Math 761 **Fall 2015**
Homework 4 **Drew Armstrong**

Problem 1. Burnside's Lemma. Let X be a G-set and for all $g \in G$ define the set

$$\mathsf{Fix}(g) := \{x \in X : g(x) = x\} \subseteq X.$$

(a) If G and X are finite, prove that

$$\sum_{g\in G} |\mathsf{Fix}(g)| = \sum_{x \in X} |\mathsf{Stab}(x)|.$$

(b) Let X/G be the set of orbits. Use part (a) to prove that

$$|X/G| = \frac{1}{|G|}\sum_{g \in G} |\mathsf{Fix}(g)|.$$

Proof. For part (a), consider the set $S := \{(g, x) \in G \times X : g(x) = x\}$. By counting the pairs $(g, x) \in S$ in two ways we obtain

$$\sum_{g\in G} |\mathsf{Fix}(g)| = |S| = \sum_{x \in X} |\mathsf{Stab}(x)|.$$

Choosing g first gives the first equation and choosing x first gives the second equation.

For part (b), let $O_1, O_2, \ldots, O_n$ be the set of orbits, so that $n = |X/G|$. Then from part (a) and the orbit-stabilizer theorem we have

$$\begin{aligned}
\sum_{g\in G} |\mathsf{Fix}(g)| &= \sum_{x\in X} |\mathsf{Stab}(x)| \\
&= \sum_{x\in X} \frac{|G)|}{|\mathsf{Orb}(x)|} \\
&= \sum_{i=1}^{n} \sum_{x\in O_i} \frac{|G|}{|\mathsf{Orb}(x)|} \\
&= \sum_{i=1}^{n} \sum_{x\in O_i} \frac{|G|}{|O_i|} \\
&= \sum_{i=1}^{n} |O_i| \frac{|G|}{|O_i|} \\
&= \sum_{i=1}^{n} |G| \\
&= n \cdot |G| \\
&= |X/G| \cdot |G|.
\end{aligned}$$

□

[Remark: Burnside's Lemma appears in Burnside's "Theory of groups of finite order" (1911), where he attributes it to Frobenius. It was also known to Cauchy in 1845. This is another example of a mathematical concept being named for the last person to discover it.]

Problem 2. The Dodecahedron. Let D be the group of rotational symmetries of a regular dodecahedron.

(a) Describe the conjugacy classes of D and use this to prove that D is simple. [Hint: Any normal subgroup is a union of conjugacy classes.]

(b) Compute the number of distinguishable ways to color the faces of a dodecahedron with k colors. [Hint: Let X be the set of all colorings, so that $|X| = k^{12}$. Many of these colorings are indistinguishable after rotation so we really want to know the number of orbits $|X/D|$. Use part (a) and Burnside's Lemma.]

(c) Prove that D is isomorphic to the alternating group A_5. [Hint: There are five cubes that can be inscribed in a dodecahedron. The action of D defines a nontrivial homomorphism $\varphi : D \to S_5$. Composing this with the "sign" homomorphism $\sigma : S_5 \to \{\pm 1\}$ gives a homomorphism $\sigma\varphi : D \to \{\pm 1\}$. Since D is simple the first homomorphism must be injective and the second must be trivial.]

Proof. For part (a), we will think of D as a subgroup of $SO(3)$. Note that two elements of D are conjugate in $GL_3(\mathbb{R})$ if and only if they represent the same linear transformation in two different bases. If, moreover, they are conjugate in $SO(3)$ then they represent the same linear transformation after a rotation, and if they are conjugate in D then they represent the same linear transformation after a rotational symmetry of the dodecahedron.

Using this idea we can describe the conjugacy classes as follows:

Name of Class	Size of Class	Geometric Description
C_1	1	identity element
C_2	20	rotate by $\pm 2\pi/3$ around vertex
C_3	15	rotate by π around edge
C_4	12	rotate by $\pm 2\pi/5$ around face
C_5	12	rotate by $\pm 4\pi/5$ around face

To count the elements we note that a rotation must be shared by an opposite pair of vertices/edges/faces. Clearly the five classes described are inequivalent because symmetries of the dodecahedron must take vertices/edges/faces to vertices/edges/faces, respectively. The only possible issue is to explain why rotation by $\pm 2\pi/5$ around a face is not conjugate to rotation by $\pm 4\pi/5$ around a face. To see this, note that the trace of a rotation by angle θ is $1 + 2\cos\theta$. Since conjugation preserves trace, and since $1 + 2\cos(2\pi/5) \neq 1 + 2\cos(4\pi/5)$, we conclude that the rotations are not conjugate.

To prove that D is simple, note that any normal subgroup $N \trianglelefteq D$ must be a union of conjugacy classes (including the identity class). Thus $|N|$ is a sum of 1 together with a subset of the numbers

$$20, 15, 12, 12.$$

But by Lagrange's Theorem we also know that $|N|$ divides $|D| = 60$. One can check that these two conditions imply that $|N| = 1$ (i.e. $N = 1$) or $|N| = 60$ (i.e. $N = D$).

For part (b), let X be the set of colorings of the faces of a fixed dodecahedron using k colors. Since there are 12 faces we have $|X| = k^{12}$. But many of these colorings are indistinguishable after rotation, so we are really interested in the number of orbits $|X/D|$. By Burnside's Lemma we only need to count the number of colorings fixed by each conjugacy class of D. We can think of each element $g \in D$ as a permutation of the 12 faces and for $x \in X$ we have $g(x) = x$ if and only if the coloring x is constant on each cycle of g.

- For $g \in C_1$ there are 12 cycles of faces, hence $|\mathsf{Fix}(g)| = k^{12}$.
- For $g \in C_2$ there are 4 cycles of faces, hence $|\mathsf{Fix}(g)| = k^4$.

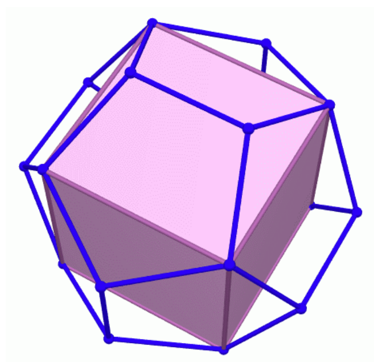
- For $g \in C_3$ there are 6 cycles of faces, hence $|\mathsf{Fix}(g)| = k^6$.
- For $g \in C_4$ there are 4 cycles of faces, hence $|\mathsf{Fix}(g)| = k^4$.
- For $g \in C_5$ there are 4 cycles of faces, hence $|\mathsf{Fix}(g)| = k^4$.

(It helps to have dodecahedron to play with.) Then from Burnside's Lemma we conclude that

$$\begin{aligned}|X/D| &= \frac{1}{|D|}\sum_{g\in G}|\mathsf{Fix}(g)| \\ &= \frac{1}{60}\left(1\cdot k^{12} + 20\cdot k^4 + 15\cdot k^6 + 12\cdot k^4 + 12\cdot k^4\right) \\ &= \frac{1}{60}k^4(k^8 + 15k^2 + 44).\end{aligned}$$

For example, there are $96 = \frac{1}{60}2^4(2^8 + 15\cdot 2^2 + 44)$ different black and white dodecahedra.

For part (c), consider the five cubes inscribed in a dodecahedron. Here is one of them:



The action of D on the set of cubes induces a homomorphism $\varphi : D \to S_5$. Since D is simple we must have either $\ker\varphi = D$ or $\ker\varphi = 1$. We know that the homomorphism is not trivial (the cubes **do** get permuted) so we conclude that φ is injective. Composing with the sign map $\sigma : S_5 \to \{\pm 1\}$ gives another homomorphism $\sigma\varphi : D \to \{\pm 1\}$. Again, since D is simple we must have $\ker(\sigma\varphi) = 1$ or $\ker(\sigma\varphi) = D$. Since D is bigger than $\{\pm 1\}$ the map can not be injective, so we must have $\ker(\sigma\varphi) = D$. Putting these two facts together gives

$$D \approx \operatorname{im}\varphi \subseteq \ker\sigma = A_5.$$

Finally, since $|D| = |A_5|$ we must have $D \approx A_5$. □

[Remark: I really like this proof for the simplicity of A_5. Unfortunately, I don't know a similarly nice proof for the simplicity of A_n when $n \geq 6$. We have no choice but to fiddle with 3-cycles.]

Problem 3. Affine Space. What is space? In general it is possible to "subtract points" to obtain a vector, but it is not possible to "add points" unless we fix an arbitrary basepoint. Let V be a vector space. We say that A is an affine space over V if there exists a "subtraction function" $[-,-] : A \times A \to V$ satisfying the following two properties:

- $[p,-] : A \to V$ is a bijection for all $p \in A$,
- $[p,q] + [q,r] = [p,r]$ for all $p,q,r \in A$.

(a) We say that a group action is free if all stabilizers are trivial and we say it is transitive if every orbit is the full set. We say that an action is regular if it is free and transitive. Prove that an affine space over a vector space V is the same thing as a regular V-set (thinking of V as an abelian group).

(b) Let A be an affine space over V and denote the induced regular action of V on A by $v(p) =$ "$p + v$". We say that a function $f : A \to A$ is affine if there exists a **linear** function $df : V \to V$ such that for all points $p \in A$ and vectors $v \in V$ we have

$$f(p + v) = f(p) + df(v).$$

In this case show that $df([p, q]) = [f(p), f(q)]$ for all $p, q \in A$, so that df is uniquely determined by f (we call it the differential of f). Prove that f is invertible if and only if df is invertible, in which case we have $d(f^{-1}) = (df)^{-1}$.

(c) Let $\mathsf{GA}(V)$ be the group of invertible affine functions $A \to A$ (called the general affine group of V). Prove that we have an isomorphism

$$\mathsf{GA}(V) \approx V \rtimes \mathsf{GL}(V)$$

where $\mathsf{GL}(V)$ acts on V in the obvious way. [Hint: Show that the differential map $d : \mathsf{GA}(V) \to \mathsf{GL}(V)$ is a group homomorphism with kernel isomorphic to V. Show that "choosing an origin" $o \in A$ defines a section $s : \mathsf{GL}(V) \to \mathsf{GA}(V)$.]

Proof. For part (a), let $[-,-] : A \times A \to V$ be a valid subtraction function. Note that for all points $p \in A$ we have $[p, p] = [p, p] + [p, p] = 2[p, p]$ and hence $[p, p] = 0$. Then for all points $p, q \in A$ we have $[p, q] + [q, p] = [p, p] = 0$ and hence $[p, q] = -[q, p]$.

Now consider a point $p \in A$ and a vector $v \in V$. Since $[p, -] : A \to V$ is a bijection, there exists a unique point, say $v(p) \in A$, satisfying the equation

$$[p, v(p)] = v.$$

We want to show that the function $V \times A \to A$ defined by $(v, p) \mapsto v(p)$ is a regular action. First we'll show that it's an action:

- For all points $p \in A$ we have $[p, 0(p)] = 0$ by definition. But we also have $[p, p] = 0$, so the injectivity of $[p, -]$ implies that $0(p) = p$.
- For all points $p \in A$ and vectors $u, v \in V$ we have $[p, v(p)] = v$ and $[v(p), u(v(p))] = u$ by definition. It follows that

$$[p, u(v(p))] = [p, v(p)] + [v(p), u(v(p))] = v + u,$$

  and hence $u(v(p)) = (v + u)(p)$.

Next we'll show that the action is regular. Given a point $p \in A$ we define

$$\mathsf{Orb}_V(p) := \{v(p) \in A : v \in V\},$$
$$\mathsf{Stab}_V(p) := \{v \in V : v(p) = p\}.$$

Since $[p, -]$ is surjective we know that for all $q \in A$ there exists $v \in V$ such that $[p, q] = v$ and hence $q = v(p)$. We conclude that $\mathsf{Orb}_V(p) = A$. Now let $v \in \mathsf{Stab}_V(p)$. Since $v(p) = p$ we have $v = [p, v(p)] = [p, p] = 0$ and hence $\mathsf{Stab}_V(p) = 0$.

Conversely, let $(v, p) \mapsto v(p)$ be a regular action of V on A and consider any two points $p, q \in A$. Since $q \in A = \mathsf{Orb}_V(p)$ there exists a vector $v \in V$ such that $v(p) = q$. If $u \in V$ is any other vector such that $u(p) = q$ then since $u(p) = v(p)$ we have $(u - v)(p) = (v - v)(p) = 0(p) = p$ and hence $u - v \in \mathsf{Stab}_V(p)$. Since $\mathsf{Stab}_V(p) = 0$ this implies that $u = v$. We have shown that for any two points $p, q \in A$ there exists a unique vector, say $v_{pq} \in V$, such that $v_{pq}(p) = q$. We will use this to define a function $[-,-] : A \times A \to V$ by

$$[p, q] := v_{pq}.$$

We want to show that this is a valid subtraction function. To show that the function $[p, -] : A \to V$ is surjective consider any vector $v \in A$ and define $q := v(p)$. By uniqueness this means that $v = v_{pq}$ so we must have $[p, q] = v_{pq} = v$. To show that $[p, -]$ is injective, consider any two points $q, r \in A$ with $v_{pq} = [p,q] = [p,r] = v_{pr}$. Then we have $q = v_{pq}(p) = v_{pr}(p) = r$. Finally, for any points $p, q, r \in A$ we have $(v_{pq} + v_{qr})(p) = v_{qr}(v_{pq}(p)) = v_{qr}(q) = r$, and hence

$$[p,q] + [p,r] = v_{pq} + v_{qr} = v_{pr} = [p,r],$$

as desired.

In summary, let A be a set and let V be a vector space. We have shown that a regular V-action $V \times A \to A$ and a subtraction function $A \times A \to V$ are equivalent structures. The equivalence is given by

$$[p,q] = v \quad \Longleftrightarrow \quad v(p) = q.$$

For part (b), let A be an affine space over V with subtraction function $(p,q) \to [p,q]$ and action $(v,p) \mapsto v(p)$. From now on we will use the more suggestive notation $v(p) =$ "$p + v$". This notation is reasonable since for all points $p \in A$ and vectors $u, v \in V$ we have

$$(p+v)+u = u(v(p)) = (v+u)(p) = p + (v+u).$$

(We will be careful not to take the notation too literally.) For posterity let me record the fact that for all $p, q \in A$ and $v \in V$ we have

$$[p,q] = v \quad \Longleftrightarrow \quad p + v = q. \tag{1}$$

In particular, we have $p + [p,q] = q$. Now we say that $f : A \to A$ is an affine function if there exists a linear function $df : V \to V$ such that for all $p \in A$ and $v \in V$ we have

$$f(p+v) = f(p) + df(v).$$

In particular, taking $v = [p,q]$ gives

$$f(q) = f(p + [p,q]) = f(p) + df([p,q])$$

and it follows from (1) that $df([p,q]) = [f(p), f(q)]$. This means that the linear function df is uniquely determined by the affine function f. We call df the differential of f.

Now consider two affine functions $f, g : A \to A$. For all $p \in A$ and $v \in V$ we have

$$\begin{aligned}(fg)(p+v) &= f(g(p+v)) \\ &= f(g(p) + dg(v)) \\ &= f(g(p)) + df(dg(v)) \\ &= (fg)(p) + (df \cdot dg)(v)\end{aligned} \tag{2}$$

and then by uniqueness of the differential we conclude that $d(fg) = df \cdot dg$. If $fg = 1$ then equation (2) says that $p + v = p + (df \cdot dg)(v)$ for all $v \in V$ and it follows that

$$v = [p, p+v] = [p, p + (df \cdot dg)(v)] = (df \cdot dg)(v).$$

Since this is true for all $v \in V$ we conclude that $df \cdot dg = 1$. Similarly, if $gf = 1$ then we have $dg \cdot df = 1$. In summary, if f is invertible with $f^{-1} = g$ then df is invertible with $(df)^{-1} = dg$.

Conversely, let $f : A \to A$ be any affine function and suppose that the differential df is invertible. To show that f is invertible, we first choose an arbitrary basepoint $o \in A$. Then for all points $p \in A$ we define a function $g : A \to A$ by

$$g(p) := o + (df)^{-1}([f(o), p]).$$

Note that for all $p \in A$ and $v \in V$ we have

$$\begin{aligned} g(p+v) &= o + (df)^{-1}([f(o), p+v]) \\ &= o + (df)^{-1}([f(o), p] + [p, p+v]) \\ &= o + (df)^{-1}([f(o), p]) + (df)^{-1}([p, p+v]) \\ &= g(p) + (df)^{-1}(v), \end{aligned}$$

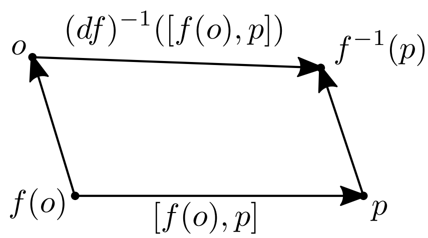
so that g is an affine function with differential $dg = (df)^{-1}$. Then for all $p \in A$ we have

$$\begin{aligned} f(g(p)) &= f(o + (df)^{-1}([f(o), p])) \\ &= f(o) + df((df)^{-1}([f(o), p])) \\ &= f(o) + [f(o), p] \\ &= p \end{aligned}$$

and

$$\begin{aligned} g(f(p)) &= o + (df)^{-1}([f(o), f(p)]) \\ &= o + (df)^{-1}(df([o, p])) \\ &= o + [o, p] \\ &= p, \end{aligned}$$

hence f is invertible with $f^{-1} = g$. [Remark: This proves a very special case of the Jacobian conjecture. (Look it up!) You might wonder where the definition of the function g came from. Let $o \in A$ be any point and suppose that the affine function $f : A \to A$ is invertible. Then applying f^{-1} to the vector $[f(o), p]$ gives the following diagram:



We conclude from the diagram that $f^{-1}(p) = o + (df)^{-1}([f(o), p])$, and this suggests how one might **define** the function f^{-1} in terms of its differential $(df)^{-1}$.]

For part (c), let me first define a map $t : V \to \mathsf{GA}(V)$ sending the vector $v \in V$ to the "translation function" $t_v : A \to A$ defined by $t_v(p) := p + v$. Note that t_v is affine with differential $dt_v = 1$. Indeed, for all $p \in A$ and $u \in V$ we have

$$t_v(p+u) = (p+u) + v = p + (u+v) = p + (v+u) = (p+v) + u = t_v(p) + u.$$

Furthermore, the map t is a group homomorphism since for all points $p \in A$ and vectors $u, v \in V$ we have

$$t_u(t_v(p)) = t_u(p+v) = (p+v) + u = p + (v+u) = t_{v+u}(p).$$

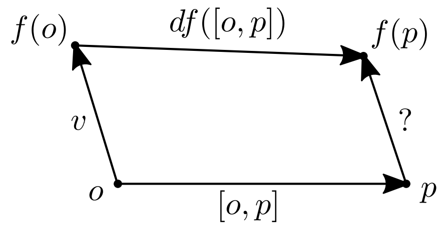
We have already seen in (2) that the differential map $d : \mathsf{GA}(V) \to \mathsf{GL}(V)$ is a homomorphism, so we obtain a sequence of homomorphisms

$$1 \longrightarrow V \xrightarrow{t} \mathsf{GA}(V) \xrightarrow{d} \mathsf{GL}(V) \longrightarrow 1 \tag{3}$$

with $\operatorname{im} t \subseteq \ker d$. I claim that, moreover, $\operatorname{im} t = \ker d$. Indeed, let $f : A \to A$ be any affine map with $df = 1$. If we choose an arbitrary basepoint $o \in A$ and define the vector $v := [o, f(o)]$ then for all $p \in A$ we have

$$\begin{aligned}
[p, f(p)] &= [p, o] + [o, f(o)] + [f(o), f(p)] \\
&= [p, o] + v + df([o, p]) \\
&= [p, o] + v + [o, p] \\
&= v,
\end{aligned}$$

and it follows from (1) that $f(p) = p + v$, that is $f = t_v$. Thus the sequence (3) is exact. [Remark: Here is the intuition behind the proof. Consider any points $o, p \in A$ and any affine function $f : A \to A$. By applying f to the vector $[o, p]$ we obtain the following diagram:



If $df = 1$ then this diagram is a "parallelogram" and it follows that $[p, f(p)] = [o, f(o)] = v$.]

Finally, I will show that "choosing an origin" $o \in A$ defines a homomorphism $s : \mathsf{GL}(V) \to \mathsf{GA}(V)$ such that for all $\varphi \in \mathsf{GL}(V)$ we have $d(s_\varphi) = \varphi$ (i.e., a section of the map d). Recall from the definition of affine space that $[o, -] : A \to V$ is a bijection. If $\varphi \in \mathsf{GL}(V)$ is any invertible linear map then we will define the function $s_\varphi : A \to A$ so that the following square commutes:

$$\begin{tikzcd}
A \arrow[r, "{[o,-]}"] \arrow[d, "s_\varphi"'] & V \arrow[d, "\varphi"] \\
A \arrow[r, "{[o,-]}"] & V
\end{tikzcd}$$

That is, for all points $p \in A$ we let $s_\varphi(p) \in A$ be the unique point such that

$$[o, s_\varphi(p)] = \varphi([o, p]). \tag{4}$$

First note that $s_\varphi : A \to A$ is an affine function with $d(s_\varphi) = \varphi$. Indeed, for all $p \in A$ and $v \in V$ we have

$$\begin{aligned}
[o, s_\varphi(p + v)] &= \varphi([o, p + v]) \\
[o, s_\varphi(p)] + [s_\varphi(p), s_\varphi(p + v)] &= \varphi([o, p] + [p, p + v]) \\
\cancel{\varphi([o, p])} + [s_\varphi(p), s_\varphi(p + v)] &= \cancel{\varphi([o, p])} + \varphi([p, p + v]) \\
[s_\varphi(p), s_\varphi(p + v)] &= \varphi(v),
\end{aligned}$$

and it follows from (1) that $s_\varphi(p + v) = s_\varphi(p) + \varphi(v)$. Then for all $\varphi, \mu \in \mathsf{GL}(V)$ and $p \in A$ we have

$$\begin{aligned}
[o, s_\varphi(s_\mu(p))] &= \varphi([o, s_\mu(p)]) \\
&= \varphi(\mu([o, p])) \\
&= (\varphi\mu)([o, p]) \\
&= [o, s_{\varphi\mu}(p)],
\end{aligned}$$

and the injectivity of $[o,-]$ implies that $s_\varphi(s_\mu(p)) = s_{\varphi\mu}(p)$. Since this is true for all $p \in A$ we conclude that $s_{\varphi\mu} = s_\varphi s_\mu$ and hence $s : \mathsf{GL}(V) \to \mathsf{GA}(V)$ is a homomorphism.

We have shown that the short exact sequence (3) is right-split. It follows from HW3.5 that

$$\mathsf{GA}(V) = t(V) \rtimes s(\mathsf{GL}(V)) \approx V \rtimes \mathsf{GL}(V).$$

And what is the action of $\mathsf{GL}(V)$ on V implied by this semi-direct product? I claim that it's the obvious action: for all linear maps $\varphi \in \mathsf{GL}(V)$ and vectors $v \in V$ we have

$$s_\varphi t_v s_\varphi^{-1} = t_{\varphi(v)}.$$

Indeed, for all $p \in A$, $v \in V$ and $\varphi \in \mathsf{GL}(V)$ we have

$$\begin{aligned}
\cancel{[o,p]} + [p, s_\varphi t_v s_\varphi^{-1}(p)] &= [o, s_\varphi t_v s_\varphi^{-1}(p)] \\
&= \varphi([o, t_v s_\varphi^{-1}(p)]) \\
&= \varphi([o, s_\varphi^{-1}(p)] + [s_\varphi^{-1}(p), t_v s_\varphi^{-1}(p)]) \\
&= \varphi(\varphi^{-1}([o,p]) + v) \\
&= \cancel{[o,p]} + \varphi(v)
\end{aligned}$$

and it follows from (1) that $s_\varphi t_v s_\varphi^{-1}(p) = p + \varphi(v) = t_{\varphi(v)}(p)$. $\square$

[Remark: No splitting of the short exact sequence (3) is better than any other splitting, just as no point of A is better than any other point. That's the whole idea. However, suppose that V is n-dimensional over a field K. If we fix an arbitrary basis for V and an arbitrary basepoint for A then we can represent the element $t_v s_\varphi \in \mathsf{GA}(V)$ as an $(n+1)\times(n+1)$ matrix

$$\left(\begin{array}{c|c} \varphi & v \\ \hline 0 & 1 \end{array}\right),$$

where φ is an $n \times n$ matrix and v is an $n \times 1$ column. One can check that matrix multiplication agrees with the relation $(t_u s_\varphi)(t_v s_\mu) = (t_u t_{\varphi(v)})(s_\varphi s_\mu)$. In other words, $\mathsf{GA}(V)$ is isomorphic to a subgroup of $\mathsf{GL}_{n+1}(K)$. I could have phrased the problem in this language from the beginning but I wanted to emphasize that the map $\mathsf{GA}(V) \hookrightarrow \mathsf{GL}_{n+1}(K)$ is not canonical.]

Problem 4. Grassmannians.

(a) Let $\mathsf{Gr}_1(r,n)$ denote the set of r-element subsets of $\{1,2,\ldots,n\}$. Show that the obvious action of the symmetric group S_n on $\mathsf{Gr}_1(r,n)$ is transitive with stabilizer isomorphic to $S_r \times S_{n-r}$. Then use orbit-stabilizer to compute $|\mathsf{Gr}_1(r,n)|$.

(b) Let K be a field and let $\mathsf{Gr}_K(r,n)$ denote the set of r-dimensional subspaces of K^n. Show that the obvious action of $\mathsf{GL}_n(K)$ on $\mathsf{Gr}_K(r,n)$ is transitive with stabilizer isomorphic to

$$\mathsf{Mat}_{r,n-r}(K) \rtimes (\mathsf{GL}_r(K) \times \mathsf{GL}_{n-r}(K)),$$

where $\mathsf{Mat}_{r,n-r}(K)$ is the additive group of $r \times (n-r)$ matrices. [Hint: Show that the stabilizer is isomorphic to the group of block matrices

$$\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right)$$

with $A \in \mathsf{GL}_r(K)$, $B \in \mathsf{GL}_{n-r}(K)$, and $C \in \mathsf{Mat}_{r,n-r}(K)$.]

(c) When K is the finite field of size q we will write $\mathsf{Gr}_q(r,n) := \mathsf{Gr}_K(r,n)$. Use orbit-stabilizer and part (b) to compute $|\mathsf{Gr}_q(r,n)|$. [Hint: Define $\mathsf{GL}_n(q) := \mathsf{GL}_n(K)$. You can assume the formula

$$|\mathsf{GL}_n(q)| = q^{\binom{n}{2}}(q-1)^n[n]_q!,$$

where $[n]_q! = [n]_q[n-1]_q\cdots[2]_q[1]_q$ and $[m]_q = 1+q+q^2+\cdots+q^{m-1}$.] Now compare your answers from parts (a) and (c).

Proof. For part (a), we define the action of $\pi \in S_n$ on a subset $X \subseteq \{1,2,\ldots,n\}$ by

$$\pi(X) := \{\pi(x_1), \pi(x_2), \ldots, \pi(x_r)\}.$$

Since $|\pi(X)| = |X|$, this defines an action of S_n on the set $\mathsf{Gr}_1(r,n)$ for any r. Now consider any two subsets $X = \{x_1,\ldots,x_r\}$ and $Y = \{y_1,\ldots,y_r\}$ in $\mathsf{Gr}_1(r,n)$. By extending these to the full set we can write

$$\{x_1,\ldots,x_r,x_{r+1},\ldots,x_n\} = \{1,2,\ldots,n\} = \{y_1,\ldots,y_r,y_{r+1},\ldots,y_n\}.$$

Then the permutation $\pi \in S_n$ defined by $\pi(x_i) := y_i$ for all i satisfies $\pi(X) = Y$ and we conclude that the action of S_n on $\mathsf{Gr}_1(r,n)$ is transitive.

Now fix a subset $X \in \mathsf{Gr}_1(r,n)$. Let $H \subseteq S_n$ be the subgroup that fixes the elements of $\{1,2,\ldots,n\} \setminus X$ pointwise and let K be the subgroup that fixes the elements of X pointwise. Note that $H \approx S_r$ and $K \approx S_{n-r}$. Then since H and K commute elementwise and intersect trivially we have $HK = H \times K \approx S_r \times S_{n-r}$. I claim that $HK = \mathsf{Stab}_{S_n}(X)$. Certainly, we have $HK \subseteq \mathsf{Stab}_{S_n}(X)$. Conversely, consider any $\pi \in S_n$ such that $\pi(X) = X$. We will define $h_\pi \in S_n$ by $h_\pi(i) := \pi(i)$ when $i \in X$ and $h_\pi(i) := i$ when $i \in \{1,2,\ldots,n\} \setminus X$. Similarly we define $k_\pi \in S_n$ by $k_\pi(i) := \pi(i)$ when $i \in \{1,2,\ldots,n\} \setminus X$ and $k_\pi(i) := i$ when $i \in X$. Note that we have $h_\pi \in H$, $k_\pi \in K$ and $\pi = h_\pi k_\pi \in HK$. It follows that $\mathsf{Stab}_{S_n}(X) \subseteq HK$, as desired.

Finally, the orbit-stabilizer theorem gives

$$\begin{aligned}|\mathsf{Gr}_1(r,n)| &= \frac{|S_n|}{|S_r \times S_{n-r}|}\\ &= \frac{|S_n|}{|S_r|\cdot|S_{n-r}|}\\ &= \frac{n!}{r!(n-r)!}.\end{aligned}$$

For part (b), we define an action of $\varphi \in \mathsf{GL}_n(K)$ on a subspace $U \subseteq K^n$ by

$$\varphi(U) := \{\varphi(u) : u \in U\}.$$

Note that $\varphi(U) \subseteq K^n$ is another subspace of the same dimension, so we obtain an action of $\mathsf{GL}_n(K)$ on the set $\mathsf{GL}_K(r,n)$ for any r. Now consider any two subspaces U and V in $\mathsf{Gr}_K(r,n)$ and choose bases $u_1,u_2,\ldots,u_r \in U$ and $v_1,v_2,\ldots,v_n \in V$. By extending these to bases for the full space we can write

$$\langle u_1,\ldots,u_r,u_{r+1},\ldots,u_n\rangle = K^n = \langle v_1,v_2,\ldots,v_r,v_{r+1},\ldots,v_n\rangle.$$

Now define a function $\varphi \in \mathsf{GL}_n(K)$ by setting $\varphi(u_i) := v_i$ for all i and extending linearly. Since $\varphi(U) = V$ we conclude that the action of $\mathsf{GL}_n(K)$ on $\mathsf{Gr}_K(r,n)$ is transitive.

Now fix a subspace $U \in \mathsf{Gr}_K(r,n)$ and choose a basis $u_1,u_2,\ldots,u_r \in U$. Extend this to a basis $u_1,\ldots,u_r,u_{r+1},\ldots,u_n$ for K^n and define the complementary subspace $U' :=$

$\langle u_{r+1}, \ldots, u_n \rangle$. We will express each vector $x \in K^n$ in this basis by writing

$$x = \left(\begin{array}{c} x_1 \\ \vdots \\ x_k \\ \hline x_{k+1} \\ \vdots \\ x_n \end{array}\right) = \left(\begin{array}{c} x_U \\ \hline x_{U'} \end{array}\right),$$

so that $x \in U$ if and only if $x_{U'} = 0$. Now consider the set of matrices

$$P := \left\{ \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) : A \in \mathsf{GL}_r(K), B \in \mathsf{GL}_{n-r}(K), C \in \mathsf{Mat}_{r,n-r}(K) \right\},$$

and observe that P is a subgroup of $\mathsf{GL}_n(K)$ with blockwise multiplication

$$\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right)\left(\begin{array}{c|c} A' & C' \\ \hline 0 & B' \end{array}\right) = \left(\begin{array}{c|c} AA' & AC' + CB' \\ \hline 0 & BB' \end{array}\right).$$

I claim that $P = \mathsf{Stab}_{\mathsf{GL}_n(K)}(U)$. Indeed, if $\varphi \in \mathsf{GL}_n(K)$ is any invertible matrix and $x \in K^n$ is any vector then we have

$$\varphi(x) = \left(\begin{array}{c|c} A & C \\ \hline D & B \end{array}\right)\left(\begin{array}{c} x_U \\ \hline x_{U'} \end{array}\right) = \left(\begin{array}{c} Ax_U + Cx_{U'} \\ \hline Dx_U + Bx_{U'} \end{array}\right) \tag{5}$$

for some matrices A, B, C, D of the correct shape. If $\varphi \in P$ (i.e. $D = 0$) and $x \in U$ (i.e. $x_{U'} = 0$) then we find that $\varphi(x)_{U'} = Dx_U + Bx_{U'} = 0$, and hence $\varphi(x) \in U$. It follows that $P \subseteq \mathsf{Stab}_{\mathsf{GL}_n(K)}(U)$. Conversely, suppose that $\varphi \in \mathsf{GL}_n(K)$ and $\varphi(x) \in U$ for all $x \in U$. From (5) this means that $Dx_U = 0$ for all $x_U \in U$ and hence $D = 0$. Then since $0 \neq \det \varphi = \det A \cdot \det B$ we must have $A \in \mathsf{GL}_r(K)$ and $B \in \mathsf{GL}_{n-r}(K)$ so that $\varphi \in P$. We conclude that $\mathsf{Stab}_{\mathsf{GL}_n(K)} \subseteq P$ as desired.

It remains to show that $P \approx \mathsf{Mat}_{r,n-r}(K) \rtimes (\mathsf{GL}_r(K) \times \mathsf{GL}_{n-r}(K))$. To do this we will let H denote the subgroup of P where $C = 0$ and $B = I$, let K denote the subgroup of P where $C = 0$ and $A = I$, and let M denote the subgroup of P where $A = I$ and $B = I$. Note that we have $H \approx \mathsf{GL}_r(K)$ and $K \approx \mathsf{GL}_{n-r}(K)$. Then we have $HK = H \times K \approx \mathsf{GL}_r(K) \times \mathsf{GL}_{n-r}(K)$ because the groups intersect trivially and commute elementwise:

$$\left(\begin{array}{c|c} A & 0 \\ \hline 0 & I \end{array}\right)\left(\begin{array}{c|c} I & 0 \\ \hline 0 & B \end{array}\right) = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right) = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & B \end{array}\right)\left(\begin{array}{c|c} A & 0 \\ \hline 0 & I \end{array}\right).$$

Note that we also have $M \approx \mathsf{Mat}_{k,n-k}(K)$ as **additive** groups because

$$\left(\begin{array}{c|c} I & C \\ \hline 0 & I \end{array}\right)\left(\begin{array}{c|c} I & C' \\ \hline 0 & I \end{array}\right) = \left(\begin{array}{c|c} I & C + C' \\ \hline 0 & I \end{array}\right).$$

Next observe that $M \cap (H \times K) = 1$ and that $M(H \times K) = P$ because

$$\left(\begin{array}{c|c} I & CB^{-1} \\ \hline 0 & I \end{array}\right)\left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right) = \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right).$$

To finish the proof we will show that M is normal in P. Indeed, we have

$$\begin{aligned}
&\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right)\left(\begin{array}{c|c} I & D \\ \hline 0 & I \end{array}\right)\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right)^{-1} \\
&= \left(\begin{array}{c|c} A & AD+C \\ \hline 0 & B \end{array}\right)\left(\begin{array}{c|c} A^{-1} & -A^{-1}CB^{-1} \\ \hline 0 & B^{-1} \end{array}\right) \\
&= \left(\begin{array}{c|c} AA^{-1} & -AA^{-1}CB^{-1}+(AD+C)B^{-1} \\ \hline 0 & BB^{-1} \end{array}\right) \\
&= \left(\begin{array}{c|c} I & -CB^{-1}+ADB^{-1}+CB^{-1} \\ \hline 0 & I \end{array}\right) \\
&= \left(\begin{array}{c|c} I & ADB^{-1} \\ \hline 0 & I \end{array}\right).
\end{aligned}$$

We conclude that $P = M \rtimes (H \times K)$ as desired, and the associated action of $H \times K$ on M is our favorite action of $\mathsf{GL}_r(K) \times \mathsf{GL}_{n-r}$ on $\mathsf{Mat}_{r,n-r}(K)$, namely

$$((A,B),D) \mapsto ADB^{-1}.$$

For part (c), note that the orbit-stabilizer theorem identifies the Grassmannian $\mathsf{Gr}_K(r,n)$ with the coset space $\mathsf{GL}_n(K)/P$, where P is the subgroup defined above. Now assume that K is the finite field of size q. We proved in class that $|\mathsf{GL}_n(q)| = q^{\binom{n}{2}}(q-1)^n[n]_q!$ and we see that $|\mathsf{Mat}_{r,n-r}(q)| = q^{r(n-r)}$ because the matrix entries are arbitrary. Note that we also have $|P| = |M \rtimes (H \times K)| = |M| \cdot |H \times K| = |M| \cdot |H| \cdot |K|$. Putting everything together gives

$$\begin{aligned}
|\mathsf{Gr}_q(r,n)| &= \frac{|\mathsf{GL}_n(q)|}{|\mathsf{Mat}_{r,n-r}(q)| \cdot |\mathsf{GL}_r(q)| \cdot |\mathsf{GL}_{n-r}(q)|} \\
&= \frac{q^{n(n-1)/2}(q-1)^n[n]_q!}{q^{r(n-r)} \cdot q^{r(r-1)/2}(q-1)^r[r]_q! \cdot q^{(n-r)(n-r-1)/2}(q-1)^{n-r}[n-r]_q!} \\
&= \frac{[n]_q!}{[r]_q!\,[n-r]_q!}.
\end{aligned}$$

If q is a prime power then one can show (for example, by induction) that this last formula is a polynomial in q with non-negative integer coefficients. More generally, if q is any element of a commutative ring R then we can use the polynomial to **define** the element $|\mathsf{Gr}_q(r,n)| \in R$. This explains the strange choice of notation in part (a). □

[Remark: The analogy between parts (a) and (c) is beautiful but I don't really understand it. Here's another beautiful thing that I don't understand. We can think of the Grassmannian $\mathsf{Gr}_{\mathbb{C}}(r,n)$ as a complex manifold. It turns out that the Betti numbers of this manifold are encoded by the numbers $|\mathsf{Gr}_q(r,n)|$. That is, we have

$$P_{\mathsf{Gr}_{\mathbb{C}}(r,n)}(t) = \sum_{k\geq 0} \dim H^k(\mathsf{Gr}_{\mathbb{C}}(r,n))t^k = \frac{[n]_{t^2}!}{[r]_{t^2}![n-r]_{t^2}!},$$

where the right hand side is interpreted as a polynomial in the formal variable t. It follows from this that the Euler characteristic of $\mathsf{Gr}_{\mathbb{C}}(r,n)$ is the binomial coefficient $\binom{n}{r}$. This was one of the motivating examples that led to the Weil Conjectures. (Look it up!)]