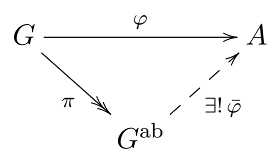
Math 661 **Fall 2013**
Homework 5 **Drew Armstrong**

Problem 0 (Abelianization). Let G be a group and for all $g,h\in G$ define the commutator $[g,h]:=ghg^{-1}h^{-1}\in G$. The subgroup of G generated by commutators is called the commutator subgroup:

$$[G,G]:=\langle [g,h]:g,h\in G\rangle.$$

(a) Prove that $[G,G]\lhd G$.
(b) Prove that the quotient $G^{\mathrm{ab}}:=G/[G,G]$ (called the abelianization of G) is abelian.
(c) If $N\lhd G$ is any normal subgroup such that G/N is abelian, prove that $[G,G]\leq N$.
(d) Put everything together to prove the **universal property of abelianization**: Given a homomorphism $\varphi:G\to A$ to an abelian group A, there exists a unique homomorphism $\bar{\varphi}:=G^{\mathrm{ab}}\to A$ such that $\varphi=\bar{\varphi}\circ\pi$, where $\pi:G\to G^{\mathrm{ab}}$ is the canonical surjection.



Proof. To show (a), first note that elements of $[G,G]$ are products of commutators and inverses of commutators. But the inverse of a commutator is also a commutator:

$$[g,h]^{-1}=(ghg^{-1}h^{-1})^{-1}=hgh^{-1}g^{-1}=[h,g].$$

Thus every element of $[G,G]$ is a product of commutators. Next note that the conjugate of a commutator is a commutator. Indeed, for all $s,g,h\in G$ we have

$$s[g,h]s^{-1}=s(ghg^{-1}h^{-1})s^{-1}=(sgs^{-1})(shs^{-1})(sg^{-1}s^{-1})(shs^{-1})=[sgs^{-1},shs^{-1}].$$

Finally, given any $x=[g_1,h_1]\cdots[g_k,h_k]\in[G,G]$ and $s\in G$ we have

$$\begin{aligned}sxs^{-1}&=s\left([g_1,h_1]\cdots[g_k,h_k]\right)s^{-1}\\&=(s[g_1,h_1]s^{-1})\cdots(s[g_k,h_k]s^{-1})\\&=[sg_1s^{-1},sh_1s^{-1}]\cdots[sg_ks^{-1},sh_ks^{-1}]\in[G,G],\end{aligned}$$

and we conclude that $[G,G]\lhd G$.

For part (b) we will write $G':=[G,G]$ to save space. To show that G/G' is abelian consider any cosets gG' and hG' with $g,h\in G$. We will be done if we can show that $[gG',hG']$ is the identity coset G'. And this is true because

$$\begin{aligned}[gG',hG']&=(gG')(hG')(gG')^{-1}(hG')^{-1}\\&=(gG')(hG')(g^{-1}G')(h^{-1}G')\\&=(ghg^{-1}h^{-1})G'\\&=[g,h]G'\\&=G'.\end{aligned}$$

For part (c), assume that $N\lhd G$ with G/N abelian and consider any $g,h\in G$. Then since $(gN)(hN)=(hN)(gN)$ we have

$$N=(gN)(hN)(gN)^{-1}(hN)^{-1}=(ghg^{-1}h^{-1})N=[g,h]N,$$

which implies that $[g,h] \in N$. Since N contains all commutators $[g,n]$ for $g,h \in G$ we conclude that $[G,G] \leq N$.

For part (d) assume we have $\varphi : G \to A$ where A is abelian and let $N = \operatorname{Ker}\varphi$. By part (c) we know that $\operatorname{Ker}\pi = [G,G] \leq N = \operatorname{Ker}\varphi$. This allows us to define a map $\bar{\varphi} : G^{\text{ab}} \to A$ by setting $\bar{\varphi}(g[G,G]) := \varphi(g)$. To see that this is well-defined, suppose that $g[G,G] = h[G,G]$, so that $gh^{-1} \in [G,G] \leq N$. Then we have $\varphi(g)\varphi(h)^{-1} = \varphi(gh^{-1}) = 1_A$, which implies that $\varphi(g) = \varphi(h)$. Also note that $\bar{\varphi} : G^{\text{ab}} \to A$ is a homomorphism because

$$\bar{\varphi}(gh[G,G]) = \varphi(gh) = \varphi(g)\varphi(h) = \bar{\varphi}(g[G,G])\bar{\varphi}(h[G,G]),$$

and note that $\varphi = \bar{\varphi} \circ \pi$ since for all $g \in G$ we have

$$\bar{\varphi} \circ \pi(g) = \bar{\varphi}(g[G,G]) = \varphi(g).$$

Finally, suppose that $F : G^{\text{ab}} \to A$ is another morphism satisfying $\varphi = F \circ \pi$. Then for all $g \in G$ we have

$$\bar{\varphi}(g[G,G]) = \varphi(g) = F(\pi(g)) = F(g[G,G])$$

so that $F = \bar{\varphi}$ as desired. ∎

Problem 1 (Splitting Lemma). Let R be a commutative ring with 1 and consider a short exact sequence of R-modules:

$$0 \longrightarrow A \xrightarrow{q} B \xrightarrow{r} C \longrightarrow 0.$$

Prove that if there exists $t : B \to A$ such that $t \circ q$ is the identity on A, then $B \approx A \oplus C$. [Hint: Define a map $\varphi : B \to A \oplus C$ by $\varphi(b) := (t(b), r(b))$. To show that φ is injective, assume that $\varphi(b) = \varphi(b')$. Show that this implies $b - b' \in \operatorname{Ker} r = \operatorname{Im} q$, and hence $b - b' = q \circ t(b - b') = q(t(b) - t(b')) = q(0) = 0$. To show that φ is surjective consider $(a,c) \in A \oplus C$. Since r and t are surjective there exist $b, b' \in B$ such that $a = t(b)$ and $c = r(b')$. Now let $x = b' + q \circ t(b - b')$ and show that $\varphi(x) = (a,c)$.]

Proof. Note that the map $\varphi(b) := (t(b), r(b))$ is an R-homomorphism because $t : B \to A$ and $r : B \to C$ are both R-homomorphisms. We must show that φ is **bijective**.

To show that φ is **injective**, consider any $b, b' \in B$ such that $\varphi(b) = \varphi(b')$. Equivalently, we have $t(b) = t(b')$ and $r(b) = r(b')$. Since $0 = r(b) - r(b') = r(b - b')$ we see that $b - b' \in \operatorname{Ker} r$. By exactness, this implies $b - b' \in \operatorname{Im} q$, hence there exists $a \in A$ with $b - b' = q(a)$. Since $t \circ q$ is the identity on A this implies

$$q \circ t \circ q(a) = q(t \circ q(a)) = q(a).$$

In other words, we have

$$b - b' = q \circ t(b - b') = q(t(b) - t(b')) = q(0) = 0,$$

where we used the assumption that $t(b) = t(b')$. We conclude that $b = b'$ as desired.

To show that φ is **surjective**, first note that t is surjective because for all $a \in A$ we have $t(q(a)) = a$. Now fix an arbitrary element $(a,c) \in A \oplus C$. Since t it surjective and r is surjective (by exactness), there exist $b, b' \in B$ such that $a = t(b)$ and $c = q(b')$. Now consider the element $x = b' + q \circ t(b - b') \in B$. Note that

$$t(x) = t(b') + t \circ q \circ t(b - b') = t(b') + t(b - b') = t(b') + t(b) - t(b') = t(b) = a$$

because $t \circ q$ is the identity. Since $\operatorname{Im} q = \operatorname{Ker} r$ we also have

$$r(x) = r(b') + r(q(t(b - b'))) = r(b') + 0 = r(b') = c,$$

and we conclude that $\varphi(x) = (t(x), r(x)) = (a,c)$ as desired. ∎

Problem 2. We say that a matrix $A \in GL(n,\mathbb{C})$ is unitary if $A^*A = I$, where A^* is the conjugate transpose. Let $U(n) \leq GL(n,\mathbb{C})$ denote the unitary group of unitary matrices.

(a) Prove that $U(n)$ is actually a group.

(b) Let $(x,y) = x^*y = \sum_i \overline{x_i} y_i$ be the standard Hermitian form on $\mathbb{C}^n$. Prove that $A \in GL(n,\mathbb{C})$ is unitary if and only if $(Ax,Ay) = (x,y)$ for all $x,y \in \mathbb{C}^n$.

(c) Prove that $A \in GL(n,K)$ is unitary if and only if its columns are orthonormal.

(d) Prove that every $A \in U(n)$ is conjugate in $U(n)$ to a diagonal matrix. [Hint: Let $A \in U(n)$. Since $\mathbb{C}$ is algebraically closed, A has an eigenvector, say $A\mathbf{v}_1 = \lambda \mathbf{v}_1$. Assume it is possible to extend this to an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for $\mathbb{C}^n$ (which it is, via the Gram-Schmidt algorithm). Letting $P = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$ gives us

$$P^{-1}AP = \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right),$$

with $A' \in U(n-1)$. By induction, A' is conjugate in $U(n-1)$ to a diagonal matrix.]

Proof. First we establish a few properties of conjugate transpose. For all column vectors $x \in \mathbb{C}^n$ let x^* denote the complex conjugate transpose row vector, and define the standard Hermitian form by $(x,y) := x^*y$. Note that for all $A \in GL(n,\mathbb{C})$ the conjugate transpose matrix A^* is characterized by

$$(Ax,y) = (x,A^*y) \quad \text{for all } x,y \in \mathbb{C}^n.$$

Indeed, if $e_i \in \mathbb{C}^n$ is a standard basis vector then for all $x \in \mathbb{C}^n$ we have

$$e_i^*A^*x = (e_i, A^*x) = (Ae_i, x) = (Ae_i)^*x,$$

and it follows that $e_i^*A^* = (Ae_i)^*$. But $e_i^*A^*$ is the i-th row of A^* and $(Ae_i)^*$ is the conjugate transpose of the i-th column of A. Now for all $A, B \in \mathbb{C}^*$ and $x,y \in \mathbb{C}^n$ we have

$$(ABx,y) = (Bx, A^*y) = (x, B^*A^*y),$$

which by the previous remarks implies that $(AB)^* = B^*A^*$. Finally, note that for all $A \in GL(n,\mathbb{C})$ we have $(A^*)^{-1} = (A^{-1})^*$ because $(A^{-1})^*A^* = (AA^{-1})^* = I^* = I$.

To show part (a), consider $A, B \in GL(n,\mathbb{C})$ such that $A^*A = I$ and $B^*B = I$. By Rank-Nullity we also have $BB^* = I$ and hence

$$\begin{aligned} (AB^{-1})^*(AB^{-1}) &= (B^{-1})^*A^*AB^{-1} \\ &= (B^{-1})^*B^{-1} \\ &= (B^*)^{-1}B^{-1} \\ &= (BB^*)^{-1} \\ &= I. \end{aligned}$$

We conclude that AB^{-1} is unitary, hence $U(n)$ is a group.

To show part (b) first suppose that $A^*A = I$. Then for all $x,y \in \mathbb{C}^n$ we have

$$(Ax,Ay) = (x, A^*Ay) = (x,y).$$

Conversely, suppose that $(Ax,Ay) = (x,y)$ for all $x,y \in \mathbb{C}^n$. Setting $y = e_i$ gives

$$x^*e_i = (x,e_i) = (Ax,Ae_i) = (x, A^*Ae_i) = x^*A^*Ae_i.$$

Since this holds for all $x \in \mathbb{C}^n$ we conclude that e_i is equal to A^*Ae_i (which is the i-th column of A^*A), hence $A^*A = I$.

To show (c), let a_i denote the i-th column of A, so that a_i^* is the i-th row of A^*. By definition the i, j entry of A^*A is $(a_i, a_j) = a_i^* a_j$ and since $A^*A = I$ we have

$$(a_i, a_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

In other words, the columns of A are orthonormal.

To show (d) consider $A \in U(n)$. We first show that A has an eigenvalue. Given any vector $x \in \mathbb{C}^n$, the $n + 1$ vectors

$$x, Ax, A^2x, \ldots, A^n x$$

cannot be linearly independent because $\mathbb{C}^n$ has dimension n. Thus there exist numbers $a_0, a_1, \ldots, a_n \in \mathbb{C}$ not all zero such that

$$0 = a_0 x + a_1 Ax + \cdots + a_n A^n x.$$

Since A commutes with its powers we can think of $a_0 + a_1 A + \cdots + a_n A^n$ as a polynomial with complex coefficients. Then since $\mathbb{C}$ is algebraically closed, there exist $c, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$\begin{aligned} 0 &= a_0 x + a_1 Ax + \cdots + a_n A^n x \\ &= (a_0 + a_1 A + \cdots + a_n A^n)x \\ &= c(A - \lambda_1 I) \cdots (A - \lambda_n I)x, \end{aligned}$$

which means that $A - \lambda_i I$ is not injective for at least one i. In other words, A has an eigenvalue. We assume that $A\mathbf{v}_1 = \lambda \mathbf{v}_1$ for some $0 \neq \mathbf{v}_1 \in \mathbb{C}^n$.

Now use the Gram-Schmidt process to extend $\mathbf{v}_1$ to an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for $\mathbb{C}^n$ and let $P = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix}$. After changing basis we obtain

$$P^{-1}AP = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)$$

with $A' \in U(n-1)$. Since the columns of P are orthonormal we have $P^*P = I$ and hence $P^{-1}AP$ is unitary. Then since the columns of $P^{-1}AP$ are orthonormal we conclude that

$$P^{-1}AP = \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right).$$

By induction, there exists $Q' \in U(n-1)$ such that $(Q')^{-1}A'Q' = D$ is diagonal. Finally, after defining the unitary matrix

$$Q = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{array}\right),$$

we see that PQ is unitary and

$$(PQ)^{-1}A(PQ) = Q^{-1}(P^{-1}AP)Q = \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & D & \\ 0 & & & \end{array}\right).$$

□

Problem 3. Prove that the center of $GL(n,K)$ is the group of scalar matrices

$$Z(GL(n,K)) = \left\{\alpha I : \alpha \in K^\times\right\} \approx K^\times.$$

Prove that the center of $SL(n,K)$ is the group of n-th roots of unity

$$Z(SL(n,K)) = \{\alpha I : \alpha \in K, \alpha^n = 1\}.$$

Assuming that $\mathbb{F}_q^\times$ is a cyclic group (this is called the Primitive Root Theorem; please don't prove it), compute the order of $PSL(n,q)$.

Proof. Let $e_{ij}(k)$ be the $n \times n$ matrix with $k \in K$ in the i,j position and zeroes elsewhere, and let $E_{ij}(k) = I + e_{ij}(k)$. Note that for $i \neq j$ we have $E_{ij}(k)^{-1} = E_{ij}(-k)$ and hence $E_{ij}(k)$ is invertible. Now suppose that $A = (a_{ij})$ is in the center of $GL(n,K)$. Since A commutes with $E_{ij}(k)$ it must also commute with $e_{ij}(k)$. But note that $Ae_{ij}(1)$ has j-th column equal the i-th column of A and zeroes elsewhere, while $e_{ij}(1)A$ has i-th row equal to the j-th row of A and zeroes elsewhere. Then the equation $Ae_{ij}(1) = e_{ij}(1)A$ says

$$\begin{array}{c} \\ i \\ \\ \end{array}
\begin{pmatrix} & & & a_{1i} & & \\ & & & \vdots & & \\ 0 & \cdots & \cdots & a_{ii} & \cdots & 0 \\ & & & \vdots & & \\ & & & \vdots & & \\ & & & a_{ni} & & \end{pmatrix}^{\;j}
= i \begin{pmatrix} & & & 0 & & \\ & & & \vdots & & \\ a_{j1} & \cdots & \cdots & a_{jj} & \cdots & a_{jn} \\ & & & \vdots & & \\ & & & \vdots & & \\ & & & 0 & & \end{pmatrix}^{\;j},$$

which implies that $a_{ij} = 0$ and $a_{ii} = a_{jj}$ for all $i \neq j$. In other words A is a scalar matrix. The invertible scalar matrices are precisely αI for $\alpha \in K^\times$. [Note that the same proof works more generally when K is a ring with 1.]

Now we wish to show that the center of $SL(n,K)$ consists of scalar matrices. Indeed, note that the matrix $E_{ij}(k)$ with $i \neq j$ has determinant 1 and hence $E_{ij}(k) \in SL(n,K)$. Then the same argument shows that every $A \in Z(SL(n,K))$ has the form αI for some $\alpha \in K$. Since the determinant of αI is α^n we must also have $\alpha^n = 1$.

Finally, let $K = \mathbb{F}_q$. By the primitive root theorem we know that $\mathbb{F}_q^\times$ is cyclic of order $q-1$, say $\mathbb{F}_q^\times = \langle g \rangle$. Then the center of $SL(n,q)$ has the form

$$Z(SL(n,q)) = \{g^x I : (g^x)^n = 1\}.$$

But note that

$$(g^x)^n = 1 \iff g^{xn} = 1 \iff xn \equiv 1 \pmod{q-1}.$$

Thus we want to solve the linear congruence $xn \equiv 1 \pmod{q-1}$. We will first solve the linear diophantine equation

$$xn + y(q-1) = 0$$

which translates to

$$-\frac{x}{y} = \frac{q-1}{n}.$$

If we let $d = \gcd(n, q-1)$ then the most general way to write this fraction is

$$-\frac{x}{y} = \frac{k(q-1)/d}{kn/d} \qquad \text{for all } k \in \mathbb{Z}$$

and it follows that the general solution is

$$(x, y) = \left(k\,\frac{q-1}{d}, -k\,\frac{n}{d}\right) \qquad \text{for all } k \in \mathbb{Z}.$$

After reducing everything mod $q-1$, we find that the general solution to $xn \equiv 0 \pmod{q-1}$ is given by

$$x \equiv k\,\frac{q-1}{d} \pmod{q-1} \qquad \text{for all } k \in \mathbb{Z}$$

and there are d distinct solutions: $0, \frac{q-1}{d}, 2\,\frac{q-1}{d}, \ldots, (d-1)\,\frac{q-1}{d}$. We conclude that

$$|Z(SL(n,q))| = d = \gcd(n, q-1).$$

Finally, we have

$$|PSL(n,q)| = \frac{|SL(n,q)|}{|Z(SL(n,q)|} = \frac{q^{\binom{n}{2}}(q^2-1)(q^3-1)\cdots(q^n-1)}{\gcd(n,q-1)}.$$

□

[Recall that $PSL(n,q)$ are the finite simple groups of "type A_{n-1}". For comparison, there exists a sequence of finite simple groups $E_6(q)$ of "type E_6" with order

$$|E_6(q)| = \frac{q^{36}(q^2-1)(q^5-1)(q^6-1)(q^8-1)(q^9-1)(q^{12}-1)}{\gcd(2,q-1)}.$$

Wow, that looks similar.]

Problem 4. Let $B \leq GL(n,K)$ be the Borel subgroup of upper triangular matrices, let $U \leq B$ be the subgroup of upper unitriangular matrices (i.e. with 1's on the diagonal) and let $T \leq B$ be the subgroup of diagonal matrices (called a maximal torus).

(a) Why is T called a torus?
(b) Prove that $B = T \ltimes U$.
(c) More generally, given $J = (n_1, \ldots, n_k) \in \mathbb{N}^k$ where $n_1 + n_2 + \cdots + n_k = n$ we define the parabolic subgroup

$$P_J = \begin{pmatrix} \boxed{*} & & & * \\ & \boxed{*} & & \\ & & \boxed{*} & \\ 0 & & & \boxed{*} \end{pmatrix} \leq GL(n,K)$$

where the diagonal blocks are square of sizes $n_1, n_2, \ldots, n_k$. We also define the unipotent radical and the Levi complement:

$$U_J = \begin{pmatrix} \boxed{I} & & & * \\ & \boxed{I} & & \\ & & \boxed{I} & \\ 0 & & & \boxed{I} \end{pmatrix} \leq P_J \quad \text{and} \quad L_J = \begin{pmatrix} \boxed{*} & & & 0 \\ & \boxed{*} & & \\ & & \boxed{*} & \\ 0 & & & \boxed{*} \end{pmatrix} \leq P_J.$$

Prove that $P_J = L_J \ltimes U_J$. [Hint: Consider the projection homomorphism $\varphi : P_J \to L_J$ Show that the kernel is U_J. Now consider any $g \in P_J$ and show that $g\varphi(g)^{-1} \in \operatorname{Ker}\varphi = U_J$. It follows that $g \in U_J \cdot \varphi(g) \subseteq U_J L_J$.]

Proof. For part (a) note that T is isomorphic to the direct product of multiplicative groups $K^\times \times K^\times \times \cdots \times K^\times$. In the case $K = \mathbb{C}$ note that $\mathbb{C}^\times$ is homotopy equivalent to a circle. In this case T is homotopy equivalent to a product of n circles, i.e., a torus. The intersection of $T \leq GL(n, \mathbb{C})$ with the subgroup of unitary matrices $U(n)$ is isomorphic to $U(1) \times U(1) \times \cdots \times U(1)$, and this **really is** a torus. The general use of the word "torus" refers to this special case.

Now we will prove (c), of which (b) is a special case. To prove $P_J = L_J \ltimes U_J$ we must show that (1) $L_J \cap U_J = 1$, (2) $U_J \triangleleft P_J$, and (3) $P_J = L_J U_J$. (1) is trivial. Now consider the function $\varphi : P_J \to L_J$ that sends all elements outisde the diagonal blocks to zero. Since matrix multiplication respects block partitions it is easy to see that this is a group homomorphism. (Also note that L_J is isomorphic to $GL(n_1, K) \times \cdots \times GL(n_k, K)$.) The kernel of φ is clearly U_J, which implies (2). Finally, consider any element

$$A = \begin{pmatrix} \boxed{A_1} & & & * \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_k} \end{pmatrix} \in P_J.$$

Note that

$$\varphi(A)^{-1} = \begin{pmatrix} \boxed{A_1^{-1}} & & & 0 \\ & \boxed{A_2^{-1}} & & \\ & & \ddots & \\ 0 & & & \boxed{A_k^{-1}} \end{pmatrix}$$

and hence

$$\varphi(A)^{-1}A = \begin{pmatrix} \boxed{A_1^{-1}A_1} & & & * \\ & \boxed{A_2^{-1}A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_k^{-1}A_k} \end{pmatrix} = \begin{pmatrix} \boxed{I} & & & * \\ & \boxed{I} & & \\ & & \boxed{I} & \\ 0 & & & \boxed{I} \end{pmatrix} \in U_J.$$

We conclude that $A \in \varphi(A)U_J \subseteq L_J U_J$, i.e., (3). □

[There is a relationship between Problems 2 and 4. As mentioned, the diagonal matrices inside $U(n)$ form an actual torus $T \approx U(1) \times \cdots \times U(1)$. You proved in Problem 2 that the conjugates of $T \leq U(n)$ cover the group. This is a general phenomenon that holds in all compact Lie groups G. The major technique of Lie theory is to express everything about G in terms of an arbitrary maximal torus $T \leq G$.]