COMPLETE REDUCIBILITY OF REPRESENTATIONS OF COMPACT GROUPS

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 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 $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.[1]

Remark 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 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. 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[2] $\mathrm{O}(V, \langle -, - \rangle) \subset \mathrm{GL}(V)$ (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.

---

[1] 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.

[2] 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.

Proposition 4. *Orthogonal (resp. unitary) representations are completely reducible.*

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. $\square$

We first consider finite groups.

Theorem 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. $\square$

Definition 6. A topological group is a group G is 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 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 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 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 [1, Chapter 7, §1, No. 2, Theorem 1].

Theorem 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 11. (1) If G is 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 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 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 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 4 shows that it is completely reducible. $\square$

Remark 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 13. Let W be a finite dimensional real vector space, and let μ be a Lebesgue measure[3] 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 13.

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

Lemma 15. *If $K \subset W$ is compact,*[4] *then so is* $\mathrm{conv}(K) \subset W$.

Proof. Let $n = \dim_{\mathbb{R}}(W)$. A classical theorem of Carathéodory states that for every $\boldsymbol{w} \in \mathrm{conv}(K)$ is a convex combination of at most $n+1$ points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_n \in K$. 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 $\boldsymbol{x}_0, \dots, \boldsymbol{x}_n \in K$, $a_0, \dots, a_n \in [0,1]$, 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 $(\boldsymbol{x}_0, \dots, \boldsymbol{x}_n, a_0, \dots, a_n)$ 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 $(a_0, \dots, a_n)$ 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 16. *If $K \subset W$ is a compact subset with $\mu(K) > 0$, then*

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

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

[4] 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

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

for all $r \geq 1$. But by Lemma 15, $\mathrm{conv}(K) \subset W$ is a compact subset of a Hausdorff space, so $c(K) \in \mathrm{conv}(K)$ as stated. □

Proof (of Theorem 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

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

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 $\langle -, - \rangle_0 \in K_0 \subset I(V)$, 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 $i \colon K \to B^+(V)$ 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

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

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

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

is an inner product. By Lemma 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 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. □

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:[5]

(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

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

with $\boldsymbol{x}, \boldsymbol{y} \in V$ and $a \in \mathbb{R} \subset \mathbb{C}$. The subset $J(V) \subset H^+(V)$ 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 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,*

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

[5] 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

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

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

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

So it will suffice to prove that

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

for all $1 \le i \le s$. 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 $a_0, \ldots, a_e \in [0,1]$ 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 $a_0, \ldots, a_e$ be equal to zero. By rearranging the $\boldsymbol{x}_i$, if necessary, we can assume that $a_0 = 0$. But then

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

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

References

[1] 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.