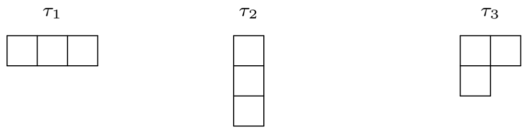
$$\begin{aligned}
\pi_{S'}^+ &= (\mathrm{Ind}_K^G \circ \mathrm{Res}_K^G)(\tau) \simeq (\mathrm{Ind}_K^G \circ \mathrm{Res}_K^G)(\sigma \otimes \sigma) \\
&\simeq (\mathrm{Ind}_K^G \circ \mathrm{Res}_K^G)(\sigma) \otimes \sigma \simeq \pi_S^- \otimes \sigma \\
\pi_{S'}^- &= (\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\sigma) \simeq (\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\tau \otimes \sigma) \\
&\simeq (\mathrm{Ind}_H^G \circ \mathrm{Res}_H^G)(\tau) \otimes \sigma \simeq \pi_S^+ \otimes \sigma.
\end{aligned}$$

Here we have used that, in general, for $H \subset G$, one has

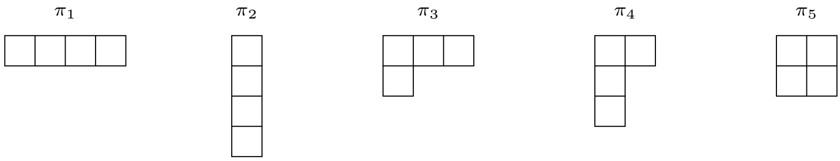
$$\mathrm{Res}_H^G(\pi \otimes \rho) \simeq \mathrm{Res}_H^G(\pi) \otimes \mathrm{Res}_H^G(\rho)$$
$$\mathrm{Ind}_H^G(\sigma \otimes \mathrm{Res}_H^G(\rho)) \simeq \mathrm{Ind}_H^G(\sigma) \otimes \rho$$

for all representations π and ρ of G and σ of H. The latter identity is called the projection formula. □

Example 11.16. For $H = \Sigma_3$, we have earlier found three irreducible finite dimensional complex representations of H, namely, the 1-dimensional trivial representation τ_1 and sign representation τ_2, and the 2-dimensional standard representation τ_3. These correspond to the following Specht representations:



Similarly, for $G = \Sigma_4$, we have earlier found five irreducible finite dimensional complex representations of G, namely, the 1-dimensional trivial representation π_1 and sign representation π_2, the 3-dimensional standard representation π_3 and its tensor product $\pi_4 = \pi_2 \otimes \pi_3$ with the sign representation, and the 2-dimensional representation π_5. These correspond to the following Specht representations:



Using Lemma 11.15, we see immediately from these listings that $\tau_2 \otimes \tau_3 \simeq \tau_3$ and that $\pi_2 \otimes \pi_5 \simeq \pi_5$. If we identify H with the subgroup of G consisting of all $g \in G$ with $g(4) = 4$, then one can also show that, in terms of Young diagrams, Res_H^G takes an irreducible G-representation π to the sum with multiplicity one of all irreducible H-representations τ corresponding to the Young diagrams obtained from the Young diagram for π by removing one box. So we have

$$\begin{aligned}
\mathrm{Res}_H^G(\pi_1) &\simeq \tau_1 \\
\mathrm{Res}_H^G(\pi_2) &\simeq \tau_2 \\
\mathrm{Res}_H^G(\pi_3) &\simeq \tau_1 \oplus \tau_3 \\
\mathrm{Res}_H^G(\pi_4) &\simeq \tau_2 \oplus \tau_3 \\
\mathrm{Res}_H^G(\pi_5) &\simeq \tau_3
\end{aligned}$$

Similarly, one can show that Ind_H^G takes an irreducible H-representation τ to the sum with multiplicity one of all irreducible G-representations π corresponding to the Young diagrams obtained from the Young diagram associated with τ by adding one box. So we find that

$$\begin{aligned}
\mathrm{Ind}_H^G(\tau_1) &\simeq \pi_1 \oplus \pi_3 \\
\mathrm{Ind}_H^G(\tau_2) &\simeq \pi_2 \oplus \pi_3 \\
\mathrm{Ind}_H^G(\tau_3) &\simeq \pi_3 \oplus \pi_4 \oplus \pi_5,
\end{aligned}$$

which is also what we have calculated directly before.

Finally, we mention that for Young diagrams S and T, Frobenius has given a formula for the value $\chi_{\pi_S}(g)$ of the character of the Specht representation π_S on an element g in the conjugacy class corresponding to T in terms of combinatorial data that can be read off from the Young diagrams S and T directly. The formula is called the Frobenius character formula.

This week's lecture will cover Chapter 7 in the book, but I will begin more generally by defining the socalled classical (matrix) groups. These will be subgroups of the groups $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{C})$, and $\mathrm{GL}_n(\mathbb{H})$ of invertible $n \times n$-matrices with entries in the real numbers, the complex numbers, and the quaternions, respectively.

If $k = (k, +, \cdot\,)$ is a ring, then we define the opposite ring $k^{\mathrm{op}} = (k, +, \star)$ to have the same set of elements and the same addition but the opposite multiplication

$$a \star b = b \cdot a.$$

If k is a division ring, then so is k^{op}.

Definition 12.1. Let k be a ring. A ring homomorphism

$$k \xrightarrow{\sigma} k^{\mathrm{op}}$$

is an antiinvolution, if $\sigma \circ \sigma = \mathrm{id}$.

In particular, an antiinvolution is an isomorphism. We remark that the identity map $\mathrm{id}_k \colon k \to k$ is an antiinvolution if and only if k is commutative. We will often write a^* or $\overline{a}$ instead of $\sigma(a)$.

Example 12.2. (1) If $k = \mathbb{R}$, then the identity map is an antiinvolution, and one can show that it is the only one.

(2) If $k = \mathbb{C}$, then the identity map and complex conjugation are antiinvolutions.

(3) If $k = \mathbb{H}$, then quaternionic conjugation, which is the map $\sigma \colon \mathbb{H} \to \mathbb{H}$ that to the quaternion $q = a + ib + jc + kd$ assigns the quaternion

$$q^* = a - ib - jc - kd$$

is an antiinvolution. The identity map $\mathrm{id}_{\mathbb{H}} \colon \mathbb{H} \to \mathbb{H}$ is not an antiinvolution.

Definition 12.3. Let k be a division ring, and let $\sigma \colon k \to k^{\mathrm{op}}$ be an antiinvolution. The adjoint matrix of $A = (a_{ij}) \in M_{m,n}(k)$ is $A^* = (a_{ji}^*) \in M_{n,m}(k)$.[31]

The number of rows in A^* is equal to the number of columns in A and vice versa. So it is only meaningful to ask whether $A = A^*$ if A is a square matrix. If k is a field and $\sigma \colon k \to k^{\mathrm{op}}$ is the identity map, then it is customary to call A^* the transpose matrix of A and to denote it by A^t instead of A^*.

Proposition 12.4. *Let k be a division ring, and let $\sigma \colon k \to k^{\mathrm{op}}$ be an antiinvolution. For all matrices A, B, and C of appropriate dimensions, the following hold:*

(I1) $(A + B)^* = A^* + B^*$
(I2) $(AB)^* = B^* A^*$
(I3) $E^* = E$
(I4) $(A^*)^* = A$

Proof. Let us prove (2). For the purpose of this proof, given $A \in M_{m,n}(k)$, we write $A^* = (a'_{ij}) \in M_{n,m}(k)$. So $a'_{ij} = a^*_{ji}$ by the definition of the adjoint matrix. We let

[31] The notation $A^\dagger$ for the adjoint matrix is also common, particularly in physics.

$A \in M_{m,n}(k)$ and $B \in M_{n,p}(k)$ with product $C = AB \in M_{m,p}(k)$ and calculate

$$c'_{ik} = c^*_{ki} = (\sum_{j=1}^{n} a_{kj} b_{ji})^* = \sum_{j=1}^{n} (a_{kj} b_{ji})^* = \sum_{j=1}^{n} b^*_{ji} a^*_{kj} = \sum_{j=1}^{n} b'_{ij} a'_{jk}.$$

This proves (2), and the remaining identities are proved analogously. □

Definition 12.5. Let k be a division ring, and let $\sigma\colon k \to k^{\mathrm{op}}$ be an antiinvolution. A square matrix $A \in M_n(k)$ is hermitian, if $A^* = A$, and it is skew-hermitian, if $A^* = -A$.

If k is a field and $\sigma\colon k \to k^{\mathrm{op}}$ is the identity map, then it is customary to say that $A \in M_n(k)$ is symmetric, if $A^t = A$, and that A is skew-symmetric, if $A^t = -A$.

We will now consider vector spaces over the division ring k, and we will always consider right vector spaces in the sense that scalars multiply from the right.

Definition 12.6. Let k be a division ring, let $\sigma\colon k \to k^{\mathrm{op}}$ be an antiinvolution, and let V be a right k-vector space. A hermitian form on V is a map

$$V \times V \xrightarrow{\langle -,- \rangle} k$$

such that the following hold for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ and $a \in k$:

(H1) $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$
(H2) $\langle \boldsymbol{x}, \boldsymbol{y} \cdot a \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \cdot a$
(H3) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
(H4) $\langle \boldsymbol{x} \cdot a, \boldsymbol{y} \rangle = a^* \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle$
(H5) $\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle^*$

Example 12.7. Let k be a division ring, and let $\sigma\colon k \to k^{\mathrm{op}}$ be an antiinvolution. Let $k^n = M_{n,1}(k)$ be the right k-vector space of column n-matrices with entries in k. If $A \in M_n(k)$ is a hermitian matrix, then the map $\langle -,- \rangle\colon k^n \times k^n \to k$ defined by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* A \boldsymbol{y}$ is a hermitian form. Conversely, if $\langle -,- \rangle\colon k^n \times k^n \to k$ is a hermitian form, then the matrix $A = (a_{i,j}) \in M_n(k)$ with entries $a_{ij} = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$ is a hermitian matrix.

If $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and if $\sigma\colon k \to k^{\mathrm{op}}$ is the identity map, complex conjugation, and quaternionic conjugation, respectively, then for all $a \in k$, $a^* = a$ if and only if $a \in \mathbb{R} \subset k$. In particular, if $\langle -,- \rangle$ is a hermitian form on a right real, complex, or quaternionic vector space V, then for all $\boldsymbol{x} \in V$, we have $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \in \mathbb{R}$.

Definition 12.8. Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $\sigma\colon k \to k^{\mathrm{op}}$ be the identity map, complex conjugation, and quaternionic conjugation, respectively. A hermitian inner product on a right k-vector space V is a hermitian form $\langle -,- \rangle\colon V \times V \to k$ such that, in addition to (H1)–(H5), the following positivity property holds:

(P) For all $\boldsymbol{0} \neq \boldsymbol{x} \in V$, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$.

Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $\sigma\colon k \to k^{\mathrm{op}}$ be the identity map, complex conjugation, and quaternionic conjugation, respectively. The standard hermitian inner product on the right k-vector space $k^n = M_{n,1}(k)$ of column n-vectors is defined to be the map $\langle -,- \rangle\colon k^n \times k^n \to$ given by the matrix product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \boldsymbol{y},$$

which is meaningful, since $\boldsymbol{x}^* \in M_{1,n}(k)$ and $\boldsymbol{y} \in M_{n,1}(k)$.

Definition 12.9. Let $(U, \langle -,-\rangle_U)$ and $(V, \langle -,-\rangle_V)$ be right real, complex or quaternionic vector spaces with hermitian inner products. A k-linear map $f\colon V \to U$ is an isometry with respect to the given hermitian inner products if

$$\langle f(\boldsymbol{x}), f(\boldsymbol{y})\rangle_U = \langle \boldsymbol{x}, \boldsymbol{y}\rangle_V$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$.

An isometry $f\colon U \to V$ is always injective, but it need not be an isomorphism. However, if it is an isomorphism, then the inverse map $f^{-1}\colon U \to V$ is automatically an isometry. In particular, an endomorphism $f\colon V \to V$ of a finite dimensional real, complex, or quaternionic vector space that is an isometry with respect to a given hermitian inner product is automatically an isometric isomorphism.

Definition 12.10. Let $(U, \langle -,-\rangle_U)$ and $(V, \langle -,-\rangle_V)$ be right real, complex or quaternionic vector spaces with hermitian inner products. Two k-linear maps $f\colon V \to U$ and $g\colon U \to V$ are adjoint with respect to the given hermitian inner products if

$$\langle \boldsymbol{x}, f(\boldsymbol{y})\rangle_U = \langle g(\boldsymbol{x}), \boldsymbol{y}\rangle_V$$

for all $\boldsymbol{x} \in U$ and $\boldsymbol{y} \in V$.

If both $g\colon U \to V$ and $h\colon U \to V$ are adjoint to $f\colon V \to U$, then $g = h$, so if an adjoint of $f\colon V \to U$ exists, then it is unique. If U and V are finite dimensional, then an adjoint always exists.

Proposition 12.11. *Let $(U, \langle -,-\rangle_U)$ and $(V, \langle -,-\rangle_V)$ be finite dimensional right real, complex, or quaternionic vector spaces with hermitian inner products, and let $f\colon V \to U$ be a linear map. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be bases of U and V that are orthonormal with respect to $\langle -,-\rangle_U$ and $\langle -,-\rangle_V$, respectively.*[32]

(1) *There exists a unique linear map $g\colon U \to V$ that is adjoint to $f\colon V \to U$ with respect to $\langle -,-\rangle_U$ and $\langle -,-\rangle_V$.*
(2) *If the matrix $A \in M_{m,n}(k)$ represents $f\colon V \to U$ with respect to the given orthonormal bases, then the adjoint matrix $A^* \in M_{n,m}(k)$ represents $g\colon U \to V$ with respect to these bases.*

Proof. We claim that if $f\colon V \to U$ and $g\colon U \to V$ are the linear maps represented by $A \in M_{m,n}(k)$ and $A^* \in M_{n,m}(k)$ with respect to the given orthonormal bases, then these two maps are adjoint with respect to the given hermitian inner products. Indeed, let $\boldsymbol{u} \in U$ and $\boldsymbol{v} \in V$, and let $\boldsymbol{x} \in k^m$ and $\boldsymbol{y} \in k^n$ be their coordinates with respect to the given bases. Since the bases are orthonormal, we find

$$\langle \boldsymbol{u}, f(\boldsymbol{v})\rangle_U = \boldsymbol{x}^* A \boldsymbol{y} = \boldsymbol{x}^* (A^*)^* \boldsymbol{y} = (A^* \boldsymbol{x})^* \boldsymbol{y} = \langle g(\boldsymbol{u}), \boldsymbol{v}\rangle_V.$$

This proves the proposition, since an adjoint map, if it exists, is unique. □

Lemma 12.12. *Let $k - \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $(U, \langle -,-\rangle)_U$ and $(V, \langle -,-\rangle_V)$ be right k-vector spaces with hermitian inner product. If $f\colon V \to U$ and $g\colon U \to V$ are adjoint with respect to $\langle -,-\rangle_U$ and $\langle -,-\rangle_V$, then $f\colon V \to U$ is a linear isometry with respect to $\langle -,-\rangle_U$ and $\langle -,-\rangle_V$ if and only if $g \circ f = \mathrm{id}_V$.*

[32] This means that $\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle_U = \delta_{ij}$ and $\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle_V = \delta_{ij}$.

Proof. We find that $f\colon V \to U$ is a linear isometry if and only if

$$\langle (g \circ f)(\boldsymbol{x}), \boldsymbol{y} \rangle_V = \langle \boldsymbol{x}, \boldsymbol{y} \rangle_V$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$. If $g \circ f = \mathrm{id}_V$, then this is certainly true, and conversely, we find, by setting $\boldsymbol{y} = (g \circ f)(\boldsymbol{x}) - \boldsymbol{x}$, that

$$\langle \boldsymbol{y}, \boldsymbol{y} \rangle_V = \langle (g \circ f)(\boldsymbol{x}) - \boldsymbol{x}, \boldsymbol{y} \rangle_V = \langle (g \circ f)(\boldsymbol{x}), \boldsymbol{y} \rangle_V - \langle \boldsymbol{x}, \boldsymbol{y} \rangle_V = 0,$$

which shows that $g \circ f = \mathrm{id}_V$, because $\langle -, - \rangle_V$ is an inner product. □

Theorem 12.13. *Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $(U, \langle -, - \rangle_U)$ and $(V, \langle -, - \rangle_V)$ be finite dimensional right k-vector spaces with hermitian inner products. Let $f\colon V \to U$ be a linear map, and let $A \in M_{m,n}(k)$ be the matrix that represents $f\colon V \to U$ with respect to bases $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ of U and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V that are orthonormal with respect to the given hermitian inner products. The following* (1)–(3) *are equivalent.*

(1) *The map $f\colon V \to U$ is a linear isometry.*
(2) *The matrix identity $A^*A = E_n$ holds.*
(3) *The family $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ of vectors in k^m consisting of the columns of A is orthonormal with respect to the standard hermitian inner product.*

In addition, the following (4)–(6) *are equivalent.*

(4) *The map $f\colon V \to U$ is an isometric isomorphism.*
(5) *The matrix A is invertible and $A^{-1} = A^*$.*
(6) *The family $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ of columns of A is a basis of k^m that is orthonormal with respect to the standard hermitian inner product.*

Proof. By Proposition 12.11, the adjoint map $g\colon U \to V$ is represented by the adjoint matrix $A^* \in M_{n,m}(k)$ with respect to the given bases, so the equilvalence of (1) and (2) follows from Lemma 12.12. The (i,j)th entry in A^*A is $\boldsymbol{a}_i^*\boldsymbol{a}_j$, which, by definition, is the standard hermitian inner product of $\boldsymbol{a}_i, \boldsymbol{a}_j \in k^m$, from which the equivalence of (2) and (3) follows. To prove the equivalence of (4) and (5), we note that $f\colon V \to U$ is an isomorphism if and only if A is invertible, in which case

$$A^{-1} = (A^*A)A^{-1} = A^*(AA^{-1}) = A^*.$$

Finally, the equivalence of (5) and (6) uses that an $n \times n$-matrix invertible if and only if the family consisting of its columns is a basis of k^n. □

Corollary 12.14. *Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $(V, \langle -, - \rangle)$ be a finite dimensional right k-vector space with hermitian inner product, and let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V that is orthonormal with respect to $\langle -, - \rangle$. Let $f\colon V \to V$ be an endomorphism, and let $A \in M_n(k)$ be the matrix that represents $f\colon V \to V$ with $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$.*

(1) *The endomorphism $f\colon V \to V$ is an isometry with respect to $\langle -, - \rangle$ if and only if $A^*A = E_n$. If so, then A is invertible and $A^{-1} = A^*$.*
(2) *The endomorphism $f\colon V \to V$ is selfadjoint*[33] *with respect to $\langle -, - \rangle$ if and only if $A^* = A$.*

Proof. The statement (1) follows from Theorem 12.13 and from the fact that a square matrix that has a right inverse is invertible. This fact, in turn, is a consequence of Gauss elimination. The statement (2) follows from Proposition 12.11. □

[33] This means that $f\colon V \to V$ and its adjoint $g\colon V \to V$ with respect to $\langle -, - \rangle$ are equal.

Remark 12.15. A matrix $P \in \mathrm{GL}_n(k)$ such that $P^* = P^{-1}$ is said to be orthogonal, if $k = \mathbb{R}$, unitary, if $k = \mathbb{C}$, and quaternionic unitary, if $k = \mathbb{H}$. A matrix $A \in M_n(k)$ such that $A^* = A$ is said to be symmetric, if $k = \mathbb{R}$, hermitian, if $k = \mathbb{C}$, and quaternionic hermitian, if $k = \mathbb{H}$.

We now define the classical groups. The subgroups

$$\begin{aligned} O(n) &= \{Q \in \mathrm{GL}_n(\mathbb{R}) \mid Q^* = Q^{-1}\} \subset \mathrm{GL}_n(\mathbb{R}) \\ U(n) &= \{U \in \mathrm{GL}_n(\mathbb{C}) \mid U^* = U^{-1}\} \subset \mathrm{GL}_n(\mathbb{C}) \\ \mathrm{Sp}(n) &= \{S \in \mathrm{GL}_n(\mathbb{H}) \mid S^* = S^{-1}\} \subset \mathrm{GL}_n(\mathbb{H}) \end{aligned}$$

are called the orthogonal group, the unitary group, and the compact symplectic group. They are topological groups with respect to the subspace topology from the metric topology on $M_n(k)$, and they are all compact. In particular, we have

$$\begin{aligned} O(1) &= \{x \in \mathrm{GL}_1(\mathbb{R}) \mid x^*x = 1\} \subset \mathrm{GL}_1(\mathbb{R}) \\ U(1) &= \{z \in \mathrm{GL}_1(\mathbb{C}) \mid z^*z = 1\} \subset \mathrm{GL}_1(\mathbb{C}) \\ \mathrm{Sp}(1) &= \{q \in \mathrm{GL}_1(\mathbb{H}) \mid q^*q = 1\} \subset \mathrm{GL}_1(\mathbb{H}), \end{aligned}$$

so as topological spaces, these are the unit 0-sphere S^0, the unit 1-sphere S^1, and the unit 3-sphere S^3, respectively. If $A = Q \in O(n)$ or $A = U \in U(n)$, then

$$\det(A)^* = \det(A^*) = \det(A^{-1}) = \det(A)^{-1}$$

so $\det(Q) \in O(1)$ and $\det(U) \in U(1)$. The subgroups

$$\begin{aligned} SO(n) &= \{Q \in O(n) \mid \det(Q) = 1\} \subset O(n) \\ SU(n) &= \{U \in U(n) \mid \det(U) = 1\} \subset U(n) \end{aligned}$$

are called the special orthogonal group and the special unitary group, respectively. There is no useful determinant of quaternionic square matrices, because the division ring $\mathbb{H}$ is noncommutative.[34]

We embed $\mathbb{C}$ in $\mathbb{H}$ as the subfield $L \subset \mathbb{H}$ consisting of all quaternions of the form $q = a + ib$. The subfield $L \subset \mathbb{H}$ is a maximal subfield, and if also $L' \subset \mathbb{H}$ is a maximal subfield, then there exists $q \in \mathbb{H}$ such that $L' = qLq^{-1}$. So every maximal subfield of $\mathbb{H}$ is isomorphic to $\mathbb{C}$, but the embedding of $\mathbb{C}$ as a maximal subfield in $\mathbb{H}$ is well-defined, up to conjugation, only. Left multiplication by $q = z_1 + jz_2 \in \mathbb{H}$ defines an L-linear map $\lambda(q)\colon \mathbb{H} \to \mathbb{H}$, and hence, a ring homomorphism

$$\mathbb{H} \xrightarrow{\lambda} \mathrm{End}_L(\mathbb{H}).$$

Since $\mathbb{H}$ is a division ring, the kernel of λ is either $\{0\}$ or $\mathbb{H}$, and since $\lambda(1) = \mathrm{id}_{\mathbb{H}} \neq 0$, we conclude that the kernel is $\{0\}$. Let us choose the basis $(1, j)$ of $\mathbb{H}$ as a right L-vector space. This defines a ring isomorphism

$$\mathrm{End}_L(\mathbb{H}) \xrightarrow{\mu} M_2(L)$$

that to an L-linear map $f\colon \mathbb{H} \to \mathbb{H}$ assigns the matrix $A = \mu(f) \in M_2(L)$ that represents $f\colon \mathbb{H} \to \mathbb{H}$ with respect to the basis $(1, j)$. The calculation

$$\begin{aligned} q \cdot 1 &= (z_1 + jz_2) \cdot 1 = 1 \cdot z_1 + j \cdot z_2 \\ q \cdot j &= (z_1 + jz_2) \cdot j = j \cdot z_1^* - 1 \cdot z_2^* \end{aligned}$$

[34] The best one has is the Dieudonné determinant in $K_1(\mathbb{H}) = (\mathbb{R}_{>0}, \cdot\,)$.

shows that the composite ring homomorphism

$$\mathbb{H} \xrightarrow{f=\mu\circ\lambda} M_2(L)$$

takes the quaternion $q = z_1 + jz_2$ to the matrix

$$f(q) = \begin{pmatrix} z_1 & -z_2^* \\ z_2 & z_1^* \end{pmatrix}.$$

A map between topological groups is an isomorphism if and only if it is both an isomorphism of groups and a homeomorphism of topological spaces.

Proposition 12.16. *The ring homomorphism $f\colon \mathbb{H} \to M_2(L)$ induces an isomorphism of topological groups $h\colon \mathrm{Sp}(1) \to SU(2)$.*

Proof. We have $q^* = (z_1 + jz_2)^* = z_1^* + z_2^* j^* = z_1^* - jz_2$. Therefore,

$$q^*q = (z_1^* - jz_2)(z_1 + jz_2) = z_1^*z_1 + jz_1z_2 - jz_2z_1 + z_2^*z_2 = z_1^*z_1 + z_2^*z_2,$$

which shows that $q \in \mathrm{Sp}(1)$ if and only if $f(q) \in SU(2)$. So the ring homomorphism $f\colon \mathbb{H} \to M_2(K)$ restricts to a group homomorphism $h\colon \mathrm{Sp}(1) \to SU(2)$, which is continuous because $f\colon \mathbb{H} \to M_2(K)$ is continuous. We wish to prove that h is both an isomorphism of groups and a homeomorphism of spaces, and to do so, it suffices to show that h is a bijection. Indeed, the inverse map of a bijective group homomorphism is automatically a group homomorphism, and the inverse map of a continuous bijection from a compact space such as $\mathrm{Sp}(1)$ to a Hausdorff space such as $SU(2)$ is automatically continuous. Now, the map h is injective, because the map f is injective, and the map h is surjective because, if

$$U = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \in SU(2),$$

then $U = f(q)$ with $q = z_{11} + jz_{21}$. This completes the proof. $\square$

Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. We define the Hilbert–Schmidt norm of $A \in M_n(k)$ by

$$\|A\| = \sqrt{\mathrm{tr}(A^*A)}\,.$$

It satisfies $\|A+B\| \leq \|A\| + \|B\|$ and $\|AB\| \leq \|A\|\|B\|$ for all $A, B \in M_n(k)$, so in particular, the exponential series

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

converges absolutely. If $[A, B] = AB - BA = 0$, then

$$\exp(A + B) = \exp(A)\exp(B),$$

but in general the left-hand side and the right-hand side are different.[35] Hence, the matrix $\exp(A)$ is invertible with inverse $\exp(-A)$, so we get a map

$$M_n(k) \xrightarrow{\exp} \mathrm{GL}_n(k).$$

Locally on $M_n(k)$, this map is a diffeomorphism. For it is a smooth map (considered as map between open subsets of $\mathbb{R}^m$) with derivative $\mathrm{id}\colon M_n(k) \to M_n(k)$, so the inverse function theorem shows that it is a diffeomorphism locally on $M_n(k)$.

[35] The difference is given by the Baker–Campbell–Hausdorff formula.

If $G \subset \mathrm{GL}_n(k)$ is one of the classical groups, then we define its Lie algebra to the be the subset $\mathfrak{g} \subset M_n(k)$ consisting of all matrices A such that $\exp(tA) \in G$, for all $t \in \mathbb{R}$. It is a real subspace of $M_n(k)$.

Proposition 12.17. *The Lie algebras of the classical groups are given by*

$$\begin{aligned} \mathfrak{o}(n) &= \{ A \in M_n(\mathbb{R}) \mid A^* + A = 0 \} \\ \mathfrak{u}(n) &= \{ A \in M_n(\mathbb{C}) \mid A^* + A = 0 \} \\ \mathfrak{sp}(n) &= \{ A \in M_n(\mathbb{H}) \mid A^* + A = 0 \} \\ \mathfrak{so}(n) &= \{ A \in \mathfrak{o}(n) \mid \mathrm{tr}(A) = 0 \} \\ \mathfrak{su}(n) &= \{ A \in \mathfrak{u}(n) \mid \mathrm{tr}(A) = 0 \} \end{aligned}$$

Proof. We prove the statements for $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$; the proofs in the remaining cases are analogous. If $A \in \mathfrak{u}(n)$, then for all $t \in \mathbb{R}$, we have

$$\exp(tA^*) = \exp(tA)^* = \exp(tA)^{-1} = \exp(-tA),$$

and since the exponential map is a local diffeomorphism, this implies that $A^* = -A$. Similarly, if $A \in \mathfrak{su}(n)$, then we have in addition that for all $t \in \mathbb{R}$,

$$\exp(nt\,\mathrm{tr}(A)) = \exp(\mathrm{tr}(tA)) = \det(\exp(tA)) = 1.$$

Since the exponential map is a local diffeomorphism, this implies that $\mathrm{tr}(A) = 0$. □

Example 12.18. The Lie algebra $\mathfrak{sp}(1) \subset \mathbb{H}$ is the 3-dimensional real subspace of purely imaginary quaternions. One can show that $\exp\colon \mathfrak{sp}(1) \to \mathrm{Sp}(1)$ is given by

$$\exp(v) = \cos|v| + \frac{v}{|v|}\sin|v|,$$

where $|v| = \sqrt{v^*v}$.

Lemma 12.19. *Let $G \subset \mathrm{GL}_n(k)$ be one of the classical groups, and let $\mathfrak{g} \subset M_n(k)$ be its Lie algebra. If $g \in G$ and $A \in \mathfrak{g}$, then $gAg^{-1} \in \mathfrak{g}$.*

Proof. Indeed, for all $t \in \mathbb{R}$, we have

$$\exp(tgAg^{-1}) = \exp(gtAg^{-1}) = g\exp(tA)g^{-1},$$

so if $\exp(tA) \in G$ and $g \in G$, then also $\exp(tgAg^{-1}) \in G$. □

Definition 12.20. The adjoint representation of the classical group $G \subset \mathrm{GL}_n(k)$ on its Lie algebra $\mathfrak{g} \subset M_n(k)$ is the real representation

$$G \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{g})$$

defined by $\mathrm{Ad}(g)(A) = gAg^{-1}$.

We consider the adjoint representation

$$\mathrm{Sp}(1) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{sp}(1))$$

of the compact symplectic group $\mathrm{Sp}(1)$ on its Lie algebra $\mathfrak{sp}(1)$ of purely imaginary quaternions, or equivalently, the adjoint representation

$$SU(2) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{su}(2))$$

of the special unitary group $SU(2)$ on its Lie algebra $\mathfrak{su}(2)$ given by the real vector space of complex 2×2-matrices that are skew-hermitian and traceless. The map that to $v \in \mathfrak{sp}(1)$ assigns $|v| = \sqrt{v^* v}$ is a norm on the real vector space $\mathfrak{sp}(1)$, and it determines a real inner product $\langle -, - \rangle$ on $\mathfrak{sp}(1)$ given by[36]

$$\langle v, w \rangle = \frac{1}{2}(|v+w|^2 - |v|^2 - |w|^2).$$

We claim that the adjoint representation takes values in the subgroup

$$SO(\mathfrak{sp}(1)) \subset \mathrm{GL}(\mathfrak{sp}(1))$$

of linear isometries with respect to $\langle -, - \rangle$ that have determinant 1. To see this, we first note that since $\mathrm{Ad}(q)(v) = qvq^{-1} = qvq^*$, we have

$$(qvq^*)^* qvq^* = qv^* q^* qvq^* = qv^* vq^* = v^* v,$$

where the last identity holds, because $v^* v$ is an element of the center $\mathbb{R}$ of $\mathbb{H}$. This shows that $\mathrm{Ad}(q)$ is a linear isometry with respect to $\langle -, - \rangle$. Therefore, the adjoint representation induces a group homomorphism

$$\mathrm{Sp}(1) \xrightarrow{\mathrm{Ad}} O(\mathfrak{sp}(1))$$

to the subgroup $O(\mathfrak{sp}(1)) \subset \mathrm{GL}(\mathfrak{sp}(1))$ of linear isometric isomorphisms. It is clearly a continuous map, and since $\mathrm{Sp}(1)$ is connected, its image is fully contained in one of the two components of $O(\mathfrak{sp}(1))$. But $\mathrm{Ad}(1)$ is the identity map of $\mathfrak{sp}(1)$, which has determinant 1, so we conclude that $\mathrm{Ad}(q)$ takes values in $SO(\mathfrak{sp}(1))$ as claimed.

Theorem 12.21. *The adjoint representation induces a group homomorphism*

$$\mathrm{Sp}(1) \xrightarrow{\mathrm{Ad}} SO(\mathfrak{sp}(1))$$

which is surjective with kernel $\{\pm 1\}$.

We first prove two lemmas. If V is a real vector space with norm $\|-\|$, then we write $S(V) = \{v \in V \mid \|v\| = 1\} \subset V$ for the unit sphere.

Lemma 12.22. *If* $H \subset SO(\mathfrak{sp}(1))$ *is a subgroup such that the restriction to* H *of the standard action by* $SO(\mathfrak{sp}(1))$ *on* $S(\mathfrak{sp}(1))$ *is transitive and such that there exists* $u \in S(\mathfrak{sp}(1))$ *with* $SO(\mathfrak{sp}(1))_u \subset H$*, then* $H = SO(\mathfrak{sp}(1))$.

Proof. Given $g \in SO(\mathfrak{sp}(1))$, we can find $h \in H$ such that $h \cdot u = g \cdot u$. But then $h^{-1} g \cdot u = u$, so $h^{-1} g \in SO(\mathfrak{sp}(1))_u \subset H$, and hence, $g = h \cdot h^{-1} g \in H$. □

Lemma 12.23. *For all* $v \in S(\mathfrak{sp}(1))$*, there exists* $g \in \mathrm{Sp}(1)$ *such that*

$$\mathrm{Ad}(g)(v) = i.$$

Proof. We will use the spectral theorem for normal operators on finite dimensional complex vector spaces. The ring homomorphism $f \colon \mathbb{H} \to M_2(\mathbb{C})$ that we considered above induces isomorphisms $h \colon \mathrm{Sp}(1) \to SU(2)$ and $h' \colon \mathfrak{sp}(1) \to \mathfrak{su}(2)$. It maps $v \in \mathfrak{sp}(1)$ to $X = h'(v) \in \mathfrak{su}(2)$ with $\det(X) = v^* v = 1$. Since the matrix X is skew-hermitian, it is normal.[37] Therefore, by the spectral theorem for normal matrices, there exists $P \in U(2)$ such that $PXP^{-1} = \mathrm{diag}(\lambda_1, \lambda_2)$, where λ_1 and

[36] Writing $v = ib + jc + kd$, we have $|v|^2 = b^2 + c^2 + d^2$.

[37] Indeed, $X^* X = (-X)X = X(-X) = XX^*$.

λ_2 are the eigenvalues of X. Since X is skew-hermitian and $\det(X) = 1$, one shows that $\lambda_1 = i$ and $\lambda_2 = -i$. So we have $P \in U(2)$ with

$$PXP^{-1} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = h'(i).$$

Since $P \in U(2)$, we have $\det(P) \in U(1)$, so we can choose $z \in U(1)$ such that $z^2 = \det(P)$. Then $U = z^{-1}P \in SU(2)$, and we still have $UXU^{-1} = h'(i)$. Hence, if $g \in \mathrm{Sp}(1)$ is the unique element with $h(g) = U$, then $\mathrm{Ad}(g)(v) = i$. □

Proof of Theorem 12.21. We apply Lemma 12.22 to the subgroup $H \subset SO(\mathfrak{sp}(1))$ given by the image of $\mathrm{Ad}\colon \mathrm{Sp}(1) \to SO(\mathfrak{sp}(1))$. Lemma 12.23 shows that H acts transitively on $S(\mathfrak{sp}(1))$, and we proceed to show that for $SO(\mathfrak{sp}(1))_i \subset H$. The matrix that represents a general element of the isotropy subgroup $SO(\mathfrak{sp}(1))_i$ with respect to the basis (i, j, k) of $\mathfrak{sp}(1)$ has the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

for some $\theta \in \mathbb{R}$. We calculate that the matrix that represent $\mathrm{Ad}(e^{it})$ with respect to the basis (i, j, k) of $\mathfrak{sp}(1)$ is given by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2t & -\sin 2t \\ 0 & \sin 2t & \cos 2t \end{pmatrix}.$$

This shows that $SO(\mathfrak{sp}(1))_i \subset H$, and therefore, we conclude from Lemma 12.22 that $H = SO(\mathfrak{sp}(1))$ as stated.

Finally, we show that $\ker(\mathrm{Ad}) = \{\pm 1\}$. If we write $g = a + ib + jc + kd$, then the identity $\mathrm{Ad}(g)(i) = gig^* = i$ implies that $a^2 - b^2 - c^2 - d^2 = 1$. Since we also have $a^2 + b^2 + c^2 + d^2 = 1$, we conclude that $a = \pm 1$ and $b = c = d = 0$, as desired. □

Corollary 12.24. *The map induced by the adjoint representation,*

$$\mathrm{Sp}(1)/\{\pm 1\} \xrightarrow{\overline{\mathrm{Ad}}} SO(\mathfrak{sp}(1)),$$

is an isomorphism of topological groups.

Proof. We have not explicitly specified the topologies on these groups before, so we do that now. We have identified both $\mathrm{Sp}(1)$ and $SO(\mathfrak{sp}(1))$ with subsets of $M_2(\mathbb{C})$, and we give both the respective subspace topologies induced from the metric topology on $M_2(\mathbb{C})$. Finally, we give $\mathrm{Sp}(1)/\{\pm 1\}$ the quotient topology induced from the topology on $\mathrm{Sp}(1)$. As a topological space, $\mathrm{Sp}(1)/\{\pm 1\}$ is compact, because $\mathrm{Sp}(1)$ is compact, and $SO(\mathfrak{sp}(1))$ is Hausdorff, because the metric topology on $M_2(\mathbb{C})$ is Hausdorff. So it suffices to show that $\overline{\mathrm{Ad}}$ is a group homomorphism and a continuous bijection. Theorem 12.21 shows that it is a group isomorphism, so it only remains to show that the map $\overline{\mathrm{Ad}}$ is continuous. By the universal property of the quotient topology, the map $\overline{\mathrm{Ad}}$ is continuous if and only if the map Ad is continuous. And by the universal property of the subspace topology, the map Ad is continuous if and only if the map

$$\mathrm{Sp}(1) \xrightarrow{\widetilde{\mathrm{Ad}}} \mathrm{End}_{\mathbb{R}}(M_2(\mathbb{C}))$$

defined by $\widetilde{\mathrm{Ad}}(g)(X) = h(g)Xh(g)^{-1}$ is continuous. This, in turn, follows from the definition of matrix multiplication and from Cramer's formula for the inverse of a matrix. $\square$

If G is a topological group, then we write $\mathrm{Rep}_{\mathbb{C}}(G)$ for the category, whose objects are complex representations (V, π) of G such that $\pi\colon G \to \mathrm{GL}(V)$ is continuous, and whose morphisms are intertwining $\mathbb{C}$-linear maps. Restriction along the continuous group homomorphism $\mathrm{Ad}\colon \mathrm{Sp}(1) \to SO(\mathfrak{sp}(1))$ defines a functor

$$\mathrm{Rep}_{\mathbb{C}}(SO(\mathfrak{sp}(1))) \xrightarrow{\mathrm{Ad}^*} \mathrm{Rep}_{\mathbb{C}}(\mathrm{Sp}(1)),$$

and Corollary 12.24 shows that this functor is a fully faithful embedding and that its essential image are the continuous complex representations (V, π) of $\mathrm{Sp}(1)$ with the property that $\pi(-1) = \mathrm{id}_V$.

Another consequence of Corollary 12.24 is that, as a topological space, $SO(3)$ is homeomorphic to the real projective space $\mathbb{P}^3(\mathbb{R})$. Indeed, as a topological space $\mathrm{Sp}(1)$ is homeomorphic to S^3, and the action of the subgroup $\{\pm 1\} \subset \mathrm{Sp}(1)$ by left multiplication is free.

We say that a topological group G is a compact group if its underlying space is compact and Hausdorff. The classical groups are all compact topological groups in this sense. It turns out that the theory of continuous finite dimensional complex representations of compact groups is completely analogous to the theory of finite dimensional complex represenations of finite groups, except that there will typically be a countably infinite number of non-isomorphic irreducible such representations. We first define the generalization to compact groups of the regular representation for finite groups. It will a representation on a complex Hilbert space $L^2(G)$, the definition of which requires some input from analysis, which we will assume.

One can show that there exists a Borel measure μ on G that is both left-invariant and right-invariant in the sense that for every Borel subset $A \subset G$ and $g \in G$,[38]

$$\mu(g \cdot A) = \mu(A) = \mu(A \cdot g),$$

and regular in the sense that for every Borel subset $A \subset G$,

$$\mu(A) = \inf\{\mu(U) \mid A \subset U,\, U \subset G \text{ open}\} = \sup\{\mu(K) \mid K \subset A \text{ compact}\}.$$

Moreover, such a measure, which is called a Haar measure, is unique up to scaling. In particular, there exists a unique Haar measure on G that is a probability measure in the sense that $\mu(G) = 1$.

Let $C^0(G, \mathbb{R})$ to be the (right) real vector space given by the set consisting of all continuous functions $\varphi \colon G \to \mathbb{R}$ equipped with pointwise vector sum and pointwise scalar multiplication. Given a Haar measure μ on G, we define a linear map

$$C^0(G, \mathbb{R}) \xrightarrow{\;I\;} \mathbb{R}$$

as follows. Given $\varphi \in C^0(G, \mathbb{R})$, we choose a real number $0 < d < 1$ and define

$$A_{n,r}(\varphi) = \{x \in G \mid nd^r \leq \varphi(x) < (n+1)d^r\} \subset G,$$

for all integers n and positive integers r. Since $\varphi \colon G \to \mathbb{R}$ is continuous and G compact, the subset $\varphi(G) \subset \mathbb{R}$ is compact and therefore bounded. It follows that for every positive integer r, the subset $A_{n,r}(\varphi) \subset G$ is non-empty for only finitely many integers n. It is a Borel subset, and hence, we may form the sum

$$\textstyle\sum_{n \in \mathbb{Z}} nd^r \mu(A_{n,r}(\varphi)) \in \mathbb{R}.$$

One may show that the limit

$$I(\varphi) = \lim_{r \to \infty} \textstyle\sum_{n \in \mathbb{Z}} nd^r \mu(A_{n,r}(\varphi)) \in \mathbb{R}$$

exists and is independent of the choice of $0 < d < 1$. Finally, one may show that the function $I \colon C^0(G, \mathbb{R}) \to \mathbb{R}$ defined in this way is indeed linear.

Similarly, let $C^0(G, \mathbb{C})$ be the (right) complex vector space given by the set of all continuous complex functions $\varphi \colon G \to \mathbb{C}$ equipped with pointwise vector sum and scalar multiplication. Let $f \colon \mathbb{R} \to \mathbb{C}$ be the canonical inclusion. Then we have the map of right real vector spaces

$$C^0(G, \mathbb{R}) \longrightarrow f_* C^0(G, \mathbb{C})$$

[38] More generally, if G is locally compact, then there exists a left-invariant, but not necessarily right-invariant, measure on G.

that to $\varphi\colon G \to \mathbb{R}$ assigns $f \circ \varphi\colon G \to \mathbb{C}$, and its adjunct map

$$f^*C^0(G,\mathbb{R}) = C^0(G,\mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C} \longrightarrow C^0(G,\mathbb{C})$$

is an isomorphism of complex vector spaces. Hence, we obtain a $\mathbb{C}$-linear map

$$C^0(G,\mathbb{C}) \xrightarrow{I_{\mathbb{C}}} \mathbb{C}$$

defined to be the adjunct of the composite $\mathbb{R}$-linear map

$$C^0(G,\mathbb{R}) \xrightarrow{I} \mathbb{R} \xrightarrow{f} f_*\mathbb{C}.$$

We will only consider $\mathbb{C}$-valued continuous functions on G, so we will abbreviate and write $C^0(G)$ instead of $C^0(G,\mathbb{C})$ and $I(\varphi)$ or $\int_G f(x)d\mu(x)$ instead of $I_{\mathbb{C}}(\varphi)$.

Given $\varphi,\psi \in C^0(G)$, we define $\varphi\psi \in C^0(G)$ to be the pointwise product of φ and ψ, and we define φ^* to be the pointwise complex conjugate of φ. Since the map $I\colon C^0(G) \to \mathbb{C}$ is $\mathbb{C}$-linear, it follows immediately that the map

$$C^0(G) \times C^0(G) \xrightarrow{\langle -,-\rangle} \mathbb{C}$$

defined by $\langle \varphi,\psi\rangle = I(\varphi^*\psi)$ is a hermitian form. Moreover, this map is a hermitian inner product. Indeed, if $\varphi \in C^0(G)$ and $\langle \varphi,\varphi\rangle = I(|\varphi|^2) = 0$, then $\varphi = 0$.

If $(V,\langle -,-\rangle)$ is a complex vector space with hermitian inner product, then the inner product gives rise to a metric $d\colon V \times V \to \mathbb{R}_{\geq 0}$ defined by

$$d(v,w) = \sqrt{\langle v-w, v-w\rangle},$$

and we say that $(V,\langle -,-\rangle)$ is a Hilbert space if the metric space (V,d) is complete.[39] If both $(U,\langle -,-\rangle_U)$ and $(V,\langle -,-\rangle_V)$ are complex vector spaces with hermitian inner products, then we say that a linear map $f\colon V \to U$ is Cauchy-continuous if for every sequence $v\colon \mathbb{Z}_{\geq 0} \to V$ that is Cauchy with respect to d_V, the sequence $f \circ v\colon \mathbb{Z}_{\geq 0} \to U$ is Cauchy with respect to d_U.[40] Let $\mathrm{Herm}_{\mathbb{C}}$ be the category, whose objects are the complex vector spaces with hermitian inner products, and whose morphisms are the Cauchy-continuous linear maps between these, and let $\mathrm{Hilb}_{\mathbb{C}}$ be the full subcategory of Hilbert spaces. In this situation, there is an adjunction

$$\mathrm{Herm}_{\mathbb{C}} \underset{i_\wedge}{\overset{i^\wedge}{\rightleftarrows}} \mathrm{Hilb}_{\mathbb{C}},$$

where the right adjoint functor $i_\wedge$ is the canonical inclusion, and where the left adjoint functor $i^\wedge$ takes a complex vector space with hermitian inner product $(U,\langle -,-\rangle_U)$ to a Hilbert space $(V,\langle -,-\rangle_V)$ such that the underlying metric space (V,d_V) is the completion of the metric space (U,d_U). The unit map

$$(U,\langle -,-\rangle_U) \xrightarrow{\eta} (\widehat{U},\langle -,-\rangle_{\widehat{U}}) = (i_\wedge \circ i^\wedge)(U,\langle -,-\rangle_U)$$

is injective and its image $\eta(U) \subset \widehat{U}$ is a dense subset of the metric space $(\widehat{U},d_{\widehat{U}})$. In the following, we will omit the hermitian inner products from the notation.

[39] This means that every sequence in V that is Cauchy with respect to d converges with respect to d. A sequence $v\colon \mathbb{Z}_{\geq 0} \to V$ is Cauchy with respect to d, if for all $\epsilon > 0$, there exists $N \in \mathbb{Z}_{\geq 0}$ such that $d(v_i,v_j) < \epsilon$, for all $i,j \geq N$, and it converges with respect to d, if there exists $v \in V$ such that for all $\epsilon > 0$, there exists $N \in \mathbb{Z}_{\geq 0}$ such that $d(v,v_i) < \epsilon$, for all $i \geq N$.

[40] Every Cauchy-continuous map between two metric spaces is continuous, and every continuous map between two complete metric spaces is Cauchy-continuous.

We return to the complex vector space with hermitian inner product $C^0(G)$. It is not a Hilbert space, unless G is finite, so we define the Hilbert space

$$L^2(G) = \widehat{C^0(G)}$$

to be its completion. As just explained the unit map

$$C^0(G) \xrightarrow{\eta} L^2(G)$$

is injective and its image is dense in $L^2(G)$. Hence, every element of $L^2(G)$ can be written, non-canonically, as a limit of a Cauchy sequence of continuous $\mathbb{C}$-valued functions on G, but a general element of $L^2(G)$ is not a $\mathbb{C}$-valued function on G, unless G is finite. In particular, the value "$f(x)$" of $f \in L^2(G)$ at $x \in G$ is not meaningful.[41] We will see below that the Hilbert space $L^2(G)$ is separable in the sense that it admits a countably dimensional dense subspace.

Lemma 13.1. *The map $I\colon C^0(G) \to \mathbb{C}$ is Cauchy-continuous.*

Proof. We must show that if the sequence $\varphi\colon \mathbb{Z}_{\geq 0} \to C^0(G)$ is Cauchy, then so is the sequence $I \circ \varphi\colon \mathbb{Z}_{\geq 0} \to \mathbb{C}$. It suffices to show that for all $\varphi, \psi \in C^0(G)$,

$$|I(\varphi) - I(\psi)| = |I(\varphi - \psi)| \leq I(|\varphi - \psi|),$$

which follows immediately from the definition of $I\colon C^0(G) \to \mathbb{C}$. □

Since $\mathbb{C}$ is complete, we conclude that $I\colon C^0(G) \to \mathbb{C}$ extends uniquely to a continuous, or equivalently, Cauchy-continuous linear map

$$L^2(G) \xrightarrow{I} \mathbb{C}.$$

Example 13.2. If G is a finite group, which we consider as a compact topological group with the discrete topology, then the Haar probability measure on G is given by the normalized counting measure that to $A \subset G$ assigns $\mu(A) = |A|/|G|$. It follows that the corresponding integral $I\colon C^0(G) \to \mathbb{C}$ is given by

$$I(f) = |G|^{-1} \textstyle\sum_{x \in G} f(x),$$

so we find that $L^2(G) = C^0(G) = \mathbb{C}[G]$.

We wish to extend the definition of the two-sided regular representation from finite groups to compact groups. So let G be a compact topological group. Given $(g_1, g_2) \in G \times G$ and $\varphi \in C^0(G)$, the formula

$$\operatorname{Reg}(g_1, g_2)(\varphi)(x) = \varphi(g_2^{-1} x g_1)$$

defines an element $\operatorname{Reg}(g_1, g_2)(\varphi) \in C^0(G)$. Moreover, since a Haar measure on G is both left-invariant and right-invariant, the map

$$C^0(G) \xrightarrow{\operatorname{Reg}(g_1, g_2)} C^0(G)$$

is a linear isometry with respect to $\langle -, - \rangle$. Indeed, we have

$$\|\operatorname{Reg}(g_1, g_2)(\varphi)\|^2 = \textstyle\int_G |\varphi(g_2^{-1} x g_1)|^2 d\mu(x) = \int_G |\varphi(x)|^2 d\mu(x) = \|\varphi\|^2.$$

[41] The linear map $\mathrm{ev}_x\colon C^0(G) \to \mathbb{C}$ defined by $\mathrm{ev}_x(\varphi) = \varphi(x)$ is not Cauchy-continuous, and hence, does not extend to a map $\mathrm{ev}_x\colon L^2(G) \to \mathbb{C}$. However, it is possible to identify $L^2(G)$ with the quotient of the complex vector space consisting of the functions $f\colon G \to \mathbb{C}$ that are Haar measurable and square-integrable by the subspace of functions that are zero almost everywhere.

In particular, it is Cauchy-continuous, and therefore, it induces a map

$$L^2(G) \xrightarrow{\text{Reg}(g_1,g_2)} L^2(G)$$

which is a linear isometry with inverse $\text{Reg}(g_1^{-1}, g_2^{-1})$. This defines a map

$$G \times G \xrightarrow{\text{Reg}} U(L^2(G))$$

to the group of linear isometric isomorphisms of $L^2(G)$.[42] We wish to say that this is a map of topological groups, so we much define a topology on $U(L^2(G))$ and show that the map is continuous. It turns out that the appropriate topology on $U(L^2(G))$ is the so-called strong operator topology.[43]

Proposition 13.3. *The two-sided regular representation*

$$G \times G \xrightarrow{\text{Reg}} U(L^2(G))$$

is continuous with respect to the strong operator topology.

Proof. The strong operator topology has the property that the map Reg in question is continuous if and only if for every $\varphi \in L^2(G)$, the composite map

$$G \times G \xrightarrow{\text{Reg}} U(L^2(G)) \xrightarrow{\text{ev}_\varphi} L^2(G)$$

is continuous. Let us write Reg_φ for this map. Since $G \times G$ is a topological group, it suffices to prove that this map is continuous at $(g_1, g_2) = (e, e)$.

We first let $\varphi \in C^0(G)$ and prove that Reg_φ is continuous at (e, e). We have

$$\|\text{Reg}_\varphi(g_1, g_2) - \text{Reg}_\varphi(e, e)\|^2 = \textstyle\int_G |\varphi(g_2^{-1}xg_1) - \varphi(x)|^2 d\mu(x)$$

and wish to prove that this quantity goes to 0 as $(g_1, g_2) \to (e, e)$. Since both φ and multiplication and inversion in G are continuous, we have every $x \in G$,

$$\lim_{(g_1,g_2)\to(e,e)} |\varphi(g_2^{-1}xg_1) - \varphi(x)|^2 = 0.$$

Moreover, for all $x \in G$, the integrand is dominated by

$$|\varphi(g_2^{-1}xg_1) - \varphi(x)|^2 \leq 4 \cdot \sup\{|\varphi(h)| \mid h \in G\},$$

so by the dominated convergence theorem for the integral, we conclude that

$$\lim_{(g_1,g_2)\to(e,e)} \textstyle\int_G |\varphi(g_2^{-1}xg_1) - \varphi(x)|^2 d\mu(x) = 0$$

as desired.

We next prove that for any $\varphi \in L^2(G)$, the map Reg_φ is continuous at (e, e). Given $\epsilon > 0$, we choose $\varphi_\epsilon \in C^0(G)$ such that $\|\varphi - \varphi_\epsilon\| < \epsilon$, which is possible,

[42] Traditionally, linear isometric isomorphisms of a Hilbert space $\mathfrak{h}$ are called unitary operators, and therefore, we write $U(\mathfrak{h})$ for the group consisting of these operators.

[43] The uniform operator topology, which is given by the operator norm, is stronger than the strong operator topology. It turns out that it is too strong for our purposes, since, even for $G = U(1)$, the map Reg is not continuous with respect to this topology.

because $C^0(G)$ is dense in $L^2(G)$. Now

$$\begin{aligned}\|\mathrm{Reg}_\varphi(g_1,g_2)-\varphi\| &\le \|\mathrm{Reg}_\varphi(g_1,g_2)-\mathrm{Reg}_{\varphi_\epsilon}(g_1,g_2)\| \\ &\qquad + \|\mathrm{Reg}_{\varphi_\epsilon}(g_1,g_2)-\varphi_\epsilon\| + \|\varphi_\epsilon-\varphi\| \\ &= 2\|\varphi-\varphi_\epsilon\| + \|\mathrm{Reg}_{\varphi_\epsilon}(g_1,g_2)-\varphi_\epsilon\| \\ &< 2\epsilon + \|\mathrm{Reg}_{\varphi_\epsilon}(g_1,g_2)-\varphi_\epsilon\|,\end{aligned}$$

and by the first case, there exists an open neighborhood $(e,e)\in U\subset G\times G$ such that $\|\mathrm{Reg}_{\varphi_\epsilon}(g_1,g_2)-\varphi_\epsilon\|<\epsilon$, for all $(g_1,g_2)\in U$. So we conclude that

$$\|\mathrm{Reg}_\varphi(g_1,g_2)-\varphi\|<3\epsilon,$$

for all $(g_1,g_2)\in U$. This proves that Reg_φ is continuous at (e,e). □

If (V,π) is a finite dimensional complex representation of G, then we define the associated space of matrix coefficients $M(\pi)$ to be the image of the map

$$V\otimes V^* \xrightarrow{\mu_\pi} C^0(G)\subset L^2(G)$$

defined by $\mu_\pi(v\otimes h)(g)=h(\pi(g)(v))$. One verifies immediately that it intertwines between $\pi\boxtimes\pi^*$ and Reg, so that we obtain a map

$$\pi\boxtimes\pi^* \xrightarrow{\mu_\pi} \mathrm{Reg}_{M(\pi)}$$

of continuous representations of $G\times G$. It is an isomorphism, if π is an irreducible representation of G, because then $\pi\boxtimes\pi^*$ is an irreducible representation of $G\times G$.

Lemma 13.4. *Let G be a compact topological group, let π_1 and π_2 be irreducible finite dimensional complex representations of G, and let $M(\pi_1),M(\pi_2)\subset L^2(G)$ be their subspaces of matrix coefficients.*

(1) *If $\pi_1\simeq\pi_2$, then $M(\pi_1)=M(\pi_2)$.*
(2) *If $\pi_1\not\simeq\pi_2$, then $M(\pi_1)\perp M(\pi_2)$.*

Proof. To prove (1), we let V_1 and V_2 be the representation spaces of π_1 and π_2, respectively, and let $h\colon V_1\to V_2$ be a linear isomorphism that is intertwining between π_1 and π_2. In this situation, the diagram

$$\begin{array}{ccc} V_1\otimes V_2^* & \xrightarrow{\mathrm{id}\otimes h^*} & V_1\otimes V_1^* \\ \downarrow{\scriptstyle h\otimes\mathrm{id}} & & \downarrow{\scriptstyle \mu_{\pi_1}} \\ V_2\otimes V_2^* & \xrightarrow{\mu_{\pi_2}} & L^2(G)\end{array}$$

commutes, and therefore,

$$M(\pi_1)=\mathrm{im}(\mu_{\pi_1})=\mathrm{im}(\mu_{\pi_1}\circ(\mathrm{id}\otimes h^*))=\mathrm{im}(\mu_{\pi_2}\circ(h\otimes\mathrm{id}))=\mathrm{im}(\mu_{\pi_2})=M(\pi_2).$$

To prove (2), we consider the composition

$$M(\pi_1)\xrightarrow{i} L^2(G)\xrightarrow{p} M(\pi_2)$$

of the canonical inclusion of $M(\pi_1)$ and the orthogonal projection onto $M(\pi_2)$. The map i is intertwining between $\mathrm{Reg}_{M(\pi_1)}$ and Reg, since $M(\pi_1)$ is a Reg-invariant subspace, and the map p is intertwining between Reg and $\mathrm{Reg}_{M(\pi_2)}$, since Reg is a unitary representation. Therefore, the composite map $p\circ i$ is intertwining between

$\mathrm{Reg}_{M(\pi_1)}$ and $\mathrm{Reg}_{M(\pi_2)}$, which are non-isomorphic irreducible finite dimensional complex representations of $G \times G$, so by Schur's lemma, $p \circ i = 0$ as stated. □

The theorem of Peter and Weyl states that if G is a compact topological group, then the two-sided regular representation of $G \times G$ decomposes as the completed direct sum of the spaces of matrix coefficients, one for each isomorphism class of irreducible finite dimensional continuous complex representations of G.

Theorem 13.5 (Peter–Weyl)**.** *Let G be a compact topological group, and let $\widehat{G}$ be the set of isomorphism classes of finite dimensional complex representations of G. For every $\sigma \in \widehat{G}$, let (V_σ, π_σ) be a representative of the class σ. The map*

$$\widehat{\bigoplus}_{\sigma \in \widehat{G}} \pi_\sigma \boxtimes \pi_\sigma^* \xrightarrow{\mu} \mathrm{Reg},$$

whose σth component is given by $\mu_{\pi_\sigma}(v \otimes h)(g) = h(\pi_\sigma(g)(v))$, is an isomorphism of continuous representations of $G \times G$.

Proof. We will only prove the theorem for compact groups G that admit a faithful continuous representation $\rho \colon G \to \mathrm{GL}_n(\mathbb{C})$; for a proof in the general case, we refer to [4, Theorem 5.4.1]. By Lemma 13.4, the canonical map

$$\bigoplus_{\sigma \in \widehat{G}} M(\pi_\sigma) \longrightarrow C^0(G)$$

is injective, and we proceed to prove that its image is dense with respect to the L^2-norm. To this end, we let $a_{ij} = \mu_\rho(e_j \otimes e_i^*) \in C^0(G)$ be the matrix coefficients of $\rho \colon G \to \mathrm{GL}_n(\mathbb{C})$ and consider the sub-$\mathbb{C}$-algebra $\mathbb{C}[G] \subset C^0(G)$ given by the image of the unique $\mathbb{C}$-algebra homomorphism

$$\mathbb{C}[X_{ij}, Y_{ij} \mid 1 \leq i, j \leq n] \longrightarrow C^0(G)$$

that to X_{ij} and $Y_{i,j}$ assign a_{ij} and a_{ij}^*. We claim that $\mathbb{C}[G] \subset C^0(G)$ is dense with respect to the L^2-norm. Indeed, by the Stone–Weierstrass theorem, $\mathbb{C}[G] \subset C^0(G)$ is dense with respect to the supremum norm $\|-\|_\infty$, and since G has finite volume $\mu(G)$, the calculation

$$\|\varphi\|_2^2 = \textstyle\int_G |\varphi(x)|^2 d\mu(x) \leq \int_G \|\varphi\|_\infty^2 d\mu(x) = \|\varphi\|_\infty^2 \mu(G)$$

shows that $\mathbb{C}[G] \subset C^0(G)$ is also dense with respect to the L^2-norm.

Now, for all $m \geq 0$, we consider the finite dimensional subspace

$$\mathrm{Fil}_m \mathbb{C}[G] \subset \mathbb{C}[G]$$

given by the image by the $\mathbb{C}$-algebra homomorphism

$$\mathbb{C}[X_{ij}, Y_{i,j} \mid 1 \leq i, j \leq n] \longrightarrow C^0(G)$$

of the subspace of polynomials of degree $\leq m$. It is Reg-invariant, since the matrix coefficients a_{ij} transform linearly under left and right translation on G, and

$$\textstyle\bigcup_{m \geq 0} \mathrm{Fil}_m \mathbb{C}[G] = \mathbb{C}[G].$$

We consider the representation $R_m \colon G \to \mathrm{GL}(\mathrm{Fil}_m \mathbb{C}[G])$ given by the restriction of the right regular representation of G on $L^2(G)$ to this subspace. Since it is finite dimensional, it decomposes as a direct sum of irreducible finite dimensional representations of G, so by Lemma 13.4, the inclusion $M(R_m) \to C^0(G)$ factors as

$$M(R_m) \longrightarrow \bigoplus_{\sigma \in \widehat{G}} M(\pi_\sigma) \longrightarrow C^0(G).$$

We define $\epsilon \colon C^0(G) \to \mathbb{C}$ to be the linear map given by $\epsilon(\varphi) = \varphi(e)$ and consider the map $\nu_m \colon \operatorname{Fil}_m \mathbb{C}[G] \to M(R_m)$ given by $\nu_m(\varphi) = \mu_{R_m}(\varphi \otimes \epsilon)$. The calculation

$$\nu_m(\varphi)(g) = \mu_{R_m}(\varphi \otimes \epsilon)(g) = \epsilon(R_m(g)(\varphi)) = R_m(g)(\varphi)(e) = \varphi(e \cdot g) = \varphi(g)$$

shows that the composite map

$$\operatorname{Fil}_m \mathbb{C}[G] \xrightarrow{\nu_m} M(R_m) \longrightarrow \bigoplus_{\sigma \in \widehat{G}} M(\pi_\sigma) \longrightarrow C^0(G)$$

is equal to the canonical inclusion, and hence, the canonical inclusion of $\mathbb{C}[G]$ into $C^0(G)$ factors as a composition

$$\mathbb{C}[G] = \bigcup_{m \geq 0} \operatorname{Fil}_m \mathbb{C}[G] \longrightarrow \bigoplus_{\sigma \in \widehat{G}} M(\pi_\sigma) \longrightarrow C^0(G).$$

Since the image of the composite map is dense with respect to the L^2-norm, so is the image of the right-hand map. This completes the proof. □

Remark 13.6. Let G be a linear compact topological group, let $\rho \colon G \to \mathrm{GL}_n(\mathbb{C})$ be a faithful continuous representation, and let $\mathbb{C}[G] \subset C^0(G)$ be the subalgebra of polynomial functions on G defined in the proof of Theorem 13.5. We claim that

$$\mathbb{C}[G] = \bigoplus_{\sigma \in \widehat{G}} M(\pi_\sigma) \subset C^0(G).$$

For otherwise, there exists $\tau \in \widehat{G}$ such that $M(\pi_\tau) \not\subset \mathbb{C}[G]$, and since $\mathbb{C}[G]$ is a direct sum of irreducible finite dimensional representations, it follows from Lemma 13.4 that $M(\pi_\tau) \perp \mathbb{C}[G]$. But this contradicts the fact that $\mathbb{C}[G] \subset \bigoplus_{\sigma \in \widehat{G}} M(\pi_\sigma)$ is dense.

Remark 13.7. In general, a unitary representation of a topological group G is defined to be a pair $(\mathfrak{h}, \pi)$ of a Hilbert space $\mathfrak{h}$ and a continuous group homomorphism

$$G \xrightarrow{\pi} U(\mathfrak{h})$$

from G to the group $U(\mathfrak{h})$ of linear isometric isomorphisms of $\mathfrak{h}$ equipped with the strong operator topology. As a consequence of the Peter–Weyl theorem, one can show that for G compact, every such representation admits a finite dimensional π-invariant subspace $V \subset \mathfrak{h}$; for a proof, see [4, p. 301]. Hence, every irreducible unitary representation of a compact topological group G is finite dimensional. By contrast, locally compact topological groups such as $G = \mathrm{GL}_n(\mathbb{C})$ that are not compact have irreducible unitary representations that are infinite dimensional.

Example 13.8. We let $G = U(1)$ and let $\tau \colon G \to \mathrm{GL}(V)$ be the standard representation on $V = \mathbb{C}$. For every $n \geq 0$, we have the representation

$$\tau_n = \operatorname{Sym}^n_{\mathbb{C}}(\tau)$$

of G on $\operatorname{Sym}^n_{\mathbb{C}}(V)$. It is an irreducible representation, because the complex vector space $\operatorname{Sym}^n_{\mathbb{C}}(V)$ is 1-dimensional. Let (e_1) be the standard basis of V so that (e_1^n) is a basis of $\operatorname{Sym}^n_{\mathbb{C}}(V)$. Then for $z \in G$, we have

$$\tau_n(z)(e_1^n) = (e_1 z)^n = e_1^n z^n.$$

The dual representation $\tau_{-n} = \tau_n^*$ is also 1-dimensional, and hence, irreducible, and

$$\tau_{-n}(z)((e_1^*)^n) = ((e_1 z)^*)^n = (e_1^*)^n z^{-n}.$$

So for all $m, n \in \mathbb{Z}$, we have $\tau_m \simeq \tau_n$ if and only if $m = n$. Up to isomorphism, these are all irreducible finite dimensional continuous complex representations of G. Hence, by the Peter–Weyl theorem, the map of unitary $G \times G$-representations

$$\widehat{\bigoplus}_{n \in \mathbb{Z}} \tau_n \boxtimes \tau_n^* \xrightarrow{\mu} \mathrm{Reg}$$

is an isomorphism.

Example 13.9. Let $G = SU(2)$ and let $\pi\colon G \to \mathrm{GL}(V)$ be the standard representation on $V = \mathbb{C}^2$. For every $n \geq 0$, we have the representation

$$\pi_n = \mathrm{Sym}^n_{\mathbb{C}}(\pi)$$

of G on the $(n+1)$-dimensional complex vector space $\mathrm{Sym}^n_{\mathbb{C}}(V)$. Let (e_1, e_2) be the standard basis of V so that $(e_1^{n-i} e_2^i \mid 0 \leq i \leq n)$ is a basis of $\mathrm{Sym}^n_{\mathbb{C}}(V)$. We let $f\colon U(1) \to SU(2)$ be the group homomorphism defined by $f(z) = \mathrm{diag}(z, z^{-1})$ and consider the representation $f^*(\pi_n)$ of $U(1)$. For $z \in U(1)$, the calculation

$$\pi_n(f(z))(e_1^{n-i} e_2^i) = (e_1 z)^{n-i}(e_2 z^{-1})^i = e_1^{n-i} e_2^i z^{n-2i}$$

shows that the $\mathbb{C}$-linear isomorphism

$$\bigoplus_{0 \leq i \leq n} \mathrm{Sym}^{n-2i}_{\mathbb{C}}(\mathbb{C}) \xrightarrow{h} \mathrm{Sym}^n_{\mathbb{C}}(V),$$

whose ith component is given by $h_i(v_i^{n-2i}) = e_1^{n-i} e_2^i v_i^{n-2i}$, is intertwining with respect to $\bigoplus_{0 \leq i \leq n} \tau_{n-2i}$ and $f^*(\pi_n)$. Therefore, every $f^*(\pi_n)$-invariant subspace of $\mathrm{Sym}^n_{\mathbb{C}}(V)$ is of the form $W = h(\bigoplus_{i \in S} \mathrm{Sym}^{n-2i}_{\mathbb{C}}(\mathbb{C}))$ with $S \subset \{0, 1, \dots, n\}$. In particular, if $x = \sum_{0 \leq i \leq n} e_1^{n-i} e_2^i x_i \in W$ and $x_i \neq 0$, then $e_1^{n-i} e_2^i \in W$.

If $W \subset \mathrm{Sym}^n_{\mathbb{C}}(V)$ is a non-zero π_n-invariant subspace, then W is in particular an $f^*(\pi_n)$-invariant subspace. Hence, there exists $0 \leq i \leq n$ such that $e_1^{n-i} e_2^i \in W$. We now consider

$$g = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in G$$

and first calculate

$$g \cdot e_1^{n-i} e_2^i = e_1^{n-i}(e_1 + e_2)^i = e_1^n + \sum_{0 < j \leq i} \binom{i}{j} e_1^{n-i-j} e_2^j,$$

which shows that $e_1^n \in W$, and next calculate

$$g^* \cdot e_1^n = (e_1 + e_2)^n = \sum_{0 \leq j \leq n} \binom{n}{j} e_1^{n-j} e_2^j,$$

which shows that $e_1^{n-j} e_2^j \in W$ for all $0 \leq j \leq n$. Therefore, $W = \mathrm{Sym}^n_{\mathbb{C}}(V)$, and hence, π_n is irreducible. We will show later that, up to isomorphism, these are all irreducible finite dimensional continuous complex representations of G. Hence, by the Peter–Weyl theorem, the map of unitary $G \times G$-representations

$$\widehat{\bigoplus}_{n \in \mathbb{Z}_{\geq 0}} \pi_n \boxtimes \pi_n^* \xrightarrow{\mu} \mathrm{Reg}$$

is an isomorphism.

Example 13.10. Let $G = SO(\mathfrak{su}(2)) \simeq SO(3)$. We recall from last time that restriction along the adjoint representation

$$SU(2) \xrightarrow{\text{Ad}} SO(\mathfrak{su}(2))$$

defines an equivalence of categories from $\operatorname{Rep}_{\mathbb{C}}(SO(\mathfrak{su}(2))$ onto the full subcategory of $\operatorname{Rep}_{\mathbb{C}}(SU(2))$ that is spanned by the representations (V, π) of $SU(2)$ for which $\pi(-I) = \text{id}_V$. Now, for the representation π_n defined in Example 13.9, we have

$$\pi_n(-I)(e_1^{n-i}e_2^i) = (-e_1)^{n-i}(-e_2)^i = (-1)^n e_1^{n-i}e_2^i.$$

So there exists $\bar{\pi}_n \in \operatorname{Rep}_{\mathbb{C}}(SO(\mathfrak{su}(2))$ such that $\pi_n \simeq \text{Ad}^*(\bar{\pi}_n) \in \operatorname{Rep}_{\mathbb{C}}(SU(2))$ if and only if $n = 2m$ is even. Therefore, by the Peter–Weyl theorem, we conclude that the map of unitary $G \times G$-representations

$$\widehat{\bigoplus}_{m \in \mathbb{Z}_{\geq 0}} \bar{\pi}_{2m} \boxtimes \bar{\pi}_{2m}^* \xrightarrow{\mu} \text{Reg}$$

is an isomorphism.

Appendix: Tensors

Let k be a field and V a vector space.[44] The tensor algebra of V is defined to be the graded associative k-algebra given by the graded k-vector space

$$T_k(V) = \bigoplus_{n \geq 0} T_k^n(V),$$

where $T_k(V) = V^{\otimes_k n}$, equipped with the multiplication given by

$$(x_1 \otimes \cdots \otimes x_m) \cdot (y_1 \otimes \cdots \otimes y_n) = x_1 \otimes \cdots \otimes x_m \otimes y_1 \otimes \cdots \otimes y_n.$$

The symmetric algebra of V is defined to be the graded commutative k-algebra given by the quotient

$$\operatorname{Sym}_k(V) = \bigoplus_{n \geq 0} \operatorname{Sym}_k^n(V) = T_k(V)/I$$

of the tensor algebra of V by the graded two-sided ideal $I \subset T_k(V)$ generated by the family $(x \otimes y - y \otimes x \mid x, y \in V)$, and the exterior algebra of V is defined to be the graded anticommutative k-algebra given by the quotient

$$\Lambda_k(V) = \bigoplus_{n \geq 0} \Lambda_k^n(V) = T_k(V)/J$$

of the tensor algebra of V by the graded two-sided ideal $J \subset T_k(V)$ generated by the family $(x \otimes x \mid x \in V)$. If $f \colon V \to U$ is a k-linear map, then the map

$$T_k^n(V) \xrightarrow{T_k^n(f)} T_k^n(U)$$

that to $x_1 \otimes \cdots \otimes x_n$ assigns $f(x_1) \otimes \cdots \otimes f(x_n)$ is k-linear and induce maps

$$\operatorname{Sym}_k^n(V) \xrightarrow{\operatorname{Sym}_k^n(f)} \operatorname{Sym}_k^n(U) \qquad \Lambda_k^n(V) \xrightarrow{\Lambda_k^n(f)} \Lambda_k^n(U)$$

that also are k-linear. This makes $T_k^n(-)$, $\operatorname{Sym}_k^n(-)$, and $\Lambda_k^n(-)$ functors from the category of k-vector spaces and k-linear maps to itself.

[44] We only use that k is a commutative ring and that V is a k-module. It is important, however, that k be commutative, so $k = \mathbb{H}$ is not an option.

In particular, if $\pi\colon G \to \mathrm{GL}(V)$ is a representation of a group G on a k-vector space V, then the composite map

$$G \xrightarrow{\ \pi\ } \mathrm{GL}(V) \xrightarrow{\mathrm{Sym}^n_k} \mathrm{GL}(\mathrm{Sym}^n_k(V))$$

is a representation of G on the k-vector space $\mathrm{Sym}^n_k(V)$, which we, by abuse of notation, denote by $\mathrm{Sym}^n_k(\pi)$. Similarly, we define k-linear representations $T^n_k(\pi)$ and $\Lambda^n_k(\pi)$ on $T^n_k(V)$ and $\Lambda^n_k(V)$.

We denote the classes of $v_1 \otimes \cdots \otimes v_n \in T^n_k(V)$ in $\mathrm{Sym}^n_k(V)$ and $\Lambda^n_k(V)$ by $v_1 \dots v_n$ and $v_1 \wedge \cdots \wedge v_n$, respectively. If $\sigma \in \Sigma_n$ is a permutation, then we

$$v_{\sigma(1)} \dots v_{\sigma(n)} = v_1 \dots v_n \in \mathrm{Sym}^n_k(V)$$

and

$$v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)} = \mathrm{sgn}(\sigma) v_1 \wedge \cdots \wedge v_n \in \Lambda^n_k(V).$$

These statements both follow immediately from the definitions. However, it is a non-trivial theorem that if $(e_i)_{i \in I}$ is a basis of V then the family

$$(e_{i_1} \otimes \cdots \otimes e_{i_n} \mid i_1, \dots, i_n \in I)$$

is a basis of $T^n_k(V)$, and that if we choose a total order "$\leq$" on I, then

$$(e_{i_1} \dots e_{i_n} \mid i_1, \dots, i_n \in I, i_1 \leq \cdots \leq i_n)$$

is a basis of $\mathrm{Sym}^n_k(V)$, and

$$(e_{i_1} \wedge \cdots \wedge e_{i_n} \mid i_1, \dots, i_n \in I, i_1 < \cdots < i_n)$$

is a basis of $\Lambda^n_k(V)$. For instance, if $\dim_k(V) = d$ and $(e_1, \dots, e_d)$ is a basis V, then the fact that $\dim_k(\Lambda^d_k(V)) = 1$ with basis $e_1 \wedge \cdots \wedge e_d$ is equivalent to the existence of the determinant.

We recall that a topological group is defined to be a group $G = (G, \mu, \iota)$ together with a topology on the set G such that the maps $\mu\colon G \times G \to G$ and $\iota\colon G \to G$ are continuous. Similarly, a Lie group is defined to be a group $G = (G, \mu, \iota)$ together with a structure of smooth manifold on the set G such that the maps $\mu\colon G \times G \to G$ and $\iota\colon G \to G$ are smooth. We first discuss smooth manifolds.

Smooth manifolds belong to geometry rather than topology. Geometric objects are pairs $(X, \mathcal{O}_X)$ of a topological space X and a sheaf of rings $\mathcal{O}_X$ on X, where for $U \subset X$ open, the set $\Gamma(U, \mathcal{O}_X)$ of sections of $\mathcal{O}_X$ over U should be thought of as the set "geometric functions" on U. The geometric functions that we allow will depend on the geometric situation that we consider. For instance, we could consider "smooth functions," "analytic functions," or "algebraic functions," but note that we have not yet assigned any precise mathematical meaning to these terms. Moreover, in some situations, the elements of $\Gamma(U, \mathcal{O}_X)$ may not be functions in the usual sense. A map of geometric objects $f\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ is a pair $f = (p, \phi)$ of a continuous map $p\colon Y \to X$ and a map of sheaves of rings $\phi\colon \mathcal{O}_X \to p_*\mathcal{O}_Y$. Let us now define sheaves properly.

Let X be a topological space, and let X_{Zar} be the category, whose objects are the open subsets $U \subset X$, and whose morphisms are

$$\mathrm{Hom}_{X_{\mathrm{Zar}}}(U, V) = \begin{cases} \{\mathrm{incl}_U^V\} & \text{if } U \subset V \\ \emptyset & \text{if } U \not\subset V. \end{cases}$$

So if $U \subset V$, then there is a unique morphism $\mathrm{incl}_U^V\colon U \to V$, and if $U \not\subset V$, then there are no morphisms from U to V. A presheaf of sets on X is a defined to be a functor $\mathcal{F}\colon X_{\mathrm{Zar}}^{\mathrm{op}} \to \mathrm{Set}$. To specify a functor $\mathcal{F}\colon X_{\mathrm{Zar}}^{\mathrm{op}} \to \mathrm{Set}$, we must specify for every open subset $U \subset X$, a set $\mathcal{F}(U)$, and for every inclusion $U \subset V$ of open subsets of X, a map $\mathcal{F}(\mathrm{incl}_U^V)\colon \mathcal{F}(V) \to \mathcal{F}(U)$. We may think of $F(U)$ as the set of "functions defined on U" and of $\mathcal{F}(\mathrm{incl}_U^V)$ as the map that to a "function defined on U" assigns the restriction of this function to a "function defined on V." To emphasize this interpretation, we also write $\Gamma(U, \mathcal{F}) = \mathcal{F}(U)$ and call it the set of sections of $\mathcal{F}$ over U, and we write $\mathrm{Res}_U^V = \mathcal{F}(\mathrm{incl}_U^V)$ and call it the restriction from V to U. A presheaf $\mathcal{F}\colon X_{\mathrm{Zar}}^{\mathrm{op}} \to \mathrm{Set}$ is defined to be a sheaf if it satisfies the following sheaf condition: For every covering $(U_i \to U)_{i \in I}$ of an open subset $U \subset X$ by open subsets $U_i \subset U$, the diagram

$$\mathcal{F}(U) \xrightarrow{\;h\;} \prod_{i \in I} \mathcal{F}(U_i) \overset{a}{\underset{b}{\rightrightarrows}} \prod_{(i,j) \in I \times I} \mathcal{F}(U_i \cap U_j)$$

is an equalizer. Here h is the unique map such that for all $i \in I$,

$$\mathrm{pr}_i \circ h = \mathrm{Res}_{U_i}^U,$$

and a and b are the unique maps such that for all $(i, j) \in I \times I$,

$$\mathrm{pr}_{(i,j)} \circ a = \mathrm{Res}_{U_i \cap U_j}^{U_i} \circ \mathrm{pr}_i$$
$$\mathrm{pr}_{(i,j)} \circ b = \mathrm{Res}_{U_i \cap U_j}^{U_j} \circ \mathrm{pr}_j .$$

That the diagram is an equalizer means that for all $(\varphi_i)_{i \in I} \subset \prod_{i \in I} \mathcal{F}(U_i)$ such that

$$a((\varphi_i)_{i \in I}) = b((\varphi_i)_{i \in I}),$$

there exists a unique $\varphi \in \mathcal{F}(U)$ such that

$$(\varphi_i)_{i\in I} = (\mathrm{Res}^U_{U_i}(\varphi))_{i\in I}.$$

Informally, the sheaf condition expresses that if we are given "functions" φ_i on U_i for all $i \in I$ such that $\varphi_i|_{U_i\cap U_j} = \varphi_j|_{U_i\cap U_j}$ for all $(i,j) \in I \times I$, then there exists a unique "function" φ on U such that $\varphi_i = \varphi|_{U_i}$ for all $i \in I$.

Example 14.1. Let X be a topological space, and let $k = \mathbb{R}$ or $k = \mathbb{C}$. The presheaf $\mathcal{O}^{\mathrm{cont}}_X \colon X^{\mathrm{op}}_{\mathrm{Zar}} \to \mathrm{Set}$, where $\mathcal{O}^{\mathrm{cont}}_X(U)$ is the set of continuous functions $\varphi\colon U \to k$, and where $\mathrm{Res}^V_U \colon \mathcal{O}^{\mathrm{cont}}_X(V) \to \mathcal{O}^{\mathrm{cont}}_X(U)$ is the map defined by $\mathrm{Res}^V_U(\varphi) = \varphi \circ \mathrm{incl}^V_U$, is a sheaf, because "being continuous" is a local property.

We define the category of presheaves of sets on X to be the category

$$\mathcal{P}(X) = \mathrm{Fun}(X^{\mathrm{op}}_{\mathrm{Zar}}, \mathrm{Set}),$$

whose objects are functors and whose morphisms are natural transformations, and we define the category of sheaves on X to be the full subcategory

$$\mathrm{Shv}(X) \subset \mathcal{P}(X)$$

spanned by the sheaves on X. One can prove that there is an adjunction

$$\mathcal{P}(X) \underset{\iota_X}{\overset{L_X}{\rightleftarrows}} \mathrm{Shv}(X)$$

where the right adjoint functor ι_X is the canonical inclusion of the subcategory of sheaves in the category of presheaves, and where the left adjoint functor L_X takes a presheaf to its associated sheaf. The functor L_X is called "sheafification."

Example 14.2. Let X be a topological space, and let $\mathcal{F} \in \mathcal{P}(X)$ be the presheaf of constant functions, $\mathcal{F}(U) = \{\varphi\colon U \to k \mid \varphi \text{ constant}\}$. It is not a sheaf, since "being constant" is not a local property. The associated sheaf $L_X(\mathcal{F}) \in \mathrm{Shv}(X)$ is the sheaf of locally constant functions, $L_X(\mathcal{F})(U) = \{\varphi\colon U \to \mathbb{R} \mid \varphi \text{ locally constant}\}$.

It is a fundamental result of Grothendieck[45] that "sheafification" preserves finite limits. (The inclusion functor ι_X preserves all limits, as does every right adjoint functor.) In particular, it preserves finite products, which implies that it takes "presheaves of rings" to "sheaves of rings." Indeed, we define a presheaves of rings and sheaves of rings to be ring objects in $\mathcal{P}(X)$ and $\mathrm{Shv}(X)$, respectively. A ring object in a category $\mathcal{C}$ with finite products is defined to be a sixtuple $(R, +, \cdot, -, 0, 1)$ of an object $R \in \mathcal{C}$, two morphisms $+, \cdot \colon R \times R \to R$, one morphism $-\colon R \to R$, and two morphisms $0, 1\colon e \to R$ that satisfy the usual ring axioms. Here the empty product $e = R^0 \in \mathcal{C}$ is a terminal object.

Let $p\colon Y \to X$ be a continous map. If $U \subset X$ is open, then $p^{-1}(U) \subset Y$ is open, so we obtain a functor $u = p^{-1} \colon X_{\mathrm{Zar}} \to Y_{\mathrm{Zar}}$. The functor

$$\mathcal{P}(Y) \xrightarrow{u^*} \mathcal{P}(X)$$

given by restriction along u has a left adjoint functor

$$\mathcal{P}(X) \xrightarrow{u_!} \mathcal{P}(Y)$$

[45] This result and many results are consequences of Grothendieck's theorem that, in the category of sets, filtered colimits and finite limits commute.

given by left Kan extension along u. More concretely, we have

$$u_!(\mathcal{F})(V) = \operatorname{colim}_{p(V)\subset U} \mathcal{F}(U),$$

where the colimit is indexed by the opposite of the "slice category"

$$(X_{\text{Zar}})_{V/}$$

with objects open subsets $U \subset X$ such that $V \subset p^{-1}(U)$ and with morphisms inclusions among such open subsets. It is a cofiltered category, so $u_!$ preserves finite limits by Grothendieck's theorem. The functor u^* preserves sheaves in the sense that there is a unique functor p_* making the diagram

$$\begin{array}{ccc} \operatorname{Shv}(Y) & \xrightarrow{\;p_*\;} & \operatorname{Shv}(X) \\ \big\downarrow{\scriptstyle \iota_X} & & \big\downarrow{\scriptstyle \iota_Y} \\ \mathcal{P}(Y) & \xrightarrow{\;u^*\;} & \mathcal{P}(X) \end{array}$$

commute, but the functor $u_!$ does not. However, the functor

$$\operatorname{Shv}(X) \xrightarrow{\;p^*\;} \operatorname{Shv}(Y)$$

defined by $p^* = L_Y \circ u_! \circ \iota_X$ is left adjoint of p_*. We call p^* the inverse image functor and we call p_* the direct image functor. So we have an adjunction

$$\operatorname{Shv}(X) \underset{p_*}{\overset{p^*}{\rightleftarrows}} \operatorname{Shv}(Y)$$

and the functor p^* preserves finite limits. In particular, it preserves ring objects.

Example 14.3. (1) Let $j\colon U \to X$ be the inclusion of an open subset. It is an open map in the sense that if $V \subset U$ is open, then so is $V = j(V) \subset X$. Therefore, the slice category $(X_{\text{Zar}})_{V/}$ has $j(V) \subset X$ as its initial object, which, in turn, implies that $j^*(\mathcal{F})(V) = \mathcal{F}(j(V))$. In this situation, we also write $\mathcal{F}|_U = j^*(\mathcal{F})$.

(2) Let $i_x\colon \{x\} \to X$ be the inclusion of a point and note that $\operatorname{Shv}(\{x\}) \simeq \text{Set}$. Indeed, a presheaf $\mathcal{G}\colon \{x\}_{\text{Zar}} \to \text{Set}$ is a sheaf if and only if $\mathcal{G}(\emptyset)$ is a one-element set, so, up to unique isomorphism, a sheaf $\mathcal{G} \in \operatorname{Shv}(\{x\})$ is determined by the set $\mathcal{G}(\{x\})$. We say that $\mathcal{F}_x = i_x^*(\mathcal{F})(\{x\})$ is the stalk of $\mathcal{F} \in \operatorname{Shv}(X)$ at $x \in X$. Concretely, we have $\mathcal{F}_x = \operatorname{colim}_{x\in U} \mathcal{F}(U)$, where the colimit is indexed by the opposite of the category of open neighborhoods $x \in U \subset X$ under inclusion. One can prove that a morphism $h\colon \mathcal{F} \to \mathcal{F}'$ in $\operatorname{Shv}(X)$ is an isomorphism if and only if the induced map of stalks $h_x\colon \mathcal{F}_x \to \mathcal{F}'_x$ is an isomorphism for all $x \in X$.[46]

The sheaf $\mathcal{O}_X^{\text{cont}}$ of continuous k-valued functions on X is a sheaf of commutative rings, and therefore, its stalk $\mathcal{O}_{X,x}^{\text{cont}}$ at $x \in X$ is a commutative ring.

Lemma 14.4. *For every $x \in X$, $\mathcal{O}_{X,x}^{\text{cont}}$ is a local ring.*

Proof. The elements $h \in \mathcal{O}_{X,x}^{\text{cont}}$ are germs of continuous k-valued functions at $x \in X$, that is, equivalence classes of pairs (U, φ) of an open neighborhood $x \in U \subset X$ and a continuous function $\varphi\colon U \to k$, where two such pairs (U_1, φ_1) and (U_2, φ_2) are equivalent, if there exists $x \in V \subset U_1 \cap U_2$ open such that $\varphi_1|_V = \varphi_2|_V$. The map $\psi_x\colon \mathcal{O}_{X,x}^{\text{cont}} \to k$ that to the class of (U, φ) assigns $\varphi(x)$ is a surjective ring

[46] We refer to this statement by saying that the Zariski topos $\operatorname{Shv}(X)$ has "enough points."

homomorphism to a field, so its kernel $\mathfrak{m}_x \subset \mathcal{O}_{X,x}^{\mathrm{cont}}$ is a maximal ideal. Now, if $h \in \mathcal{O}_{X,x}^{\mathrm{cont}}$ and $h \notin \mathfrak{m}_x$, then we can represent h by a pair (U, φ) such that $\varphi(u) \neq 0$ for all $u \in U$. This shows that h is invertible with h^{-1} given by the class of the pair (U, ψ), where $\psi(u) = \varphi(u)^{-1}$. This proves the lemma. □

Let $p\colon Y \to X$ be a continuous map. We define the map

$$\mathcal{O}_X^{\mathrm{cont}} \xrightarrow{\phi} p_*\mathcal{O}_Y^{\mathrm{cont}}$$

of sheaves of rings on X as follows. If $U \subset X$ is open with $V = p^{-1}(U) \subset Y$, then

$$\mathcal{O}_X^{\mathrm{cont}}(U) \xrightarrow{\phi_U} (p_*\mathcal{O}_Y^{\mathrm{cont}})(U) = \mathcal{O}_Y^{\mathrm{cont}}(V)$$

is the ring homomorphism that to $\varphi\colon U \to k$ continuous assigns $\varphi \circ p|_V \colon V \to k$. By adjunction, it determines and is determined by a map

$$p^*\mathcal{O}_X^{\mathrm{cont}} \xrightarrow{\tilde{\phi}} \mathcal{O}_Y^{\mathrm{cont}}$$

of sheaves of rings on Y. We will abuse notation and write also ϕ instead of $\tilde{\phi}$ for this map. The induced map of stalks at $y \in Y$ is a ring homomorphism

$$\mathcal{O}_{X,x}^{\mathrm{cont}} = i_x^*\mathcal{O}_X^{\mathrm{cont}} \simeq (p \circ i_y)^*\mathcal{O}_X^{\mathrm{cont}} \simeq i_y^*p^*\mathcal{O}_X^{\mathrm{cont}} \xrightarrow{\phi_y} i_y^*\mathcal{O}_Y = \mathcal{O}_{Y,y}^{\mathrm{cont}},$$

where the indicated isomorphisms are the unique natural isomorphisms between different choices of left adjoint functors of the functor $i_{x*} = (p \circ i_y)_*$.

Lemma 14.5. *The ring homomorphism $\phi_y\colon \mathcal{O}_{X,x}^{\mathrm{cont}} \to \mathcal{O}_{Y,y}^{\mathrm{cont}}$ is a local homomorphism.*

Proof. That ϕ_y is a local homomorphism means that it is a ring homomorphism and that $(\phi_y)^{-1}(\mathfrak{m}_y) = \mathfrak{m}_x$, or equivalently, that the following diagram commutes.

$$\begin{array}{ccc} \mathcal{O}_{X,x}^{\mathrm{cont}} & \xrightarrow{\psi_x} & k \\ \downarrow{\scriptstyle \phi_y} & & \downarrow{\scriptstyle \mathrm{id}} \\ \mathcal{O}_{Y,y}^{\mathrm{cont}} & \xrightarrow{\psi_y} & k \end{array}$$

Now, if $h \in \mathcal{O}_{X,x}^{\mathrm{cont}}$ is represented by the pair (U, φ), where $x \in U \subset X$ is an open neighborhood and $\varphi\colon U \to k$ is a continuous map, then $y \in V = p^{-1}(U) \subset Y$ is an open neighborhood, and the pair $(V, \varphi \circ p|_V)$ represents $p_y(h) \in \mathcal{O}_{Y,y}$. So

$$\psi_y(\phi_y(h)) = (\varphi \circ p|_V)(y) = \varphi(p(y)) = \varphi(x) = \psi_x(h),$$

as desired. □

We will consider other kinds of "functions," but we always want them to retain the properties that we proved in Lemmas 14.4 and 14.5 for continuous functions. We encode these properties in the following definition.

Definition 14.6. (1) A locally ringed space is a pair $(X, \mathcal{O}_X)$ of a topological space X and a sheaf of rings $\mathcal{O}_X$ such that for all $x \in X$, the stalk $\mathcal{O}_{X,x}$ is a local ring.

(2) A map of locally ringed spaces is a pair $f = (p, \phi)\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ of a continuous map $p\colon Y \to X$ and a map $\phi\colon \mathcal{O}_X \to p_*\mathcal{O}_Y$ of sheaves of rings on X

with the property that given $y \in Y$ with image $x = p(y) \in X$, the induced map of stalks $\phi_y \colon \mathcal{O}_{X,x} \to \mathcal{O}_{Y,y}$ is a local ring homomorphism.

If $(X, \mathcal{O}_X)$ is a locally ringed space, if $U \subset X$ is open, and if $\varphi \in \mathcal{O}_X(U)$, then we define its value $\varphi(x)$ at $x \in U$ to be the image of φ by the composite map

$$\mathcal{O}_X(U) \xrightarrow{i_U} \mathcal{O}_{X,x} \xrightarrow{\psi_x} k(x) = \mathcal{O}_{X,x}/\mathfrak{m}_x.$$

We note that the value $\varphi(x) \in k(x)$ is an element of a field $k(x)$ that may vary with $x \in U$. It may also happen that $\varphi \neq 0$ even though $\varphi(x) = 0$ for all $x \in U$.

We now define the "geometric functions" relevant for smooth manifolds, namely, the smooth functions. However, our discussion below applies mutatis mutandis to holomorphic functions and complex manifolds and to analytic functions and real analytic manifolds. Let $U \subset \mathbb{R}^n$ be an open subset. A function $\varphi \colon U \to \mathbb{R}$ is defined to be smooth if the partial derivatives $\partial^k\varphi/\partial x_{i_1} \dots \partial x_{i_k} \colon U \to \mathbb{R}$ exist and are continuous for all $k \geq 0$ and $1 \leq i_1, \dots, i_k \leq n$. The sheaf of standard smooth functions on U is defined to be the subsheaf $\mathcal{O}_U^{\mathrm{sm}} \subset \mathcal{O}_U^{\mathrm{cont}}$ given by

$$\Gamma(V, \mathcal{O}_U^{\mathrm{sm}}) = \{\varphi \colon V \to \mathbb{R} \mid \varphi \text{ smooth}\} \subset \Gamma(V, \mathcal{O}_U^{\mathrm{cont}})$$

for all $V \subset U$ open. We say that a locally ringed space $(X, \mathcal{O}_X)$ is an affine smooth manifold, if there exists an isomorphism of locally ringed spaces

$$(X, \mathcal{O}_X) \xrightarrow{f=(p,\phi)} (U, \mathcal{O}_U^{\mathrm{sm}})$$

with $U \subset \mathbb{R}^n$ open. The number n is uniquely determined by $(X, \mathcal{O}_X)$ and is called the dimension of the affine smooth manifold.

Definition 14.7. A smooth manifold[47] is a locally ringed space $(X, \mathcal{O}_X)$ for which there exists an open covering $(U_i \to X)_{i \in I}$ such that for all $i \in I$, $(U_i, \mathcal{O}_X|_{U_i})$ is an affine smooth manifold. A morphism $f \colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ between smooth manifolds is a morphism of locally ringed spaces.

Remark 14.8. (1) If $(X, \mathcal{O}_X)$ is a smooth manifold, then $\mathcal{O}_X$ is canonically isomorphic to a subsheaf of $\mathcal{O}_X^{\mathrm{cont}}$. Indeed, by definition, this is true locally, so by the sheaf condition, it is also true globally. Moreover, if $f = (p, \phi) \colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ is a morphism between smooth manifolds, then the diagram

$$\begin{array}{ccc} \mathcal{O}_X & \xrightarrow{\phi} & p_*\mathcal{O}_Y \\ \big\downarrow & & \big\downarrow \\ \mathcal{O}_X^{\mathrm{cont}} & \xrightarrow{\phi} & p_*\mathcal{O}_Y^{\mathrm{cont}} \end{array}$$

commutes, and therefore, the top horizontal map is uniquely determined by the bottom horizontal map. Therefore, we may view "being smooth" as the property of the continuous map $p \colon Y \to X$ that a map $\phi \colon \mathcal{O}_X \to p_*\mathcal{O}_Y$ making the diagram commute exist. In particular, the map ϕ is uniquely determined by the map p. An isomorphism between smooth manifolds is traditionally called a diffeomorphism.

[47] In the literature, the requirement that X be Hausdorff is often included in the definition of a smooth manifold, but we will not do so. Note that "being Hausdorff" is not a local property.

(2) We define the dimension of a smooth manifold $(X, \mathcal{O}_X)$ to be the map

$$X \xrightarrow{\dim} \mathbb{Z}_{\geq 0}$$

that to $x \in X$ assigns $n = \dim(x)$, if there exists $x \in U \subset X$ open with $(U, \mathcal{O}_X|_U)$ an affine smooth manifold of dimension n. It is well-defined and locally constant, and if it is constant with value n, then we say that $(X, \mathcal{O}_X)$ has pure dimension n or that $(X, \mathcal{O}_X)$ is a smooth n-manifold. We define a chart of $(X, \mathcal{O}_X)$ around $x \in X$ to be a pair (U, h) of an open neighborhood $x \in U \subset X$ and a diffeomorphism

$$(U, \mathcal{O}_X|_U) \xrightarrow{h} (V, \mathcal{O}_V^{\mathrm{sm}})$$

with $V \subset \mathbb{R}^{\dim(x)}$ an open subset.

Proposition 14.9. *The category of smooth manifolds and their morphisms admits finite products. More precisely, if $f\colon (Z, \mathcal{O}_Z) \to (X, \mathcal{O}_Y)$ and $g\colon (Z, \mathcal{O}_Z) \to (Y, \mathcal{O}_Y)$ are morphisms between smooth manifolds, then, up to unique isomorphism, there is a unique sheaf of rings $\mathcal{O}_{X\times Y}$ on $X \times Y$ such that $(X \times Y, \mathcal{O}_{X\times Y})$ is a smooth manifold and such that, in the diagram*

$$\begin{array}{ccccc} & & (Z, \mathcal{O}_Z) & & \\ & \swarrow^{f} & \downarrow{\scriptstyle (f,g)} & {}^{g}\searrow & \\ (X, \mathcal{O}_X) & \xleftarrow{p} & (X \times Y, \mathcal{O}_{X\times Y}) & \xrightarrow{q} & (Y, \mathcal{O}_Y), \end{array}$$

the projections p and q and the unique map (f, g) that makes the diagram commute are morphisms of smooth manifolds.

Proof. Up to isomorphism, there is a unique sheaf $\mathcal{O}_{X\times Y}$ on $X \times Y$ such that given $(x, y) \in X \times Y$ and charts $h\colon (U, \mathcal{O}_X|_U) \to (A, \mathcal{O}_A^{\mathrm{sm}})$ and $k\colon (V, \mathcal{O}_Y|_V) \to (B, \mathcal{O}_B^{\mathrm{sm}})$ around $x \in X$ and $y \in Y$, respectively, the map

$$(U \times V, \mathcal{O}_{X\times Y}|_{U\times V}) \xrightarrow{h\times k} (A \times B, \mathcal{O}_{A\times B}^{\mathrm{sm}})$$

is a chart around $(x, y) \in X \times Y$.[48] Since the subsets of the form $U \times V \subset X \times Y$, where $U \subset X$ and $V \subset Y$ are open, form a basis for the product topology, this shows that $(X \times Y, \mathcal{O}_{X\times Y})$ is a smooth manifold. That the maps p, q, and (f, g) are smooth can be checked locally in charts, where it is clear. □

One can construct new smooth manifolds by gluing existing smooth manifolds together. To state the result, we introduce some terminology. In general, we define a morphism $(s, t)\colon R \to Y \times Y$ in a category $\mathcal{C}$ that admits finite products to be an equivalence relation if for all $Z \in \mathcal{C}$, the induced map of sets

$$\mathrm{Map}(Z, R) \xrightarrow{(s,t)} \mathrm{Map}(Z, Y) \times \mathrm{Map}(Z, Y)$$

exhibits $\mathrm{Map}(Z, R)$ as an equivalence relation on the set $\mathrm{Map}(Z, Y)$ in the usual sense. In particular, the morphism (s, t) is a monomorphism.

A map $f\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ of smooth manifolds is étale if there exists an open covering $(V_i \to Y)_{i\in I}$ such that each $f|_{V_i}\colon (V_i, \mathcal{O}_Y|_{V_i}) \to (f(V_i), \mathcal{O}_X|_{f(V_i)})$ is

[48] There is a canonical map $\mathcal{O}_X \otimes_k \mathcal{O}_Y \to \mathcal{O}_{X\times Y}$ of sheaves of k-algebras on $X \times Y$, but it is not an isomorphism. Rather the target is a suitable completion of the source.

a diffeomorphism. It is an open immersion if, in addition, the map $f\colon Y \to X$ is injective. The image $f(Y) \subset X$ of an étale morphism is an open subset.

Proposition 14.10. *Given an equivalence relation of smooth manifolds*

$$(R, \mathcal{O}_R) \xrightarrow{(s,t)} (Y, \mathcal{O}_Y) \times (Y, \mathcal{O}_Y)$$

such that $Y = \coprod_{i \in I} Y_i$ *and* $R = \coprod_{(i,j) \in I \times I} U_{i,j}$ *and such that* s *and* t *restrict to open immersions* $s|_{U_{i,j}} \colon U_{i,j} \to Y_i$ *and* $t|_{U_{i,j}} \colon U_{i,j} \to Y_j$*, the coequalizer*

$$(R, \mathcal{O}_R) \underset{t}{\overset{s}{\rightrightarrows}} (Y, \mathcal{O}_Y) \xrightarrow{f} (X, \mathcal{O}_X)$$

exists. Moreover, the morphism f *is étale.*

Proof. Suppose that $s = (g, \sigma)$ and $t = (h, \tau)$. We let $X = Y/R$ with the quotient topology, and let $p\colon Y \to X$ be the canonical projection. It is the coequalizer of $g, h\colon R \to Y$ in the category of topological spaces and continuous maps. We claim that for all $i \in I$, the map $p|_{Y_i}\colon Y_i \to f(Y_i)$ is a homeomorphism. First, it is a bijection, since the maps $g|_{U_{i,i}}\colon U_{i,i} \to Y_i$ and $h_{U_{i,i}}\colon U_{i,i} \to Y_i$ necessarily are equal. Indeed, they are both open immersions and the diagonal map $\Delta\colon Y_i \to Y_i \times Y_i$ factors through $(g,h)|_{U_{i,i}}\colon U_{i,i} \to Y_i \times Y_i$, since (s,t) is an equivalence relation. Second, it is an open map. Indeed, if $V \subset Y_i$ is an open subset, then so is the subset

$$p^{-1}(p(V)) = \textstyle\coprod_{j \in I} (h \circ g^{-1})(V \cap U_{i,j}) \subset \coprod_{j \in I} Y_j = Y.$$

This shows that $p|_{Y_i}\colon Y_i \to p(Y_i) \subset X$ is a homeomorphism.

Finally, the sheaf of rings $\mathcal{O}_X$ given by the equalizer

$$\mathcal{O}_X \xrightarrow{\phi} p_*\mathcal{O}_Y \underset{p_*(\tau)}{\overset{p_*(\sigma)}{\rightrightarrows}} q_*\mathcal{O}_R,$$

where $q = p \circ g = p \circ h$, makes $(X, \mathcal{O}_X)$ a smooth manifold and makes the diagram in the statement a coequalizer in the category of smooth manifolds and morphisms of smooth manifolds. ∎

Remark 14.11. The morphisms $s, t\colon (R, \mathcal{O}_R) \to (Y, \mathcal{O}_Y)$ in Proposition 14.10 are étale, but they are a very particular kind of étale morphisms. We would like the result to hold more generally for every étale equivalence relation, that is, for every equivalence relation $(s,t)\colon (R, \mathcal{O}_R) \to (Y \times Y, \mathcal{O}_{Y \times Y})$ such that s and t are étale, but this is not true.[49] To remedy this, one builds the larger category of smooth stacks in which the result holds for every étale equivalence relation.

Example 14.12. (1) Let $\mathbb{A}^1_k = (\mathbb{R}, \mathcal{O}^{\mathrm{sm}}_{\mathbb{R}})$ be the affine line, let $\mathbb{A}^1_k \smallsetminus \{0\} \subset \mathbb{A}^1_k$ be the open complement of $\{0\} \subset \mathbb{A}^1_k$, and let (s,t) be the equivalence relation with

$$R = R_{11} \sqcup R_{12} \sqcup R_{22} = \mathbb{A}^1_k \sqcup (\mathbb{A}^1_k \smallsetminus \{0\}) \sqcup \mathbb{A}^1_k \underset{t}{\overset{s}{\rightrightarrows}} Y = Y_1 \sqcup Y_2 = \mathbb{A}^1_k \sqcup \mathbb{A}^1_k$$

where the maps $s, t\colon R_{12} \to Y_1$ are defined to be the canonical inclusion and the map $t \mapsto t^{-1}$, respectively. The coequalizer $(X, \mathcal{O}_X)$ is the projective line $\mathbb{P}^1_k$.

[49] A counterexample is $(s,t)\colon \mathbb{Z} \times S^1 \to S^1 \times S^1$, where $s(n,z) = z$ and $t(n,z) = w^n z$, where w is some fixed irrational roation of the circle S^1.

(2) We consider the equivalence relation defined as in (1), except that we now define both $s, t\colon R_{12} \to Y_1$ to be the canonical inclusion. The coequalizer $(X, \mathcal{O}_X)$ is an affine line with a double point at the origin. The space X is not Hausdorff.

We will use Proposition 14.10 to construct the tangent bundle of a smooth manifold. It is a functor that to a smooth manifold $(X, \mathcal{O}_X)$ assigns a morphism

$$T(X, \mathcal{O}_X) = (TX, \mathcal{O}_{TX}) \xrightarrow{p_X} (X, \mathcal{O}_X)$$

of smooth manifolds together with a structure of real vector space on the fiber

$$T(X, \mathcal{O}_X)_x = p_X^{-1}(x) \subset T(X, \mathcal{O}_X)$$

for all $x \in X$, and that to a morphism $f\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ of smooth manifolds assigns a commutative diagram[50] of morphisms of smooth manifolds

$$\begin{array}{ccc} T(Y, \mathcal{O}_Y) & \xrightarrow{df} & T(X, \mathcal{O}_X) \\ \downarrow{\scriptstyle p_Y} & & \downarrow{\scriptstyle p_X} \\ (Y, \mathcal{O}_Y) & \xrightarrow{f} & (X, \mathcal{O}_X) \end{array}$$

such that for all $y \in Y$ with image $x = f(y) \in X$, the induced map of fibers

$$T(Y, \mathcal{O}_Y)_y \xrightarrow{df_y} T(X, \mathcal{O}_X)_x$$

is linear. The "chain rule" is the statement that this assignment is a functor.

First, if $U \subset \mathbb{R}^m$ is an open subset, then we define

$$T(U, \mathcal{O}_U^{\mathrm{sm}}) = (U \times \mathbb{R}^m, \mathcal{O}_{U \times \mathbb{R}^m}^{\mathrm{sm}}) \xrightarrow{p_U} (U, \mathcal{O}_U^{\mathrm{sm}})$$

to the projection on the first factor. We define the structure of real vector space on the fiber $T(U, \mathcal{O}_U^{\mathrm{sm}})_x$ by $(x, \boldsymbol{v}) + (x, \boldsymbol{w}) = (x, \boldsymbol{v} + \boldsymbol{w})$ and $(x, \boldsymbol{v}) \cdot a = (x, \boldsymbol{v} \cdot a)$, where $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^m$ and $a \in \mathbb{R}$. If $f\colon (V, \mathcal{O}_V^{\mathrm{sm}}) \to (U, \mathcal{O}_U^{\mathrm{sm}})$ is a morphism of smooth manifolds with $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ open, then we define

$$T(V, \mathcal{O}_V^{\mathrm{sm}}) \xrightarrow{df} T(U, \mathcal{O}_U^{\mathrm{sm}})$$

to be the morphism of smooth manifolds defined by

$$df(y, \boldsymbol{v}) = (f(y), D_{\boldsymbol{v}} f(y)),$$

where $y \in V$ and $\boldsymbol{v} \in \mathbb{R}^n$, and where

$$D_{\boldsymbol{v}} f(y) = \lim_{h \to 0} (f(y + \boldsymbol{v}h) - f(y))/h$$

is the directional derivative. If $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ and $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ are the standard bases for $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, and if we write $f(y) = \sum_{i=1}^m \boldsymbol{e}_i f_i(y)$, then

$$D_{\boldsymbol{e}_j} f(y) = \sum_{i=1}^m \boldsymbol{e}_i \cdot (\partial f_i / \partial y_j)(y).$$

[50] This diagram is cartesian, if f is étale, but not in general.

It follows that the diagram

$$\begin{array}{ccc} T(V, \mathcal{O}_V^{\mathrm{sm}}) & \xrightarrow{df} & T(U, \mathcal{O}_U^{\mathrm{sm}}) \\ \big\downarrow{\scriptstyle p_V} & & \big\downarrow{\scriptstyle p_U} \\ (V, \mathcal{O}_V^{\mathrm{sm}}) & \xrightarrow{f} & (U, \mathcal{O}_U^{\mathrm{sm}}) \end{array}$$

commutes, and that for all $y \in V$ with image $x = f(y) \in U$, the induced map

$$T(V, \mathcal{O}_V^{\mathrm{sm}})_y \xrightarrow{df_y} T(U, \mathcal{O}_U^{\mathrm{sm}})_x$$

is linear. Moreover, the chain rule from calculus shows that

$$d(f \circ g) = df \circ dg$$

for all composable morphisms of smooth manifolds

$$(W, \mathcal{O}_W^{\mathrm{sm}}) \xrightarrow{g} (V, \mathcal{O}_V^{\mathrm{sm}}) \xrightarrow{f} (U, \mathcal{O}_U^{\mathrm{sm}})$$

with $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$, and $W \subset \mathbb{R}^p$ open.

Second, given any smooth manifold $(X, \mathcal{O}_X)$, we let $(Y_i, h_i \colon Y_i \to V_i)_{i \in I}$ be a family of charts with $V_i \subset \mathbb{R}^{n_i}$. The canonical map

$$(Y, \mathcal{O}_Y) = \textstyle\coprod_{i \in I} (Y_i, \mathcal{O}_X|_{Y_i}) \xrightarrow{f} (X, \mathcal{O}_X)$$

is étale, the canonical inclusion

$$(R, \mathcal{O}_R) = (Y, \mathcal{O}_Y) \times_{(X, \mathcal{O}_X)} (Y, \mathcal{O}_Y) \xrightarrow{(s,t)} (Y, \mathcal{O}_Y) \times (Y, \mathcal{O}_Y)$$

is an equivalence relation, and the diagram

$$(R, \mathcal{O}_R) \underset{t}{\overset{s}{\rightrightarrows}} (Y, \mathcal{O}_Y) \xrightarrow{f} (X, \mathcal{O}_X)$$

is a coequalizer. We have $Y = \coprod_{i \in I} Y_i$ and $R = \coprod_{(i,j) \times I \times I} U_{i,j}$ with $U_{i,j} = Y_i \cap Y_j$, so the existence of the coequalizer also is a consequence of Proposition 14.10. We now define $p_X \colon T(X, \mathcal{O}_X) \to (X, \mathcal{O}_X)$ to be the induced morphism of coequalizers

$$\begin{array}{ccccc} T(R, \mathcal{O}_R) & \underset{dt}{\overset{ds}{\rightrightarrows}} & T(Y, \mathcal{O}_Y) & \xrightarrow{df} & T(X, \mathcal{O}_X) \\ \big\downarrow{\scriptstyle p_R} & & \big\downarrow{\scriptstyle p_Y} & & \big\downarrow{\scriptstyle p_X} \\ (R, \mathcal{O}_R) & \underset{t}{\overset{s}{\rightrightarrows}} & (Y, \mathcal{O}_Y) & \xrightarrow{f} & (X, \mathcal{O}_X), \end{array}$$

and we give the fiber $T(X, \mathcal{O}_X)_x$ the unique structure of real vector space such that for any $y \in Y$ with $f(y) = x$, the induced map of fibers

$$T(Y, \mathcal{O}_Y)_y \xrightarrow{df_y} T(X, \mathcal{O}_X)_x$$

is a linear isomorphism. To see that $p_X \colon T(X, \mathcal{O}_X) \to (X, \mathcal{O}_X)$ is well-defined, up to canonical isomorphism, one has to prove two things. First, one much show that the equivalence relation (ds, dt) satisfies the hypothesis of Proposition 14.10, which is not difficult. Second, if $p'_X \colon T(X, \mathcal{O}_X)' \to (X, \mathcal{O}_X)$ is obtained as above but beginning with a different choice of family of charts $(Y_i', h_i' \colon Y_i' \to V_i')_{i \in I'}$, then one must produce a canonical diffeomorphism g making the diagram

$$\begin{CD} T(X, \mathcal{O}_X) @>{g}>> T(X, \mathcal{O}_X)' \\ @V{p_X}VV @VV{p_X'}V \\ (X, \mathcal{O}_X) @= (X, \mathcal{O}_X) \end{CD}$$

commute. This is more delicate, since we have not characterized the tangent bundle by some universal property, and therefore, there is not a unique choice of "canonical" diffeomorphism.[51] We will not go further into this here.

Definition 14.13. A tangent vector field on a smooth manifold $(X, \mathcal{O}_X)$ is a morphism of smooth manifolds $\boldsymbol{v} \colon (X, \mathcal{O}_X) \to T(X, \mathcal{O}_X)$ such that $p_X \circ \boldsymbol{v} = \mathrm{id}_X$.

We note that the value of the map $\boldsymbol{v}$ at $x \in X$ is a vector $\boldsymbol{v}(x) \in T(X, \mathcal{O}_X)_x$ in a vector space that varies with x. We give the set $\mathrm{Vect}(X, \mathcal{O}_X)$ of tangent vector fields on $(X, \mathcal{O}_X)$ the structure of a left $\mathcal{O}_X(X)$-module, where

$$(\boldsymbol{v} + \boldsymbol{w})(x) = \boldsymbol{v}(x) + \boldsymbol{w}(x)$$
$$(\varphi \cdot \boldsymbol{v})(x) = \varphi(x) \cdot \boldsymbol{v}(x)$$

for $\boldsymbol{v}, \boldsymbol{w} \in \mathrm{Vect}(X, \mathcal{O}_X)$ and $\varphi \in \mathcal{O}_X(X)$.

Let $(X, \mathcal{O}_X)$ be a smooth manifold, and let $\boldsymbol{v} \in \mathrm{Vect}(X, \mathcal{O}_X)$ be a tangent vector field. The directional derivative along $\boldsymbol{v}$ is a k-linear map of sheaves

$$\mathcal{O}_X \xrightarrow{D_{\boldsymbol{v}}} \mathcal{O}_X,$$

which we now define. We must define, for all $U \subset X$ open, a k-linear map

$$\mathcal{O}_X(U) \xrightarrow{D_{\boldsymbol{v},U}} \mathcal{O}_X(U)$$

such that for all $U \subset V \subset X$ open, the diagram

$$\begin{CD} \mathcal{O}_X(V) @>{D_{\boldsymbol{v},V}}>> \mathcal{O}_X(V) \\ @V{\mathrm{Res}^V_U}VV @VV{\mathrm{Res}^V_U}V \\ \mathcal{O}_X(U) @>{D_{\boldsymbol{v},U}}>> \mathcal{O}_X(U) \end{CD}$$

commutes. We first note that the smooth tangent vector field $\boldsymbol{v}$ on $(X, \mathcal{O}_X)$ restricts to a smooth tangent vector field $\boldsymbol{v}|_U$ on $(U, \mathcal{O}_X|_U)$ for all $U \subset X$ open. Indeed, if $j \colon U \to X$ is the open immersion of U in X, then the diagram

$$\begin{CD} T(U, \mathcal{O}_X|_U) @>{dj}>> T(X, \mathcal{O}_X) \\ @V{p_U}VV @VV{p_X}V \\ (U, \mathcal{O}_X|_U) @>{j}>> (X, \mathcal{O}_X) \end{CD}$$

[51] It would of course be much better to give a global definition of the tangent bundle similar to the definition $p_X \colon T(X, \mathcal{O}_X) = \mathrm{Spec}(\mathrm{Sym}_{\mathcal{O}_X}(\Omega^1_{X/k})) \to (X, \mathcal{O}_X)$ in algebraic geometry. The liquid theory of Clausen–Scholze now makes this possible.

is cartesian, and therefore, we may define $\boldsymbol{v}|_U\colon (U, \mathcal{O}_X|_U) \to T(U, \mathcal{O}_X|_U)$ to be the unique morphism such that $dj \circ \boldsymbol{v}|_U = \boldsymbol{v} \circ j$ and $p_U \circ \boldsymbol{v}|_U = \mathrm{id}_U$. Next, we may view $\varphi \in \mathcal{O}_X(U)$ as a morphism of smooth manifolds $\varphi\colon (U, \mathcal{O}_X|_U) \to (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}})$, so we have the commutative diagram

$$\begin{array}{ccc} T(U, \mathcal{O}_X|_U) & \xrightarrow{d\varphi} & T(\mathbb{R}, \mathcal{O}_{\mathbb{R}}) \\ \big\downarrow{\scriptstyle p_U} & & \big\downarrow{\scriptstyle p_{\mathbb{R}}} \\ U & \xrightarrow{\varphi} & \mathbb{R}. \end{array}$$

We also have a tangent vector field $\boldsymbol{w}$ on $(\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}})$ defined by $\boldsymbol{w}(t) = (t, \boldsymbol{e}_1)$, and we now define $D_{\boldsymbol{v},U}(\varphi) \in \mathcal{O}_X|_U(U)$ to be the unique element such that

$$d\varphi \circ \boldsymbol{v}|_U = \boldsymbol{w} \cdot D_{v,U}(\varphi).$$

It is clear from the definition that the map $D_{\boldsymbol{v},U}$ is k-linear and that if $U \subset V \subset X$ are open subsets, then $D_{\boldsymbol{v},U} \circ \mathrm{Res}_U^V = \mathrm{Res}_U^V \circ D_{\boldsymbol{v},V}$. Therefore, we have defined a k-linear map of sheaves $D_{\boldsymbol{v}}\colon \mathcal{O}_X \to \mathcal{O}_X$ as desired.

In general, given a morphism $f\colon (X, \mathcal{O}_X) \to (S, \mathcal{O}_S)$ of locally ringed spaces and a right $\mathcal{O}_X$-module $\mathcal{F}$, an $f^*\mathcal{O}_S$-linear morphism of sheaves

$$\mathcal{O}_X \xrightarrow{\delta} \mathcal{F}$$

is an $f^*\mathcal{O}_S$-linear derivation if for all $U \subset X$ open and $\varphi, \psi \in \mathcal{O}_X(U)$,

$$\delta_U(\varphi \cdot \psi) = \delta_U(\varphi) \cdot \psi + \delta_U(\psi) \cdot \varphi.$$

We write $\mathrm{Der}_{\mathcal{O}_S}(\mathcal{O}_X, \mathcal{F})$ for the set of $f^*\mathcal{O}_S$-linear derivations $\delta\colon \mathcal{O}_X \to \mathcal{F}$. It has a structure of abelian group given by the pointwise sum of derivations. Moreover, if $h\colon \mathcal{F} \to \mathcal{F}$ is an $\mathcal{O}_X$-linear morphism and if $\delta\colon \mathcal{O}_X \to \mathcal{F}$ is an $f^*\mathcal{O}_S$-linear derivation, then $h \circ \delta\colon \mathcal{O}_X \to \mathcal{F}$ again is an $f^*\mathcal{O}_S$-linear derivation. So $(h, \delta) \mapsto h \circ \delta$ defines a structure of left $\mathrm{End}_{\mathcal{O}_X}(\mathcal{F})$-module on the abelian group $\mathrm{Der}_{\mathcal{O}_S}(\mathcal{O}_X, \mathcal{F})$.

Lemma 14.14. *If $(X, \mathcal{O}_X)$ is a smooth manifold, then for all $\boldsymbol{v} \in \mathrm{Vect}(X, \mathcal{O}_X)$, the directional derivative $D_{\boldsymbol{v}}\colon \mathcal{O}_X \to \mathcal{O}_X$ is a k-linear derivation.*

Proof. Given $\boldsymbol{v} \in \mathrm{Vect}(X, \mathcal{O}_X)$, an open subset $U \subset X$, and a point $x \in U$, we give a formula for $D_{\boldsymbol{v},U}(\varphi)(x)$ for $\varphi \in \mathcal{O}_X|_U(U)$. There exists a smooth curve $\gamma\colon (I, \mathcal{O}_I^{\mathrm{sm}}) \to (U, \mathcal{O}_X|_U)$ defined on an open interval $0 \in I \subset \mathbb{R}$ such that $\gamma(0) = x$ and such that, in the diagram

$$\begin{array}{ccccc} T(I, \mathcal{O}_I^{\mathrm{sm}}) & \xrightarrow{d\gamma} & T(U, \mathcal{O}_X|_U) & \xrightarrow{d\varphi} & T(\mathbb{R}, \mathcal{O}_{\mathbb{R}}) \\ \big\downarrow{\scriptstyle p_I} & & \big\downarrow{\scriptstyle p_U} & & \big\downarrow{\scriptstyle p_{\mathbb{R}}} \\ (I, \mathcal{O}_I) & \xrightarrow{\gamma} & (U, \mathcal{O}_X|_U) & \xrightarrow{\varphi} & (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}), \end{array}$$

we have $(d\gamma \circ \boldsymbol{w}|_I)(0) = \boldsymbol{v}|_U(x) = (\boldsymbol{v}|_U \circ \gamma)(0)$. Therefore,

$$(d\varphi \circ \boldsymbol{v}|_U)(x) = (d\varphi \circ \boldsymbol{v}|_U \circ \gamma)(0) = (d\varphi \circ d\gamma \circ \boldsymbol{w}|_I)(0) = (d(\varphi \circ \gamma) \circ \boldsymbol{w}|_I)(0),$$

from which we obtain the formula

$$D_{\boldsymbol{v},U}(\varphi)(x) = (\varphi \circ \gamma)'(0).$$

Hence, for all $\varphi, \psi \in \Gamma(U, \mathcal{O}_X)$, we have

$$\begin{aligned} D_{\boldsymbol{v},U}(\varphi \cdot \psi)(x) &= ((\varphi \cdot \psi) \circ \gamma)'(0) = ((\varphi \circ \gamma) \cdot (\psi \circ \gamma))'(0) \\ &= (\varphi \circ \gamma)'(0) \cdot (\psi \circ \gamma)(0) + (\psi \circ \gamma)'(0) \cdot (\varphi \circ \gamma)(0) \\ &= D_{\boldsymbol{v},U}(\varphi)(x) \cdot \psi(x) + D_{\boldsymbol{v},U}(\psi)(x) \cdot \varphi(x), \end{aligned}$$

and since $x \in U$ was arbitrary, we conclude that

$$D_{\boldsymbol{v},U}(\varphi \cdot \psi) = D_{\boldsymbol{v},U}(\varphi) \cdot \psi + D_{\boldsymbol{v},U}(\psi) \cdot \varphi$$

as desired. □

We now obtain the promised global description of the left $\mathcal{O}_X(X)$-module of tangent vector fields.

Proposition 14.15. *Let $(X, \mathcal{O}_X)$ be a smooth manifold. The directional derivative*

$$\mathrm{Vect}(X, \mathcal{O}_X) \xrightarrow{D} \mathrm{Der}_k(\mathcal{O}_X, \mathcal{O}_X)$$

is an isomorphism of left $\mathcal{O}_X(X)$-modules.

Proof. For all open subsets $U \subset V \subset X$, we have a commutative diagram

$$\begin{array}{ccc} \mathrm{Vect}(V, \mathcal{O}_X|_V) & \xrightarrow{D_V} & \mathrm{Der}_k(\mathcal{O}_X|_V, \mathcal{O}_X|_V) \\ \downarrow \mathrm{Res}^V_U & & \downarrow \mathrm{Res}^V_U \\ \mathrm{Vect}(U, \mathcal{O}_X|_U) & \xrightarrow{D_U} & \mathrm{Der}_k(\mathcal{O}_X|_U, \mathcal{O}_X|_U), \end{array}$$

so the family $(D_U)_{U \subset X}$ is a morphism of presheaves of left $\mathcal{O}_X$-modules

$$\mathcal{V}ect(X, \mathcal{O}_X) \xrightarrow{D} \mathcal{D}er_k(\mathcal{O}_X, \mathcal{O}_X).$$

Both of these presheaves are in fact sheaves, because they are defined in terms of by local conditions. We will prove that this morphism of sheaves is an isomorphim. Since the map in the statement is obtained from this morphism of sheaves by applying the global sections functor $\Gamma(X, -)$, this will prove the proposition.

Since the statement that the map of sheaves in question is an isomorphism is local on X, we may assume that $(X, \mathcal{O}_X)$ is equal to $(U, \mathcal{O}_U^{\mathrm{sm}})$ with $U \subset \mathbb{R}^n$ open. We may further assume that $U \subset \mathbb{R}^n$ is convex, since every open subset of $\mathbb{R}^n$ admits a covering by convex open subsets. So it suffices to prove that for $U \subset \mathbb{R}^n$ convex open, the directional derivative

$$\mathrm{Vect}(U, \mathcal{O}_U^{\mathrm{sm}}) \xrightarrow{D} \mathrm{Der}_k(\mathcal{O}_U^{\mathrm{sm}}, \mathcal{O}_U^{\mathrm{sm}})$$

is an isomorphism of left $\mathcal{O}_U^{\mathrm{sm}}(U)$-modules. The left-hand $\mathcal{O}_U^{\mathrm{sm}}(U)$-module is free of rank n, and a basis is given by the family $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ of vector fields defined by $\boldsymbol{w}_i(x) = (x, \boldsymbol{e}_i)$, where $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ is the standard basis of $\mathbb{R}^n$. By the definition of the directional derivative, we have

$$D_{\boldsymbol{w}_i}(\varphi) = \partial\varphi / \partial x_i,$$

so we must prove that the family $(\partial/\partial x_1, \dots, \partial/\partial x_n)$ of derivations is a basis of the left $\mathcal{O}_U^{\mathrm{sm}}(U)$-module $\mathrm{Der}_k(\mathcal{O}_U^{\mathrm{sm}}, \mathcal{O}_U^{\mathrm{sm}})$. It is linearly independent, since

$$\partial x_i/\partial x_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

and to show that it also generates the left $\mathcal{O}_U^{\mathrm{sm}}(U)$-module $\mathrm{Der}_k(\mathcal{O}_U^{\mathrm{sm}}, \mathcal{O}_U^{\mathrm{sm}})$, we prove that for all $\delta \in \mathrm{Der}_k(\mathcal{O}_U^{\mathrm{sm}}, \mathcal{O}_U^{\mathrm{sm}})$, the following identity holds,

$$\delta = \textstyle\sum_{i=1}^n \delta(x_i) \cdot \partial/\partial x_i.$$

It suffices to show that for all $\delta \in \mathrm{Der}_k(\mathcal{O}_U^{\mathrm{sm}}, \mathcal{O}_U^{\mathrm{sm}})$, $\varphi \in \mathcal{O}_U^{\mathrm{sm}}(U)$, and $a \in U$,

$$\delta(\varphi)(a) = \textstyle\sum_{i=1}^n \delta(x_i)(a) \cdot (\partial\varphi/\partial x_i)(a).$$

Indeed, the sheaf $\mathcal{F} = \mathcal{O}_U^{\mathrm{sm}}$ has the special property that a section $\psi \in \mathcal{O}_U^{\mathrm{sm}}(U)$ is zero if and only if all its values $\psi(a) \in \mathcal{F}(a) = \mathcal{F}_a \otimes_{\mathcal{O}_{U,a}^{\mathrm{sm}}} k(a)$ are zero. Now, since we assumed that the open subset $U \subset \mathbb{R}^n$ is convex, Corollary 14.21 below shows that there exist unique $\varphi_{i,j} \in \mathcal{O}_U^{\mathrm{sm}}(U)$ such that

$$\varphi(x) = \varphi(a) + \textstyle\sum_{i=1}^n (x_i - a_i)(\partial\varphi/\partial x_i)(a) + \sum_{i,j=1}^n (x_i - a_i)(x_j - a_j)\varphi_{i,j}(x),$$

and since δ is a k-linear derivation, the desired identity ensues. □

Example 14.16. If $(X, \mathcal{O}_X)$ is a smooth manifold, and if $h \colon U \to V$ is a chart with $V \subset \mathbb{R}^n$ open, then the family of derivations $(\delta_1, \dots, \delta_n)$, where

$$\delta_i(\varphi)(x) = (\partial(\varphi \circ h^{-1})/\partial x_i)(h(x)),$$

is a basis of the left $\mathcal{O}_X(U)$-module $\mathrm{Der}_k(\mathcal{O}_X|_U, \mathcal{O}_X|_U)$. Hence, there is a unique basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of the left $\mathcal{O}_X(U)$-module $\mathrm{Vect}(U, \mathcal{O}_X|_U)$ such that $D_{\boldsymbol{v}_i} = \delta_i$.

According to Proposition 14.15, tangent vector fields may analogously be defined to be k-linear derivations $\delta \colon \mathcal{O}_X \to \mathcal{O}_X$. This definition has the advantage of being truly global. We define the "Lie bracket"

$$\mathrm{Der}_k(\mathcal{O}_X, \mathcal{O}_X) \otimes_k \mathrm{Der}_k(\mathcal{O}_X, \mathcal{O}_X) \xrightarrow{[-,-]} \mathrm{Der}_k(\mathcal{O}_X, \mathcal{O}_X)$$

to be the map that to $\delta_1 \otimes \delta_2$ assigns the k-linear morphism

$$[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1.$$

To verify that $[\delta_1, \delta_2] \in \mathrm{Der}_k(\mathcal{O}_X, \mathcal{O}_X)$, we let $\varphi, \psi \in \Gamma(U, \mathcal{O}_X|_U)$ and calculate

$$\begin{aligned} [\delta_1, \delta_2](\varphi \cdot \psi) &= \delta_1(\delta_2(\varphi \cdot \psi)) - \delta_2(\delta_1(\varphi \cdot \psi)) \\ &= \delta_1(\delta_2(\varphi) \cdot \psi + \varphi \cdot \delta_2(\psi)) - \delta_2(\delta_1(\varphi) \cdot \psi + \varphi \cdot \delta_1(\psi)) \\ &= \delta_1(\delta_2(\varphi)) \cdot \psi + \delta_2(\varphi) \cdot \delta_1(\psi) + \delta_1(\varphi) \cdot \delta_2(\psi) + \varphi \cdot \delta_1(\delta_2(\psi)) \\ &\quad - \delta_2(\delta_1(\varphi)) \cdot \psi - \delta_1(\varphi) \cdot \delta_2(\psi) - \delta_2(\varphi) \cdot \delta_1(\psi) - \varphi \cdot \delta_2(\delta_1(\psi)) \\ &= [\delta_1, \delta_2](\varphi) \cdot \psi + \varphi \cdot [\delta_1, \delta_2](\psi). \end{aligned}$$

It is clear that the map $[-,-]$ is k-linear in both arguments so that we obtain the stated map. A similar and equally straightforward calculation shows that given three k-linear derivations δ_1, δ_2, and δ_3, the "Jacobi identity"

$$[[\delta_1, \delta_2], \delta_3] + [[\delta_2, \delta_3], \delta_1] + [[\delta_3, \delta_1], \delta_2] = 0$$

holds. This makes $\mathrm{Der}_k(\mathcal{O}_X, \mathcal{O}_X)$ a Lie algebra over k.[52]

We proved earlier that the category of smooth manifolds and morphisms of smooth manifolds has finite products. It does not have all fiber products, but the implicit function theorem shows that it does have some fiber products. Given a cartesian square of smooth manifolds and morphism of smooth manifolds

$$\begin{array}{ccc} (Y', \mathcal{O}_{Y'}) & \xrightarrow{g'} & (Y, \mathcal{O}_Y) \\ \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ (X', \mathcal{O}_{X'}) & \xrightarrow{g} & (X, \mathcal{O}_X), \end{array}$$

we say that f' is the base-change of f along g. If such a square exists for given f and g, then we say that the base-change of f along g exists.

A morphism of smooth manifolds $f\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ is a submersion[53] (resp. an immersion) if for all $y \in Y$ with image $x = f(y) \in X$, the differential

$$T(Y, \mathcal{O}_Y)_y \xrightarrow{df_y} T(X, \mathcal{O}_X)_x$$

is surjective (resp. injective). We note that, in this case, it follows from linear algebra that $\dim(y) \geq \dim(x)$ (resp. $\dim(y) \leq \dim(x)$).

Theorem 14.17 (Implicit function theorem)**.** *In the category of smooth manifolds and morphisms of smooth manifolds, the base-change of a submersion along any morphism exists and is a submersion.*

Proof. This is based on the inverse function theorem. It states that a morphism of smooth manifolds, which is both an immersion and a submersion, is étale. The proof has a number of steps. First, if $(Y, \mathcal{O}_Y) = (X \times Z, \mathcal{O}_{X \times Z})$ and f is the projection on the first factor, then the base-change along any g exists with $Y' = X' \times Z$, with f' the projection on the first factor, and with $g' = g \times \mathrm{id}_Z$. Second, the inverse function theorem shows if f is any submersion, then for all $y \in Y$, we can find open neighborhoods $y \in V \subset Y$, $x = f(y) \in U \subset X$, and $0 \in W \subset \mathbb{R}^p$ together with a diffeomorphism h making the diagram

$$\begin{array}{ccc} (V, \mathcal{O}_Y|_V) & \xrightarrow{h} & (U \times W, \mathcal{O}_{X \times W}|_{U \times W}) \\ \downarrow{\scriptstyle f|_U} & & \downarrow{\scriptstyle p} \\ (U, \mathcal{O}_X|_U) & = & (U, \mathcal{O}_X|_U), \end{array}$$

where p is the canonical projection, commute. Hence, it follows from the first step that the base-change of $f|_U$ along any morphism g exists and is a submersion. Finally, we use Proposition 14.10 to glue together the local solutions obtained in the second step to a global solution. To do so, we also use the fact that the base-change of an open immersion along any morphism exists and is an open immersion and the fact that base-change along an open immersion preserves both coequivalizers and submersions. □

[52] This Lie algebra is infinite dimensional, unless X is finite. We will define the Lie algebra of a Lie group to be a subalgebra of this Lie algebra.

[53] In algebraic geometry, the analogue of submersions are called smooth morphisms. It is for this reason, that I say "morphism of smooth manifolds" instead of "smooth map."

Remark 14.18. Let $f\colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ be a morphism of smooth manifolds. We say that $y \in Y$ is a regular point of f if df_y is surjective and that $x \in X$ is a regular value of f if every $y \in Y$ with $f(y) = x$ is a regular point. Therefore, given a morphism $g\colon (X', \mathcal{O}_{X'}) \to (X, \mathcal{O}_X)$ for which there exists $g(X') \subset U \subset X$ open such that every $x \in U$ is a regular point of f, then the base-change of f along g exists and is equal to the base-change of $f|_{f^{-1}(U)}$ along g.

Example 14.19. Let $Y = M_n(\mathbb{R})$, and let $X \subset M_n(\mathbb{R})$ be the subset of symmetric matrices. So Y and X are both real vector spaces of dimension n^2 and $(n+1)n/2$, respectively, which we view as smooth manifolds of the same dimensions. The map $f\colon (Y, \mathcal{O}_Y^{\mathrm{sm}}) \to (X, \mathcal{O}_X^{\mathrm{sm}})$ defined by $f(A) = A^*A$ is smooth, and we claim that

$$T(Y, \mathcal{O}_Y^{\mathrm{sm}})_A \xrightarrow{df_A} T(X, \mathcal{O}_X^{\mathrm{sm}})_{f(A)}$$

is surjective for all $A \in Y$ with $f(A) = E \in X$. To see this, we use the identity maps of Y and X as charts and calculate

$$\begin{aligned} df_A(B) &= \lim_{h \to 0}(f(A + hB) - f(A))/h \\ &= \lim_{h \to 0}((A + hB)^*(A + hB) - A^*A)/h \\ &= \lim_{h \to 0}(A^*A + hA^*B + hB^*A + h^2B^*B - A^*A)/h \\ &= A^*B + B^*A. \end{aligned}$$

Now, if $f(A) = A^*A = E$, then given $C = C^* \in X$, we set $B = \frac{1}{2}AC$ and calculate

$$df_A(B) = A^*B + B^*A = \tfrac{1}{2}A^*AC + \tfrac{1}{2}C^*A^*A = \tfrac{1}{2}(C + C^*) = C.$$

So the implicit function theorem shows that the base-change

$$\begin{array}{ccc} (O(n), \mathcal{O}_{O(n)}) & \xrightarrow{g'} & (Y, \mathcal{O}_Y^{\mathrm{sm}}) \\ \big\downarrow{\scriptstyle f'} & & \big\downarrow{\scriptstyle f} \\ (\{E\}, \mathcal{O}_{\{E\}}) & \xrightarrow{g} & (X, \mathcal{O}_X^{\mathrm{sm}}) \end{array}$$

exists; see Remark 14.18. Hence, the subspace $O(n) \subset M_n(\mathbb{R})$ of orthogonal matrices has a structure of smooth manifold of dimension $n^2 - (n+1)n/2 = n(n-1)/2$.

Appendix: Hadamard's lemma

We have used the following result, commonly referred to as Hadamard's lemma.

Lemma 14.20. *Let $U \subset \mathbb{R}^n$ be an open subset that is star-convex with respect to $a \in U$, and let $\varphi\colon U \to \mathbb{R}$ is a smooth function. Then there exists unique smooth functions $\varphi_i\colon U \to \mathbb{R}$ such that for all $x \in U$,*

$$\varphi(x) = \varphi(a) + \textstyle\sum_{i=1}^n (x_i - a_i)\varphi_i(x).$$

Moreover, for all $1 \le i \le n$, $\varphi_i(a) = (\partial\varphi/\partial x_i)(a)$.

Proof. We define $h\colon [0,1] \to \mathbb{R}$ by $h(t) = \varphi(a + (x-a)t)$, which is possible by the assumption that U' be star-convex with respect to a, and calculate that

$$\begin{aligned}\varphi(x) - \varphi(a) &= h(1) - h(0) = \textstyle\int_0^1 (dh/dt)(t)dt \\ &= \textstyle\int_0^1 \sum_{i=1}^n (\partial\varphi/\partial x_i)(a + (x-a)t)(x_i - a_i)dt \\ &= \textstyle\sum_{i=1}^n (x_i - a_i) \int_0^1 (\partial\varphi/\partial x_i)(a + (x-a)t)dt.\end{aligned}$$

So the lemma holds with $\varphi_i(x) = \int_0^1 (\partial\varphi/\partial x_i)(a + (x-a)t)dt$. □

Corollary 14.21. *Let $U \subset \mathbb{R}^n$ be an open subset that is star-convex with respect to $a \in U$, and let $\varphi\colon U \to \mathbb{R}$ is a smooth function. Then there exists unique smooth functions $\varphi_{i,j}\colon U \to \mathbb{R}$ such that for all $x \in U$,*

$$\varphi(x) = \varphi(a) + \textstyle\sum_{i=1}^n (x_i - a_i)(\partial\varphi/\partial x_i)(a) + \sum_{i,j=1}^n (x_i - a_i)(x_j - a_j)\varphi_{i,j}(x).$$

Proof. We first write $\varphi(x)$ as in the statement of Lemma 14.20 and the apply the lemma again to write each of the functions $\varphi_i\colon U \to \mathbb{R}$ as

$$\varphi_i(x) = \varphi_i(a) + \textstyle\sum_{j=1}^n (x_j - a_j)\varphi_{i,j}(x) = (\partial\varphi/\partial x_i)(a) + \sum_{j=1}^n (x_j - a_j)\varphi_{i,j}(x)$$

with $\varphi_{i,j}\colon U \to \mathbb{R}$ smooth. □

Definition 15.1. A Lie group is a group object $G = ((G, \mathcal{O}_G), \mu, \iota)$ in the category of smooth manifolds and morphisms of smooth manifolds.[54] A morphism of Lie groups is a homomorphism of group objects in the category of smooth manifolds and morphisms of smooth manifolds.

One defines complex Lie groups similarly to be a group objects in the category of complex manifolds and morphism of complex manifolds.

There is a "forgetful" functor from the category of Lie groups and morphisms of Lie groups to that of topological groups and continuous group homomorphisms that to $((G, \mathcal{O}_G), \mu, \iota)$ assigns (G, μ, ι). One can prove that this functor is fully faithful,[55] so in particular, the sheaf $\mathcal{O}_G$ is uniquely determined, up to unique isomorphism, by the remaining data. Hence, we may view "being a Lie group" as a property of a topological group.

Example 15.2. By using the implicit function theorem, we see that the classical groups all are (real) Lie groups. The groups $\mathrm{GL}_n(\mathbb{C})$ and $\mathrm{SL}_n(\mathbb{C})$ are examples of complex Lie groups.

If $((G, \mathcal{O}_G), \mu, \iota)$ is a Lie group, then we may consider the tangent space

$$\mathfrak{g} = T(G, \mathcal{O}_G)_e$$

of the smooth manifold $(G, \mathcal{O}_G)$ at the identity element $e \in G$. It is a real vector space of dimension $n = \dim(e)$. We proceed to show that the group structure morphisms μ and ι give rise to a structure of Lie algebra $[-,-]$ on this real vector space. Let us first define Lie algebras.

Definition 15.3. Let k be a field. A Lie algebra over k is a pair $\mathfrak{g} = (\mathfrak{g}, [-,-])$ of a right k-vector space $\mathfrak{g}$ and a k-linear map $[-,-]\colon \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ such that:

(LA1) For all $x \in \mathfrak{g}$, $[x, x] = 0$.
(LA2) For all $x, y, z \in \mathfrak{g}$, $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

A morphism of Lie algebras $f\colon (\mathfrak{h}, [-,-]) \to (\mathfrak{g}, [-,-])$ is a k-linear map $f\colon \mathfrak{h} \to \mathfrak{g}$ such that for all $x, y \in \mathfrak{h}$, $f([x, y]) = [f(x), f(y)]$.

We call $[-,-]$ the "Lie bracket" and we refer to (LA1) and (LA2) by saying that the Lie bracket is alternating and satisfies the Jacobi identity, respectively. It follows that the Lie bracket is antisymmetric in that for all $x, y \in \mathfrak{g}$, $[x, y] = -[y, x]$. We warn the reader that the Lie bracket is neither associative nor does it have an identity element, except in trivial cases. A Lie algebra $\mathfrak{a}$ is defined to be abelian if $[x, x] = 0$ for all $x \in \mathfrak{a}$.

Example 15.4. (1) An associative k-algebra A determines a Lie algebra with the same underlying k-vector space as A and with Lie bracket $[a, b] = a \cdot b - b \cdot a$. In particular, if V is a right k-vector space, then $\mathrm{End}_k(V)$ is an associative k-algebra under composition of k-linear maps. The associated Lie algebra is denoted $\mathfrak{gl}(V)$.

[54] If (G, μ, ι) is a topological group and if $\{e\} \subset G$ is closed, then the space G is automatically Hausdorff. For the diagonal $\Delta(G) \subset G \times G$ is the preimage by the continuous map $\mu \circ (\mathrm{id} \times \iota)$ of the closed subset $\{e\} \subset G$ and hence closed.

[55] See [3, Theorem 9.2.16].

(2) If $(X, \mathcal{O}_X)$ is a smooth manifold, then the real vector space $\operatorname{Der}_k(\mathcal{O}_X, \mathcal{O}_X)$ has a structure of real Lie algebra with Lie bracket $[-,-]$ defined by[56]

$$[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1.$$

Hence, there is a unique structure of Lie algebra on $\operatorname{Vect}(X, \mathcal{O}_X)$ for which the directional derivative is an isomorphism of Lie algebras.

In particular, a morphism of Lie groups $\pi \colon G \to \mathrm{GL}(V)$ gives rise to a morphism of Lie algebras $d\pi_e \colon \mathfrak{g} \to \mathfrak{gl}(V)$. So a representation of a Lie group determines a representation of its Lie algebra. This assignment is a functor, and we will prove that its restriction to the full subcategory of connected Lie groups is faithful.

Let $G = ((G, \mathcal{O}_G), \mu, \iota)$ be a Lie group. Given $g \in G$, we write

$$(G, \mathcal{O}_G) \xrightarrow{L_g} (G, \mathcal{O}_G)$$

for the morphism of smooth manifolds defined by $L_g(x) = \mu(g, x) = gx$ and call it "left multiplication by $g \in G$." The map L_g is not a group homomorphism, but it is an automorphism of smooth manifolds, so we get a map

$$G \xrightarrow{L} \operatorname{Aut}(G, \mathcal{O}_G)$$

from G to the group of automorphism of the smooth manifold $(G, \mathcal{O}_G)$, and this map is a group homomorphism. We wish to consider the induced actions on the "space" of tangent vector fields. We first prove a general result.

Proposition 15.5. *Let $f \colon (Y, \mathcal{O}_Y) \to (X, \mathcal{O}_X)$ be a morphism of smooth manifolds, and let $D_{\boldsymbol{u}} \in \operatorname{Der}_k(\mathcal{O}_X, \mathcal{O}_X)$ and $D_{\boldsymbol{v}} \in \operatorname{Der}_k(\mathcal{O}_Y, \mathcal{O}_Y)$ be the directional derivatives along two tangent vector fields $\boldsymbol{u} \in \operatorname{Vect}(X, \mathcal{O}_X)$ and $\boldsymbol{v} \in \operatorname{Vect}(Y, \mathcal{O}_Y)$, respectively. The following statements are equivalent.*

(a) *The diagram of smooth manifolds and morphisms of smooth manifolds*

$$\begin{array}{ccc} T(Y, \mathcal{O}_Y) & \xrightarrow{df} & T(X, \mathcal{O}_X) \\ \uparrow{\scriptstyle \boldsymbol{v}} & & \uparrow{\scriptstyle \boldsymbol{u}} \\ (Y, \mathcal{O}_Y) & \xrightarrow{f} & (X, \mathcal{O}_X) \end{array}$$

commutes.

(b) *The diagram of sheaves of $\mathcal{O}_X$-modules and k-linear maps*

$$\begin{array}{ccc} \mathcal{O}_X & \xrightarrow{f^\sharp} & f_*\mathcal{O}_Y \\ \downarrow{\scriptstyle D_{\boldsymbol{u}}} & & \downarrow{\scriptstyle f_*D_{\boldsymbol{v}}} \\ \mathcal{O}_X & \xrightarrow{f^\sharp} & f_*\mathcal{O}_Y \end{array}$$

commutes.

56 If $\delta_1, \delta_2 \colon \mathcal{O}_X \to \mathcal{O}_X$ are k-linear derivations, then $\delta_1 \circ \delta_2 \colon \mathcal{O}_X \to \mathcal{O}_X$ is typically not a k-linear derivation. So (2) is not a special case of (1).

Proof. We first assume (a) and prove (b). We must show that for all $U \subset X$ open with $V = f^{-1}(U) \subset Y$ and for all $\varphi \in \Gamma(V, \mathcal{O}_Y)$, the identity

$$D_{\boldsymbol{v}}(\varphi \circ f|_V) = f|_V \circ D_{\boldsymbol{u}}(\varphi)$$

holds. But this follows from the chain rule. Indeed, we consider the diagram

$$\begin{array}{ccccc} T(V, \mathcal{O}_Y|_V) & \xrightarrow{df|_V} & T(U, \mathcal{O}_X|_U) & \xrightarrow{d\varphi} & T(\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \\ \uparrow{\scriptstyle \boldsymbol{v}|_V} & & \uparrow{\scriptstyle \boldsymbol{u}|_U} & & \uparrow{\scriptstyle \boldsymbol{w}} \\ (V, \mathcal{O}_Y|_V) & \xrightarrow{f|_V} & (U, \mathcal{O}_X|_U) & \xrightarrow{\varphi} & (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \end{array}$$

where $\boldsymbol{w}$ is the constant vector field defined by $\boldsymbol{w}(t) = (t, \boldsymbol{e}_1)$. We have

$$\begin{aligned} &D_{\boldsymbol{v}}(\varphi \circ f|_V) \cdot (\boldsymbol{w} \circ \varphi \circ f|_V) = d(\varphi \circ f|_V) \circ \boldsymbol{v}|_V = d\varphi \circ d(f|_V) \circ \boldsymbol{v}|_V \\ &= d\varphi \circ \boldsymbol{u}|_U \circ f|_V = (D_{\boldsymbol{u}}(\varphi) \circ f|_V) \cdot (\boldsymbol{w} \circ \varphi \circ f|_V) \end{aligned}$$

where the first and last identity hold by the definition of $D_{\boldsymbol{u}}$ and $D_{\boldsymbol{v}}$, the second identity holds by the chain rule, and the third identity holds by (a).

We next assume (b) and prove (a). Since p_X is a submersion, the implicit function theorem shows that the base-change of p_X along f exists,

$$\begin{array}{ccc} T(X, \mathcal{O}_X)' & \xrightarrow{f'} & T(X, \mathcal{O}_X) \\ \downarrow{\scriptstyle p'_X} & & \downarrow{\scriptstyle p_X} \\ (Y, \mathcal{O}_Y) & \xrightarrow{f} & (X, \mathcal{O}_X). \end{array}$$

We repeat the definition of the directional derivative to define a map

$$\mathrm{Vect}(X, \mathcal{O}_X)' \xrightarrow{D'} \mathrm{Der}_k(\mathcal{O}_X, f_*\mathcal{O}_Y)$$

from the set of morphism of smooth manifolds $\boldsymbol{s} \colon (Y, \mathcal{O}_Y) \to T(X, \mathcal{O}_X)'$ such that $p'_X \circ \boldsymbol{s} = \mathrm{id}_Y$ to the set of k-linear derivations $\delta \colon \mathcal{O}_X \to f_*\mathcal{O}_Y$. Given $U \subset X$ open with $V = f^{-1}(U) \subset Y$ and $\varphi \in \Gamma(U, \mathcal{O}_X)$, we consider the diagram

$$\begin{array}{ccccc} T(U, \mathcal{O}_X|_U)' & \xrightarrow{(f|_V)'} & T(U, \mathcal{O}_X|_U) & \xrightarrow{d\varphi} & T(\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \\ \downarrow{\scriptstyle p'_U} & & \downarrow{\scriptstyle p_U} & & \downarrow{\scriptstyle p_{\mathbb{R}}} \\ (V, \mathcal{O}_Y|_V) & \xrightarrow{f|_V} & (U, \mathcal{O}_X|_U) & \xrightarrow{\varphi} & (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \end{array}$$

and define $D'_{\boldsymbol{s}}(\varphi) \in \Gamma(U, f_*\mathcal{O}_Y)$ to be the unique element such that

$$D'_{\boldsymbol{s}}(\varphi) \cdot (\boldsymbol{w} \circ \varphi \circ f|_V) = d\varphi \circ (f|_V)' \circ \boldsymbol{s}|_V.$$

Now, the two composites $\boldsymbol{u} \circ f$ and $df \circ \boldsymbol{v}$ of the morphisms in the top diagram in the statement are both elements of $\mathrm{Vect}(X, \mathcal{O}_X)'$, and we have

$$D'_{\boldsymbol{u} \circ f} = f^\sharp \circ D_{\boldsymbol{u}} = f_* D_{\boldsymbol{v}} \circ f^\sharp = D'_{df \circ \boldsymbol{v}}.$$

Indeed, the first and last identity follow immediately from the definitions of D and D', and the middle identity is (b). Hence, it will suffice to prove that the map D' is injective.[57] To this end, we proceed as in the proof of Proposition 15 last time.

[57] The map D' need not be surjective, because the sheaf $f_*\mathcal{O}_Y$ can be very complicated.

We first observe that the map in question is equal to the map of global sections induced by a map of sheaves of $\mathcal{O}_Y$-modules

$$\mathcal{V}ect(X, \mathcal{O}_X)' \xrightarrow{D'} \mathcal{D}er_k(\mathcal{O}_X, f_*\mathcal{O}_Y).$$

Hence, we may assume that $(X, \mathcal{O}_X)$ is equal to $(U, \mathcal{O}_U^{\mathrm{sm}})$ with $U \subset \mathbb{R}^m$ an open subset. In this case, the $\Gamma(Y, \mathcal{O}_Y)$-module $\mathrm{Vect}(X, \mathcal{O}_X)'$ is free of rank m, and a basis is given by the family $(\boldsymbol{s}_1, \dots, \boldsymbol{s}_m)$ with $\boldsymbol{s}_i = \boldsymbol{w}_i \circ f$, where $\boldsymbol{w}_i(x) = (x, \boldsymbol{e}_i)$ and where $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ is the standard basis of $\mathbb{R}^m$. Moreover, we have

$$D'_{\boldsymbol{s}_i} = D'_{\boldsymbol{w}_i \circ f} = f^\sharp \circ D_{\boldsymbol{w}_i} = f^\sharp \circ (\partial/\partial x_i),$$

and since $f^\sharp \colon \mathcal{O}_X \to f_*\mathcal{O}_Y$ is a ring homomorphism, we find that

$$D'_{\boldsymbol{s}_j}(x_i) = f^\sharp \circ (\partial x_i/\partial x_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

This shows that the family $(f^\sharp \circ (\partial/\partial x_1), \dots, f^\sharp \circ (\partial/\partial x_m))$ is linear independent, which, in turn, shows that D' is injective as desired. □

Now, if $(X, \mathcal{O}_X)$ is a smooth manifold, then we obtain a group homomorphism

$$\mathrm{Aut}(X, \mathcal{O}_X) \xrightarrow{\tau} \mathrm{Aut}_k(\mathrm{Vect}(X, \mathcal{O}_X))$$

defined by $\tau(f)(\boldsymbol{v}) = \boldsymbol{u}$, where $\boldsymbol{u}, \boldsymbol{v} \in \mathrm{Vect}(X, \mathcal{O}_X)$ are related as in the statement of Proposition 15.5. We note that the map $\tau(f)$ is not a $\Gamma(X, \mathcal{O}_X)$-linear automorphism, but, instead, it is a $\Gamma(X, \mathcal{O}_X)$-linear isomorphism

$$\mathrm{Vect}(X, \mathcal{O}_X) \xrightarrow{\tau(f)} f^{\sharp *}\,\mathrm{Vect}(X, \mathcal{O}_X)$$

from $\mathrm{Vect}(X, \mathcal{O}_X)$ to the left $\Gamma(X, \mathcal{O}_X)$ obtained from $\mathrm{Vect}(X, \mathcal{O}_X)$ by extension of scalars along $f^\sharp \colon \Gamma(X, \mathcal{O}_X) \to \Gamma(X, \mathcal{O}_X)$. We will not explore this further here and will simply consider $\tau(f)$ as a k-linear automorphism of $\mathrm{Vect}(X, \mathcal{O}_X)$. However, it is clear from Proposition 15.5 that for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathrm{Vect}(X, \mathcal{O}_X)$,

$$[\tau(f)(\boldsymbol{v}_1), \tau(f)(\boldsymbol{v}_2)] = \tau(f)([\boldsymbol{v}_1, \boldsymbol{v}_2]),$$

so we may view τ as a group homomorphism

$$\mathrm{Aut}(X, \mathcal{O}_X) \xrightarrow{\tau} \mathrm{Aut}_k(\mathrm{Vect}(X, \mathcal{O}_X), [-,-])$$

to the group of automorphisms of the real Lie algebra of tangent vector fields on the smooth manifold $(X, \mathcal{O}_X)$.

We return to the case of a Lie group G. We define "left translation of tangent vector fields" to be the composite group homomorphism

$$G \xrightarrow{L} \mathrm{Aut}(G, \mathcal{O}_G) \xrightarrow{\tau} \mathrm{Aut}_k(\mathrm{Vect}(G, \mathcal{O}_G), [-,-]),$$

and we define a "left-invariant tangent vector field" to be a tangent vector field $\boldsymbol{v}$ that is fixed under left translation by every $g \in G$.

Definition 15.6. The Lie algebra of a Lie group G is the sub-Lie algebra

$$(\mathrm{Lie}(G), [-,-]) = (\mathrm{Vect}(G, \mathcal{O}_G), [-,-])^G$$

of left-invariant tangent vector fields.[58]

We also write $\mathfrak{g}$ instead of $\mathrm{Lie}(G)$. We now show that the k-vector space $\mathrm{Lie}(G)$ is finite dimensional, and that the assignment of $\mathrm{Lie}(G)$ to G extends to a functor from the category of Lie groups and morphisms of Lie groups to the category of Lie algebras and morphisms of Lie algebras.[59]

Proposition 15.7. *If G is a Lie group, then the map $\epsilon_G \colon \mathfrak{g} \to T(G, \mathcal{O}_G)_e$ defined by $\epsilon_G(\boldsymbol{v}) = \boldsymbol{v}(e)$ is a k-linear isomorphism. Moreover, if $f \colon H \to G$ is a morphism of Lie groups, then the unique k-linear map $\mathrm{Lie}(f)$ that makes the diagram*

$$\begin{array}{ccc} \mathrm{Lie}(H) & \xrightarrow{\mathrm{Lie}(f)} & \mathrm{Lie}(G) \\ \downarrow{\scriptstyle \epsilon_H} & & \downarrow{\scriptstyle \epsilon_G} \\ T(H, \mathcal{O}_H)_e & \xrightarrow{df_e} & T(G, \mathcal{O}_G)_e \end{array}$$

commute is a morphism of Lie algebras.

Proof. A tangent vector field $\boldsymbol{u} \in \mathrm{Vect}(G, \mathcal{O}_G)$ is left-invariant if for all $g \in G$,

$$\boldsymbol{u}(g) = dL_{g,e}(\boldsymbol{u}(e)),$$

so the first part of the statement is clear. To prove the second part of the statement, we note that if $\boldsymbol{v} \in \mathrm{Vect}(H, \mathcal{O}_H)$, then $\boldsymbol{u} = \mathrm{Lie}(f)(\boldsymbol{v}) \in \mathrm{Vect}(G, \mathcal{O}_G)$ is characterized as the unique left-invariant vector field such that $\boldsymbol{u} \circ f = df \circ \boldsymbol{v}$. Equivalently, by Proposition 15.5, the directional derivative $D_{\boldsymbol{u}} \in \mathrm{Der}_k(\mathcal{O}_G, \mathcal{O}_G)$ is characterized in terms of $D_{\boldsymbol{v}} \in \mathrm{Der}_k(\mathcal{O}_H, \mathcal{O}_H)$ by the properties that (1) the diagram

$$\begin{array}{ccc} \mathcal{O}_G & \xrightarrow{f^\sharp} & f_*\mathcal{O}_H \\ \downarrow{\scriptstyle D_{\boldsymbol{u}}} & & \downarrow{\scriptstyle f_* D_{\boldsymbol{v}}} \\ \mathcal{O}_G & \xrightarrow{f^\sharp} & f_*\mathcal{O}_H \end{array}$$

commutes, and (2) for all $g \in G$, the diagram

$$\begin{array}{ccc} \mathcal{O}_G & \xrightarrow{L_g^\sharp} & L_{g*}\mathcal{O}_G \\ \downarrow{\scriptstyle D_{\boldsymbol{u}}} & & \downarrow{\scriptstyle L_{g*} D_{\boldsymbol{u}}} \\ \mathcal{O}_G & \xrightarrow{L_g^\sharp} & L_{g*}\mathcal{O}_G \end{array}$$

[58] We could of course just as well have chosen to use right-invariant tangent vector fields, but note that, in general, being left-invariant and being right-invariant are different properties.

[59] The assignment of $\mathrm{Vect}(G, \mathcal{O}_G)$ to G does not extend to a functor between these categories.

commutes. More generally, if $s \in \mathrm{End}_k(\mathcal{O}_G)$ and $t \in \mathrm{End}_k(\mathcal{O}_H)$ are any k-linear morphisms, then we may ask that (1) the diagram

$$\begin{array}{ccc} \mathcal{O}_G & \xrightarrow{f^\sharp} & f_*\mathcal{O}_H \\ \downarrow{\scriptstyle s} & & \downarrow{\scriptstyle f_*t} \\ \mathcal{O}_G & \xrightarrow{f^\sharp} & f_*\mathcal{O}_H \end{array}$$

commutes, and (2) for all $g \in G$, the diagram

$$\begin{array}{ccc} \mathcal{O}_G & \xrightarrow{L_g^\sharp} & L_{g*}\mathcal{O}_G \\ \downarrow{\scriptstyle s} & & \downarrow{\scriptstyle L_{g*}s} \\ \mathcal{O}_G & \xrightarrow{L_g^\sharp} & L_{g*}\mathcal{O}_G \end{array}$$

commutes. Let us write $s \sim t$ if this is the case. We now let $\boldsymbol{v}_i \in \mathrm{Lie}(H)$, and let $\boldsymbol{u}_i = \mathrm{Lie}(f)(\boldsymbol{v}_i) \in \mathrm{Lie}(G)$ so that $D_{\boldsymbol{u}_i} \sim D_{\boldsymbol{v}_i}$. Then the composite k-linear morphisms $D_{\boldsymbol{u}_1} \circ D_{\boldsymbol{u}_2}, D_{\boldsymbol{u}_2} \circ D_{\boldsymbol{u}_1} \in \mathrm{End}_k(\mathcal{O}_H)$ and $D_{\boldsymbol{v}_1} \circ D_{\boldsymbol{v}_2}, D_{\boldsymbol{v}_2} \circ D_{\boldsymbol{v}_1} \in \mathrm{End}_k(\mathcal{O}_G)$ also satisfy that $D_{\boldsymbol{u}_1} \circ D_{\boldsymbol{u}_2} \sim D_{\boldsymbol{v}_1} \circ D_{\boldsymbol{v}_2}$ and $D_{\boldsymbol{u}_2} \circ D_{\boldsymbol{u}_1} \sim D_{\boldsymbol{v}_2} \circ D_{\boldsymbol{v}_1}$. But then

$$[D_{\boldsymbol{u}_1}, D_{\boldsymbol{u}_2}] = D_{\boldsymbol{u}_1} \circ D_{\boldsymbol{u}_2} - D_{\boldsymbol{u}_2} \circ D_{\boldsymbol{u}_1} \sim D_{\boldsymbol{v}_1} \circ D_{\boldsymbol{v}_2} - D_{\boldsymbol{v}_2} \circ D_{\boldsymbol{v}_1} = [D_{\boldsymbol{v}_1}, D_{\boldsymbol{v}_2}],$$

which shows that

$$[\boldsymbol{u}_1, \boldsymbol{u}_2] = \mathrm{Lie}(f)([\boldsymbol{v}_1, \boldsymbol{v}_2]),$$

as desired. □

Remark 15.8. Let G be a Lie group, and let us identify $\mathfrak{g} = T(G, \mathcal{O}_G)_e$. The Lie bracket on $\mathfrak{g}$ may also be defined as follows. The group structure on $(G, \mathcal{O}_G)$ induces a group structure on $T(G, \mathcal{O}_G)$, and the maps

$$\mathfrak{g} \xrightarrow{\ i\ } T(G, \mathcal{O}_G) \underset{\mathbf{0}}{\overset{p}{\rightleftarrows}} G$$

where $i = i_e$ is the kernel of $p = p_G$ and where $\mathbf{0} = \mathbf{0}_G$ is the zero section, all are morphisms of Lie groups.[60] Moreover, they exhibit the Lie group $T(G, \mathcal{O}_G)$ as the semidirect product of the Lie group G and the k-vector space $\mathfrak{g}$ considered as a Lie group under addition. This determines a morphism of Lie groups

$$G \xrightarrow{\mathrm{Ad}} \mathrm{Aut}_k(\mathfrak{g})$$

called the adjoint representation. The induced map of tangent spaces at the identity element $e \in G$ is a k-linear map

$$\mathfrak{g} \xrightarrow{\mathrm{ad}} \mathrm{End}_k(\mathfrak{g})$$

the adjunct of which is a k-linear map

$$\mathfrak{g} \otimes \mathfrak{g} \xrightarrow{[-,-]} \mathfrak{g}.$$

[60] The fact that i is a group homomorphism was the subject of the problem set for week 14.

To see that it satisfies the Jacobi identity, we argue as follows. If $f\colon H \to G$ is a morphism of Lie groups, then the map $\mathrm{Lie}(f) = df_e\colon \mathfrak{h} \to \mathfrak{g}$ satisfies

$$\mathrm{Lie}(f)([x,y]) = [\mathrm{Lie}(f)(x), \mathrm{Lie}(f)(y)]$$

for all $x, y \in \mathfrak{g}$. Moreover, if $G = \mathrm{GL}(V)$, then bracket $[-,-]$ defined here is equal to the one defined in Example 15.4. In particular, for $\mathrm{ad} = \mathrm{Lie}(\mathrm{Ad})$, we find that

$$\mathrm{ad}([x,y]) = [\mathrm{ad}(x), \mathrm{ad}(y)] = \mathrm{ad}(x) \circ \mathrm{ad}(y) - \mathrm{ad}(y) \circ \mathrm{ad}(x)$$

for all $x, y \in \mathfrak{g}$, which is equivalent to the Jacobi identity.

We next compare the representation theory of a Lie group G to that of its Lie algebra $\mathfrak{g}$. We will restrict our attention to representations (V, π), where V is a finite dimensional complex vector spaces, and where

$$G \xrightarrow{\ \pi\ } \mathrm{GL}(V)$$

is a morphism of Lie groups. If we apply the Lie algebra functor to this morphism, then we obtain a morphism of Lie algebras

$$\mathfrak{g} \xrightarrow{\ \mathrm{Lie}(\pi)\ } \mathfrak{gl}(V).$$

Hence, a representation π of a Lie group G on a finite dimensional complex vector space V gives rise to the representation $\mathrm{Lie}(\pi)$ of the Lie algebra $\mathfrak{g}$ on the same vector space V. In particular, if $\mathrm{Lie}(\pi)$ is irreducible, then π is necessarily also irreducible. We will now use the existence and uniqueness theorem for solutions to ordinary differential equations to show that if G is connected, then the representation π is completely determined by the representation $\mathrm{Lie}(\pi)$.

A global flow on a smooth manifold $(X, \mathcal{O}_X)$ is defined to be a left action

$$(\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \times (X, \mathcal{O}_X) \xrightarrow{\ \phi\ } (X, \mathcal{O}_X)$$

in the category of smooth manifolds and morphisms of smooth manifolds, of the group object $\mathbb{R} = ((\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}), +, -)$ on the object $(X, \mathcal{O}_X)$. There is a unique tangent vector field $\boldsymbol{v} \in \mathrm{Vect}(X, \mathcal{O}_X)$ that makes the diagram

$$\begin{array}{ccc} T(\mathbb{R} \times X, \mathcal{O}_{\mathbb{R}\times X}) & \xrightarrow{\ d\phi\ } & T(X, \mathcal{O}_X) \\ \uparrow{\scriptstyle \boldsymbol{w}\times 0} & & \uparrow{\scriptstyle \boldsymbol{v}} \\ (\mathbb{R} \times X, \mathcal{O}_{\mathbb{R}\times X}) & \xrightarrow{\ \phi\ } & (X, \mathcal{O}_X) \end{array}$$

commute. Indeed, let $i\colon X \to \mathbb{R} \times X$ be the inclusion defined by $i(x) = (0, x)$. Since $\phi \circ i = \mathrm{id}_X$, we are forced to define v to be the composite morphism

$$\boldsymbol{v} = \boldsymbol{v} \circ \phi \circ i = d\phi \circ (\boldsymbol{w} \times 0) \circ i,$$

and with this definition, we have

$$\boldsymbol{v} \circ \phi = d\phi \circ (\boldsymbol{w} \times 0) \circ i \circ \phi = d\phi \circ (\boldsymbol{w} \circ 0),$$

where the second non-trivial identity holds, because ϕ is an action. We say that $\boldsymbol{v}$ is the infinitesimal generator of the flow ϕ.

Conversely, given $\boldsymbol{v} \in \mathrm{Vect}(X, \mathcal{O}_X)$, the existence and uniqueness theorem for solutions to ordinary differential equations shows that there exists a morphism of smooth manifolds $\phi\colon (U, \mathcal{O}_{\mathbb{R}\times X}|_U) \to (X, \mathcal{O}_X)$ with $\{0\} \times X \subset U \subset \mathbb{R} \times X$ open which makes the diagram

$$\begin{array}{ccc} T(U, \mathcal{O}_{\mathbb{R}\times X}|_U) & \xrightarrow{d\phi} & T(X, \mathcal{O}_X) \\ \uparrow{\scriptstyle (\boldsymbol{w}\times 0)|_U} & & \uparrow{\scriptstyle \boldsymbol{v}} \\ (U, \mathcal{O}_{\mathbb{R}\times X}|_U) & \xrightarrow{\phi} & (X, \mathcal{O}_X) \end{array}$$

commute and satisfies $\phi(0, x) = x$ and $\phi(s, \phi(t, x)) = \phi(s{+}t, x)$ whenever this makes sense. We say that ϕ is a local flow with infinitesimal generator $\boldsymbol{v}$. In particular, if there exists a global flow ϕ with infinitesimal generator $\boldsymbol{v}$, then ϕ is uniquely determined by $\boldsymbol{v}$. If this is the case, then we say that $\boldsymbol{v}$ is complete.

If G is a Lie group, and if $v \in \mathfrak{g}$ is a left-invariant vector field, then, by using the group structure, one shows that every local flow with infinitesimal generator $\boldsymbol{v}$ extends uniquely to a global flow $\phi = \phi_{\boldsymbol{v}}$ with infinitesimal generator $\boldsymbol{v}$. We define the exponential map of the Lie group G to be the map

$$\mathfrak{g} \xrightarrow{\exp} G$$

given by $\exp(\boldsymbol{v}) = \phi_{\boldsymbol{v}}(1, e)$. We remark that exp is not a group homomorphism, unless the Lie algebra $\mathfrak{g}$ is abelian.

Theorem 15.9. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. The exponential map is a morphism of smooth manifolds*

$$(\mathfrak{g}, \mathcal{O}_{\mathfrak{g}}^{\mathrm{sm}}) \xrightarrow{\exp} (G, \mathcal{O}_G).$$

Moreover, it is étale at $0 \in \mathfrak{g}$.[61]

Proof. The structure of group object on the smooth manifold $(G, \mathcal{O}_G)$ gives rise to a structure of group object on $T(G, \mathcal{O}_G)$. Moreover, there is a left-invariant tangent vector field $\boldsymbol{u}$ on the Lie group $T(G, \mathcal{O}_G)$ such that for every left-invariant tangent vector field $\boldsymbol{v}$ on $(G, \mathcal{O}_G)$, the diagram

$$\begin{array}{ccc} T(G, \mathcal{O}_G) & \xrightarrow{d\boldsymbol{v}} & T(T(G, \mathcal{O}_G)) \\ \uparrow{\scriptstyle \boldsymbol{v}} & & \uparrow{\scriptstyle \boldsymbol{u}} \\ (G, \mathcal{O}_G) & \xrightarrow{\boldsymbol{v}} & T(G, \mathcal{O}_G) \end{array}$$

commutes. Now, there is a global flow $\varphi_{\boldsymbol{u}}$ on $T(G, \mathcal{O}_G)$ with infinitesimal generator $\boldsymbol{u}$, and it follows from the uniqueness of solutions to ordinary differential equations that for every $\boldsymbol{v} \in \mathrm{Vect}(G, \mathcal{O}_G)$ with global flow $\varphi_{\boldsymbol{v}}$ on $(G, \mathcal{O}_G)$, the diagram

$$\begin{array}{ccc} (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \times (G, \mathcal{O}_G) & \xrightarrow{\varphi_{\boldsymbol{v}}} & (G, \mathcal{O}_G) \\ \downarrow{\scriptstyle \mathrm{id}\times \boldsymbol{v}} & & \downarrow{\scriptstyle \boldsymbol{v}} \\ (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \times T(G, \mathcal{O}_G) & \xrightarrow{\varphi_{\boldsymbol{u}}} & T(G, \mathcal{O}_G) \end{array}$$

[61] The exponential map may have critical points. One can show that $x \in \mathfrak{g}$ is a critical point for exp if and only if some $0 \neq \lambda \in 2\pi i\mathbb{Z} \subset \mathbb{C}$ is an eigenvalue of $\mathrm{ad}(x) \in \mathrm{End}_k(\mathfrak{g})$.

commutes. Therefore, the exponential map is equal to the composite map

$$(\mathfrak{g}, \mathcal{O}_{\mathfrak{g}}^{\mathrm{sm}}) \xrightarrow{i_1 \times i_e} (\mathbb{R}, \mathcal{O}_{\mathbb{R}}^{\mathrm{sm}}) \times T(G, \mathcal{O}_G) \xrightarrow{\varphi_{\boldsymbol{u}}} T(G, \mathcal{O}_G) \xrightarrow{p_G} (G, \mathcal{O}_G)$$

and since each of these three maps is a morphism of smooth manifolds, so is the exponential map. Finally, it follows immediately from the definition that

$$\mathfrak{g} = T(\mathfrak{g}, \mathcal{O}_{\mathfrak{g}})_0 \xrightarrow{d\exp_0} T(G, \mathcal{O}_G)_e$$

is equal to the isomorphism ϵ_G in Proposition 15.7, and therefore, the inverse function theorem shows that exp is étale at $0 \in \mathfrak{g}$ as stated. □

Corollary 15.10. *If G is a connected Lie group, then every $g \in G$ can be written as a product $g = \exp(x_1) \cdots \exp(x_n)$ with $n \geq 0$ and $x_1, \dots, x_n \in \mathfrak{g}$.*

Proof. By Theorem 15.9, there exists open subsets $0 \in U \subset \mathfrak{g}$ and $e \in V \subset G$ such that $\exp|_U \colon (U, \mathcal{O}_U^{\mathrm{sm}}) \to (V, \mathcal{O}_G|_V)$ is a diffeomorphism. Hence, it suffices to show that the subgroup $H \subset G$ generated by V is equal to G.[62] Since $V \subset G$ is open, so is $H \subset G$. But then $gH \subset G$ is open, for all $g \in G$, which implies that

$$H = G \smallsetminus (\textstyle\bigcup_{g \in G \smallsetminus H} gH) \subset G$$

is closed. Since G is connected, we conclude that $H = G$ as desired. □

Corollary 15.11. *Let G and H be Lie groups. If H is connected, then the map*

$$\operatorname{Hom}(H, G) \xrightarrow{\mathrm{Lie}} \operatorname{Hom}(\mathfrak{h}, \mathfrak{g})$$

is injective.

Proof. Let $f \colon H \to G$ be a morphism of Lie groups. The diagram

$$\begin{array}{ccc} \mathfrak{h} & \xrightarrow{\exp_H} & H \\ \Big\downarrow{\scriptstyle \mathrm{Lie}(f)} & & \Big\downarrow{\scriptstyle f} \\ \mathfrak{g} & \xrightarrow{\exp_G} & G \end{array}$$

commutes, by naturality of the exponential map. By Corollary 15.10, every element of H is a product of elements of $\exp_H(\mathfrak{h}) \subset H$. Since f is a group homomorphism, this implies that it is uniquely determined by the map $\mathrm{Lie}(f)$. □

We use the last corollary to show that if π_1 and π_2 are two finite dimensional real or complex representations of a connected Lie group G, then $\pi_1 \simeq \pi_2$ if and only if $\mathrm{Lie}(\pi_1) \simeq \mathrm{Lie}(\pi_2)$. In effect, we prove the following more precise result.

Corollary 15.12. *Let $\pi_1 \colon G \to \mathrm{GL}(V_1)$ and $\pi_2 \colon G \to \mathrm{GL}(V_2)$ be representations of a connected Lie group on finite dimensional real or complex vector spaces. A linear isomorphism $f \colon V_1 \to V_2$ intertwines between π_1 and π_2 if and only if it intertwines between $\mathrm{Lie}(\pi_1)$ and $\mathrm{Lie}(\pi_2)$.*

[62] Here we also use that $\exp(x)^{-1} = \exp(-x)$, since $[x, -x] = -[x, x] = 0$.

Proof. That f intertwines between π_1 and π_2 means that the diagram of Lie groups

$$\begin{array}{ccc} & \xrightarrow{\pi_1} & \mathrm{GL}(V_1) \\ G & & \big\downarrow{\scriptstyle c_f} \\ & \xrightarrow{\pi_2} & \mathrm{GL}(V_2) \end{array}$$

where $c_f(h) = f \circ h \circ f^{-1}$, commutes. But then the diagram of Lie algebras

$$\begin{array}{ccc} & \xrightarrow{\mathrm{Lie}(\pi_1)} & \mathfrak{gl}(V_1) \\ \mathfrak{g} & & \big\downarrow{\scriptstyle \mathrm{Lie}(c_f)} \\ & \xrightarrow{\mathrm{Lie}(\pi_2)} & \mathfrak{gl}(V_2) \end{array}$$

commutes, and since $\mathrm{Lie}(c_f)(h) = f \circ h \circ f^{-1}$, this shows that f intertwines between $\mathrm{Lie}(\pi_1)$ and $\mathrm{Lie}(\pi_2)$. This part of the statement only uses that $\mathrm{Lie}(-)$ is a functor and not that G is connected. Conversely, if f intertwines between $\mathrm{Lie}(\pi_1)$ and $\mathrm{Lie}(\pi_2)$, then the bottom diagram commutes, and since G is connected, this implies, by Corollary 15.11 that the top diagram commutes. □

This is marvelous! To a large extent, we have replaced the differential geometric problem of finding representations of a Lie group with the linear algebraic problem of finding representations of its Lie algebra. We illustrate this for $G = SU(2)$, which is a compact connected Lie group. We have already proved that for every integer $n \geq 0$, the representation π_n given by the nth symmetric power

$$\pi_n = \mathrm{Sym}^n_{\mathbb{C}}(\pi)$$

of the standard representation π of $SU(2)$ on $V = \mathbb{C}^2$ is an irreducible representation of dimension $n+1$. The associated representation of the Lie algebra $\mathfrak{g} = \mathfrak{su}(2)$ is a morphism of real Lie algebras

$$\mathfrak{su}(2) \xrightarrow{\mathrm{Lie}(\pi_n)} f_*\mathfrak{gl}(\mathrm{Sym}^n_{\mathbb{C}}(V))$$

from the real Lie algebra $\mathfrak{su}(2)$ to the real Lie algebra obtained by restriction of scalars along $f\colon \mathbb{R} \to \mathbb{C}$ from the complex Lie algebra $\mathfrak{gl}(\mathrm{Sym}^n_{\mathbb{C}}(V))$. The adjunct of $\mathrm{Lie}(\pi_n)$ is a morphism of complex Lie algebras

$$\mathfrak{su}(2)_{\mathbb{C}} = f^*\mathfrak{su}(2) \xrightarrow{\widetilde{\mathrm{Lie}(\pi_n)}} \mathfrak{gl}(\mathrm{Sym}^n_{\mathbb{C}}(V)).$$

We have earlier identified $\mathfrak{su}(2)$ with the real vector space of traceless skew-hermitian complex 2×2-matrices. It has a basis given by the family (A_1, A_2, A_3), where

$$A_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad A_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

The Lie bracket on $\mathfrak{su}(2)$ is given by $[A, B] = AB - BA$. Similarly, the complex Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ of the complex Lie group $\mathrm{SL}_2(\mathbb{C})$ is given by the complex vector space of all traceless complex 2×2-matrices with the Lie bracket given by the same formula. So the inclusion of the set of traceless skew-hermitian complex 2×2-matrices in the set of all traceless complex 2×2-matrices defines a morphism of real Lie algebras $\mathfrak{su}(2) \to f_*\mathfrak{sl}(2,\mathbb{C})$, the adjunct of which is a morphism

$$\mathfrak{su}(2)_\mathbb{C} = f^*\mathfrak{su}(2) \longrightarrow \mathfrak{sl}(2,\mathbb{C}).$$

of complex Lie algebras. We claim that the latter map is an isomorphism. Indeed, one readily verifies that the family (A_1, A_2, A_3) is a basis of both complex vector spaces. Moreover, under this identification, the representation

$$\mathfrak{sl}(2,\mathbb{C}) \xrightarrow{\widetilde{\mathrm{Lie}(\pi_n)}} \mathfrak{gl}(\mathrm{Sym}^n_\mathbb{C}(V))$$

is equivalent to the nth symmetric power of the standard representation of the complex Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ on V.

Now, the complex vector space $\mathfrak{sl}(2,\mathbb{C})$ has the much more convenient basis given by the family (X, H, Y), where[63]

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{and} \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Indeed, in this basis, the Lie bracket is given by the simple formulas

$$[X,Y] = H, \qquad [H,X] = 2X, \qquad \text{and} \qquad [H,Y] = -2Y.$$

The complex representations of $\mathfrak{sl}(2,\mathbb{C})$ can be completely understood, and this, in turn, is the starting point for understanding the representation theory of all complex reductive Lie algebras and Lie groups. Serre's book [5] is a very readable introduction to this beautiful theory.

Let $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(V)$ be a representation on a complex vector space V, which, at the moment, we do not assume to be finite dimensional. We write $V^\lambda \subset V$ for the eigenspace corresponding to the eigenvalue $\lambda \in \mathbb{C}$ of $\pi(H)\colon V \to V$, and we say that $\boldsymbol{x} \in V^\lambda$ has weight λ. The canonical map

$$\bigoplus_{\lambda\in\mathbb{C}} V^\lambda \longrightarrow V$$

is always injective. If the dimension of V is finite, then it is also surjective, but, in general, this is not the case. If x has weight λ, then the calculation

$$\begin{aligned}
(\pi(H)\circ\pi(X))(x) &= \pi([H,X])(\boldsymbol{x}) + (\pi(X)\circ\pi(H))(\boldsymbol{x}) \\
&= \pi(2X)(\boldsymbol{x}) + \pi(X)(\lambda\boldsymbol{x}) \\
&= (\lambda+2)\pi(X)(\boldsymbol{x}) \\
(\pi(H)\circ\pi(Y))(\boldsymbol{x}) &= \pi([H,Y])(\boldsymbol{x}) + (\pi(Y)\circ\pi(H))(\boldsymbol{x}) \\
&= -\pi(2Y)(\boldsymbol{x}) + \pi(Y)(\lambda\boldsymbol{x}) \\
&= (\lambda-2)\pi(Y)(\boldsymbol{x})
\end{aligned}$$

shows that $\pi(X)(\boldsymbol{x})$ has weight $\lambda+2$ and that $\pi(Y)(\boldsymbol{x})$ has weight $\lambda-2$. We say that an element $\boldsymbol{e} \in V$ is primitive of weight λ if $\boldsymbol{e} \neq \boldsymbol{0}$ and if $\pi(H)(\boldsymbol{e}) = \lambda\boldsymbol{e}$ and $\pi(X)(\boldsymbol{e}) = \boldsymbol{0}$.

[63] The alternative notation e, h, and f for these matrices is also common.

Theorem 15.13. *Let π be an irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$ on a complex vector space V of finite dimension $n+1$. The following hold.*

(1) *There exists a primitive element $\boldsymbol{e} \in V$ of weight $\lambda = n$.*
(2) *The family $(\boldsymbol{e}_0, \ldots, \boldsymbol{e}_n)$, where $\boldsymbol{e}_k = \pi(Y)^k(\boldsymbol{e})/k!$, is a basis of V.*
(3) *In this basis, the representation π is given by*

$$\pi(H)(\boldsymbol{e}_k) = (\lambda - 2k)\boldsymbol{e}_k$$

$$\pi(X)(\boldsymbol{e}_k) = \begin{cases} \mathbf{0} & \text{if } k = 0 \\ (\lambda - k + 1)\boldsymbol{e}_{k-1} & \text{if } 0 < k \leq n \end{cases}$$

$$\pi(Y)(\boldsymbol{e}_k) = \begin{cases} (k+1)\boldsymbol{e}_{k+1} & \text{if } 0 \leq k < n \\ \mathbf{0} & \text{if } k = n. \end{cases}$$

Conversely, the formulas (3) *define an irreducible representation of the complex Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ on a complex vector space with basis $(\boldsymbol{e}_0, \ldots, \boldsymbol{e}_n)$.*

Proof. Since $\mathbb{C}$ is algebraically closed, there exists an eigenvector $\boldsymbol{x} \in V$ of the linear endomorphism $\pi(H)\colon V \to V$. The vectors $\pi(X)^k(\boldsymbol{x})$ with $k \geq 0$ are either eigenvectors of $\pi(H)$ or zero. Since V is finite dimensional, there exists a maximal $k \geq 0$ such that $\boldsymbol{e} = \pi(X)^k(\boldsymbol{x}) \neq \mathbf{0}$ and $\pi(X)(\boldsymbol{e}) = \mathbf{0}$. Hence, this element $\boldsymbol{e}$ is a primitive element of some weight $\lambda \in \mathbb{C}$.

Now, for all $k \geq 0$, we consider the elements $\boldsymbol{e}_k \in V$ defined by

$$\boldsymbol{e}_k = \pi(Y)^k(\boldsymbol{e})/k!,$$

and we also set $\boldsymbol{e}_{-1} = \mathbf{0}$. We claim that for all $k \geq 0$, the following hold:

(a) $\pi(H)(\boldsymbol{e}_k) = (\lambda - 2k)\boldsymbol{e}_k$
(b) $\pi(Y)(\boldsymbol{e}_k) = (k+1)\boldsymbol{e}_{k+1}$
(c) $\pi(X)(\boldsymbol{e}_k) = (\lambda - k + 1)\boldsymbol{e}_{k-1}$.

Indeed, (b) holds, by definition, and (a) holds by the observation that $\pi(Y)$ lowers weight by 2. We prove (c) by induction on $k \geq -1$, the case $k = -1$ being trivial. Assuming that (c) holds for $k < m$, the calculation

$$\begin{aligned} m\pi(X)(\boldsymbol{e}_m) &= (\pi(X) \circ \pi(Y))(\boldsymbol{e}_{m-1}) \\ &= \pi([X,Y])(\boldsymbol{e}_{m-1}) + (\pi(Y) \circ \pi(X))(\boldsymbol{e}_{m-1}) \\ &= \pi(H)(\boldsymbol{e}_{m-1}) + (\lambda - m + 2)\pi(Y)(\boldsymbol{e}_{m-2}) \\ &= (\lambda - 2m + 2 + (\lambda - m + 2)(m-1))\boldsymbol{e}_{m-1} \\ &= m(\lambda - m + 1)\boldsymbol{e}_{m-1}, \end{aligned}$$

shows that (c) holds for $k = m$. This proves the claim.

Next, if the elements $\boldsymbol{e}_k$ with $k \geq 0$ all are non-zero, then $(\boldsymbol{e}_k)_{k \geq 0}$ is a family of eigenvectors for $\pi(H)$ with pairwise distinct eigenvalues. But then this family is linearly independent, which is not possible, because V is finite dimensional. We also observe from (b) that $\boldsymbol{e}_k = \mathbf{0}$ implies that $\boldsymbol{e}_{k+1} = \mathbf{0}$. So there exists $m \geq 0$ such that $\boldsymbol{e}_k \neq \mathbf{0}$ for $0 \leq k \leq m$ and $\boldsymbol{e}_k = \mathbf{0}$ for $k > m$. Moreover, by (c), we have

$$\mathbf{0} = \pi(X)(\boldsymbol{e}_{m+1}) = (\lambda - m)\boldsymbol{e}_m,$$

so we conclude that $\lambda = m$.

Finally, it follows immediately from (a)–(c) that the subspace $W \subset V$ spanned by $(\boldsymbol{e}_0, \dots, \boldsymbol{e}_m)$ is π-invariant. It is also non-zero, since $\mathbf{0} \neq \boldsymbol{e} = \boldsymbol{e}_0 \in W$, and since (V, π) was assumed to be irreducible, we conclude that $W = V$ and $m = n$. ☐

Corollary 15.14. *Let $n \geq 0$ be an integer.*

(1) *The complex Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ has a unique isomorphism class of irreducible complex representations of dimension $n + 1$.*

(2) *The real Lie algebra $\mathfrak{su}(2)$ has a unique isomorphism class of irreducible complex representations of dimension $n + 1$.*

(3) *The real Lie group $SU(2)$ has a unique isomorphism class of irreducible complex representations of dimension $n + 1$.*

Proof. First, (1) follows immediately from Theorem 15.13. Second, (2) follows from (1) and from the extension-of-scalars/restriction-of-scalars adjunction, since we have an isomorphism of complex Lie algebras $\mathfrak{su}(2)_{\mathbb{C}} \to \mathfrak{sl}(2, \mathbb{C})$. Finally, we conclude from (2) and from Corollary 15.12 that the connected Lie group $SU(2)$ has at most one isomorphism class of irreducible complex representations of dimension $n + 1$. But we have already proved that $\pi_n \colon SU(2) \to \mathrm{GL}(\mathrm{Sym}^n_{\mathbb{C}}(V))$ is such a representation, so (3) follows. ☐

Example 15.15. The adjoint representation

$$SU(2) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{su}(2))$$

is a 3-dimensional real representation. One can show that the adjoint representation is irreducible, and that its complexification

$$SU(2) \xrightarrow{\mathrm{Ad}_{\mathbb{C}}} \mathrm{GL}(\mathfrak{su}(2)_{\mathbb{C}})$$

also is irreducible. Therefore, by Corollary 15.14, it is isomorphic to the symmetric square π_2 of the standard representation $\pi = \pi_1$.

In elementary particle physics, a gauge theory begins with a compact Lie group G of "internal symmetries," and the complexified adjoint representation

$$G \xrightarrow{\mathrm{Ad}_{\mathbb{C}}} \mathrm{GL}(\mathfrak{g}_{\mathbb{C}})$$

provides the "gauge bosons" of the theory; they are the elements of a basis of the complex vector space $\mathfrak{g}_{\mathbb{C}}$. For example, physicists write (W^+, W^0, W^-) for the basis (X, H, Y) of $\mathfrak{su}_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C})$. Its elements are the W-bosons, which mediate the weak force. Let me explain what this means. The "elementary fermions" in the gauge theory are basis elements of certain irreducible finite dimensional complex representations of G. The selection of the irreducible representations that should be considered the "elementary fermions" of the theory, however, is entirely empirical. If $\pi \colon G \to \mathrm{GL}(V)$ is an irreducible finite dimensional complex representation, then

$$\mathfrak{g}_{\mathbb{C}} \xrightarrow{\mathrm{Lie}(\pi)_{\mathbb{C}}} \mathfrak{gl}(V)$$

is a representation of the complexified Lie algebra on V, and moreover, this map is intertwining with respect to the G-representations $\mathrm{Ad}_{\mathbb{C}}$ on the domain and $\mathrm{End}(\pi)$ on the target. It is by means of this Lie algebra representation that the gauge bosons acts on the elementary fermions. See the article [1] by Baez–Huerta for more on this.

References

[1] J. Baez and J. Huerta, *The algebra of grand unified theories*, Bull. Amer. Math. Soc. **47** (2010), 483–552.

[2] N. Bourbaki, *Integration II. Chapters 7–9. Translated from the 1963 and 1969 French originals by Sterling K. Berberian.*, Elements of Mathematics, Springer-Verlag, Berlin, 2004.

[3] J. Hilgert and K.-H. Neeb, *Structure and geometry of Lie groups*, Springer Monographs in Mathematics, Springer, New York, 2012.

[4] E. Kowalski, *An Introduction to the Representation Theory of Groups*, Graduate Studies in Mathematics, vol. 135, Amer. Math. Soc., Providence, RI, 2014.

[5] J.-P. Serre, *Complex semisimple Lie algebras. Translated from the French by G. A. Jones*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2001.