Math 504 HW6

Rohan Mukherjee

December 1, 2023

1. (a) First notice that for any i, j, and any matrix A, $E_{ij} \cdot A$ is the matrix with A's jth row at row i, and zeros everywhere else. First, clearly $I_n + aE_{ij}$ can generate $U_n^{(n)}$, since $U_n^{(n)} = \{ I_n \}$. Suppose that every element of $U_n^{(k)}$ can be generated by the $I_n + aE_{ij}$ for some $1 \leq k \leq n-1$. Notice that we can move row i of a to row $i-k$ by multiplying by the matrix $E_{i-k,i}$. So now let A be a matrix with the first $k-1$ super diagonals 0. Notice that $(i, i+k-1)$ is the index of the element in the ith row on the $k-1$th super diagonal. The matrix we are going to want to multiply by is the following: replace the $j-k$th row of A with the $j-k$th row minus the last $j-k$ indices of the jth row times $A_{j-1,j+k-2}$. When we move the $j-k$th row up to row j via multiplying by $A_{j-1,j+k-2}E_{j-k,j}$, we will get the correct $j-k$th row of A. For example, if you wanted the following matrix:

$$\begin{pmatrix} 1 & 5 & 2 & 3 \\ 0 & 1 & 6 & 4 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

You would want to multiply the matrix,

$$\begin{pmatrix} 1 & 0 & 2-5\cdot 6 & 3-5\cdot 4 \\ 0 & 1 & 0 & 4-6\cdot 7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

On the left by $(I_4+5E_{1,2})(I_4+6E_{2,3})(I_4+7E_{3,4})$ (The reader may verify this numerically with the Mathematica file I have uploaded here).

(b) Time for the significantly easier problem. Notice that $(I_n + E_{ii}) \cdot (I_n + E_{jj}) = I_n + I_n E_{ii} + I_n E_{jj} + E_{ii}E_{jj}$. $E_{ii} \cdot E_{jj}$ is the matrix with E_{jj}'s ith row at row i, which is just the zero matrix for $i \neq j$. So for a diagonal matrix D with $d_1, \ldots, d_n$ on the main diagonal, none of which are zero, we have concluded that

$$D = \prod_{i=1}^{n} (I_n + (d_i - 1)E_{ii})$$

(c) Notice that for any matrix A, $(I_n + aE_{ii}) \cdot A$ is the matrix with A having row i equal to $a + 1$ times itself by our numerous observations from part (a). Thus, the inverse of $I_n + aE_{ii}$ is just $I_n + ((a+1)^{-1} - 1)E_{ii}$. Equivalently, notice that $E_{ij}E_{kl}$ can only have a nonzero entry at il, otherwise you are taking either a 0 row or a 0 column. Now, the il entry is just $\mathbb{I}_{j=k}$ (the indicator of $j = k$). So, $E_{ij} \cdot E_{kl} = \mathbb{I}_{j=k}E_{il}$, and in particular, $(I + aE_{ii})E_{j\ell} = E_{j\ell} + a\mathbb{I}_{i=j}E_{i\ell}$. $(I + aE_{ii}) \cdot (I + ((a+1)^{-1} - 1)E_{ii}) = I + aE_{ii} + (a+1)^{-1}E_{ii} - E_{ii} + a(a+1)^{-1}E_{ii} - aE_{ii} = I$. Notice next that $A(B + C)A^{-1} = (AB + AC)A^{-1} = ABA^{-1} + ACA^{-1}$. We can now say that, letting $d = (a+1)^{-1} - 1$,

$$(I + aE_{ii})(I + bE_{j\ell})(I + aE_{ii})^{-1} = I + b(I + aE_{ii})E_{j\ell}(I_n + dE_{ii})$$

From the results above, we have that

$$(I + aE_{ii})E_{j\ell} = E_{j\ell} + a\mathbb{I}_{i=j}E_{i\ell}$$

Next notice that

$$E_{j\ell}(I_n + dE_{ii}) = E_{j\ell} + d\mathbb{I}_{\ell=i}E_{ji}$$

So,

$$\begin{aligned}
&(E_{j\ell} + a\mathbb{I}_{i=j}E_{i\ell})(I_n + dE_{ii}) = E_{j\ell} + d\mathbb{I}_{\ell=i}E_{ji} + a\mathbb{I}_{i=j}(E_{i\ell} + d\mathbb{I}_{\ell=i}E_{ii}) \\
&= E_{j\ell} + d\mathbb{I}_{\ell=i}E_{ji} + a\mathbb{I}_{i=j}E_{i\ell} + ad\mathbb{I}_{i=j=\ell}E_{ii}
\end{aligned}$$

Our final result is just $I + bE_{j\ell} + db\mathbb{I}_{\ell=i}E_{ji} + ab\mathbb{I}_{i=j}E_{i\ell} + abd\mathbb{I}_{i=j=\ell}E_{ii}$.

(d) First notice that $B_n = T_nU_n$. Given a matrix $B \in B_n$, B is the product of the matrix with just B's main diagonal, and U, where U's ith row is B's ith row divided by B_{ii}. Also, if $D \in T_n$ is nonidentity, then D has a diagonal entry that is not 1, which is not strictly upper diagonal. So, $T_n \cap U_n = \{ 1 \}$, proving the above claim. Now we claim that the if $U \in U_n$ then $(I+aE_{ii})U(I+aE_{ii})^{-1} \in U_n$. Write $U = \prod(I+a_iE_{\alpha_i\beta_i})$. Then $(I+aE_{ii})U(I+aE_{ii})^{-1} = \prod(I+aE_{ii})(I+a_iE_{\alpha_i\beta_i})(I+aE_{ii})^{-1}$ (The right element will cancel with its inverse each time, leaving only the first and last). By our above calculation, since for strictly upper triangular matrices their generators will never have $j = \ell$, $(I+aE_{ii})(I+a_iE_{\alpha_i\beta_i})(I+aE_{ii})^{-1} \in U_n$. Every term in this product is in U_n thus it is in U_n. Now let D be an arbitrary diagonal matrix, and write $D = \prod(I+a_iE_{ii})$. Then,

$$DUD^{-1} = (I+a_1E_{11})\cdots(I+a_nE_{nn})U(I+a_nE_{nn})^{-1}\cdots(I+a_1E_{11})$$

$(I+a_nE_{nn})U(I+a_nE_{nn})^{-1} \in U_n$ by our calculation above. What's left is $(I+a_1E_{11})\cdots(I+a_{n-1}E_{n-1,n-1})V(I+a_{n-1}E_{n-1,n-1})^{-1}$ for some other strictly upper triangular matrix V. So this is again in U_n, and continuing on in this fashion will show that the entire product is in U_n. Now,

$$[B_n, U_n] = \{\ [x,y] \mid x \in B_n, y \in U_n\ \} = \{\ [UD, V] \mid U, V \in U_n, D \in T_n\ \}$$

By the above claims. From exercise 3 in section 5.4, we have that $[ab, c] = (b^{-1}[a,c]b)[b,c]$, so

$$[UD, V] = D^{-1}[U,V]D \cdot [D, V]$$

Since $[U,V] \in U_n$, $D^{-1}[U,V]D \in U_n$ by the above. Similarly, $[D,V] = D^{-1}V^{-1}DV$, which is in U_n since $D^{-1}V^{-1}D$ is. We have proven that $[B_n, U_n] \leq U_n$. We shall now show we have $I + bE_{j\ell} \in [B_n, U_n]$ for any $j \neq \ell$, which would complete the proof. We claim that $(I+E_{jj})(I+bE_{j\ell})(I+E_{jj})^{-1}(I+bE_{j\ell})^{-1} = I + bE_{j\ell}$. Notice first that,

$$(I+bE_{j\ell})(I-bE_{j\ell}) = I + bE_{j\ell} - bE_{j\ell} - b^2\mathbb{I}_{l=j}E_{jl} = I$$

Since $j < l$ (importantly, they are not equal). Next, notice that,

$$(I+E_{jj})(I+bE_{j\ell})(I+E_{jj})^{-1} = I + bE_{j\ell} + db\mathbb{I}_{\ell=j}E_{jj} + b\mathbb{I}_{j=j}E_{j\ell} + bd\mathbb{I}_{j=j=\ell}E_{jj} = I + 2bE_{j\ell}$$

Finally,

$$(I+bE_{j\ell}+bE_{j\ell})(I-bE_{j\ell}) = I + bE_{j\ell}(I-bE_{j\ell}) = I + bE_{j\ell} - b^2E_{j\ell}^2$$

At last, $E_{j\ell}^2 = \mathbb{I}_{\ell=j}E_{j\ell} = 0$, since $j \neq \ell$. So $[B_n, U_n]$ contains all the generators of U_n, and hence $U_n \leq [B_n, U_n]$, which completes the proof that $[B_n, U_n] = U_n$. We now claim that $[B_n, B_n] \leq U_n$. Indeed, let $X, Y \in B_n$, and write $X = DU$, and $Y = TV$, for $D, T \in T_n$ and $U, V \in U_n$. Now,

$$X^{-1}Y^{-1}XY = U^{-1}D^{-1}V^{-1}T^{-1}DUTV = U^{-1}D^{-1}V^{-1}DT^{-1}UTV$$

Since D, T are diagonal they commute. Now, $T^{-1}UT, D^{-1}V^{-1}D \in U_n$ by the above, so this is a product of things in U_n and hence is in U_n. We now have that $U_n = [B_n, U_n] \leq [B_n, B_n] \leq U_n$, since each $U \in U_n$ is also in B_n, so $[B_n, B_n] = U_n$. Since $[B_n, U_n] = U_n$, we have shown that the $B_n^k = U_n$ for all $k \geq 1$, so B_n is not nilpotent.

(e) Above we saw that $[B_n, B_n] = U_n$. Now, we shall show that $[U_n^{(k)}, U_n^{(k)}] \leq U_n^{(k-1)}$. Let $A, B \in U_n^{(k)}$. Then we have that $A_{i,i+j} = B_{i,i+j} = 0$ for $j < k$ by definition. We show now that $(AB)_{i,i+k} = a_i + b_i$, where a_i is the element in the ith row of the kth superdiagonal. We have,

$$(AB)_{i,i+k} = (\underbrace{0, \dots, 0}_{i}, 1, \underbrace{0, \dots, 0}_{k}, a_i, *) \cdot (\underbrace{*}_{i}, b_i, \underbrace{0, \dots, 0}_{k}, 1, 0, \dots)^T = b_i + a_i$$

Note in particular that if $A \in U_n^{(k)}$ has $a_1, \dots, a_{n-k}$ as its kth super diagonal, then A^{-1} has $-a_1, \dots, -a_{n-k}$ as its kth super diagonal since AA^{-1} has all zeros on the kth superdiagonal. Thus, $A^{-1}B^{-1}$ has $-a_i - b_i$ on it's kth super diagonal, and so $ABA^{-1}B^{-1}$ has $a_i + b_i - a_i - b_i = 0$ on it's kth superdiagonal, proving the above claim. Since if $H \leq G$ then $[H, H] \leq [G, G]$, the commutators will eventually be the trivial group, which shows that B_n is solvable.

2. We claim that each group H such that $[G : H] = p$ is maximal. Writing $|G| = p^a$, we would have that $|H| = p^{a-1}$, so the only larger group containing H would be of size p^a which is just G, so H is maximal. By theorem 1 on page 188 we have that every maximal subgroup of P of index p is normal in P. We claim if A, B are normal in G then $A \cap B$ is normal in G. This follows since given $x \in A \cap B$, $gxg^{-1} \in A$ and $gxg^{-1} \in B$. By induction the intersection of finitely many normal subgroups is normal, so we have that $\Phi(G)$ is normal.

3. (a) First we prove that $G^p \leq \Phi(G)$. Let $g \in G$ be arbitrary, and consider the quotient group G/H for any (maximal) subgroup H of index p. Then $|gH| \mid |G/H| = p$, so $g^pH = H$, thus $g^p \in H$. Next, for any x, y, $x^{-1}y^{-1}xyH = (x^{-1}H)(y^{-1}H)(xH)(yH) = H$, since $G/H \cong \mathbb{Z}/p$ is abelian, we have that $[x, y] \in H$. By the next part, $G/\Phi(G)$ is an elementary abelian p-group, and by the exact same reasoning so is $G/G^p[G, G]$. By the part after that, we must have $\Phi(G) \leq G^p[G, G]$, which shows that $\Phi(G) = G^p[G, G]$.

(b) We start by proving the following lemma:

Lemma 1. *Let G be a group and $H \trianglelefteq G$.*

(1) $G/[G,G]$ is abelian

(2) G/H is abelian iff $[G,G] \trianglelefteq H$.

Proof. (1) Let $x, y \in G$. For any $[a,b] \in [G,G]$, we have that $xy[a,b] = yxx^{-1}y^{-1}xy[a,b] = yx[x,y][a,b] \in yx[G,G]$. So, $xy[G,G] \leq yx[G,G]$, and since x, y were arbitrary, this shows that $xy[G,G] = yx[G,G]$.

(2) Suppose that G/H is abelian. Then $[x,y]H = x^{-1}y^{-1}xyH = (x^{-1}H)(y^{-1}H)(xH)(yH) = (x^{-1}H)(xH)(y^{-1}H)(yH) = H$, so $[G,G] \leq H$. Since $[G,G]$ is normal in G, it is normal in H, proving this direction.

Next suppose that $G' \trianglelefteq H$. By the third isomorphism theorem, we have that,

$$\frac{G/[G,G]}{H/[G,G]} \cong G/H$$

And since $G/[G,G]$ is abelian, any quotient of it is abelian, which completes the proof.

□

The above lemma shows that $G/\Phi(G)$ is abelian. By the fundemental theorem of finitely generated abelian groups, $G/\Phi(G)$ is a direct product of cyclic groups, say $\mathbb{Z}/n_1 \times \cdots \times \mathbb{Z}/n_k$. Notice next that since $x^p \in \Phi(G)$ for every $x \in G$, $x\Phi(G)$ has order either 1 or p. If any of the n_i were neither 1 or p, then G would have an element of order $n_i \neq 1$ and $n_i \neq p$ a contradiction. So $G/\Phi(G)$ is an elementary abelian p group.

(c) Suppose that $N \trianglelefteq G$ and that G/N is an elementary abelian p group. The proof of the fundamental theorem for finitely generated abelian groups shows that for some $xN \neq N$, there exists some $M \leq G/N$ such that $G/N = M \times \langle xN \rangle$. M is now a group of order $|G/N|/p$, so repeating this process until we run out of nonidentity elements, we can say that $G/N = \langle x_1 N \rangle \times \cdots \times \langle x_k N \rangle$. We claim that $N_j/N = \prod_{i \neq j} \langle x_1 N \rangle$ is a maximal subgroup. This is clear just by order considerations–N_j/H has order $|G/N|/p$. We also claim that $\bigcap_{j=1}^{k} N_j/N = N$. This intersection is just the identity–the element $(a_1 N, \ldots, a_k N)$ must be N in every slot (since it will be in N_j/N for each j, which has an N in slot j). Now, we shall prove the following lemma, which will complete the proof:

Lemma 2. *Suppose that $N \trianglelefteq G$ and that H/N is maximal in G/N. Then H is maximal in G.*

Proof. Suppose otherwise, then there would be some $H < L < G$. Now, $L/N < G/N$ because L/N has order strictly smaller than G/N (by Langrange's theorem). Similarly, $H/N < L/N$ since H/N has order strictly smaller than L/N. This shows that H/N is not maximal–a contradiction. $\square$

It follows that each N_j is maximal in G. Finally, since $\bigcap_{j=1}^k N_j/N = (\bigcap_{j=1}^k N_j)/N = N$, we have that for any $\ell \in \bigcap_{j=1}^k N_j$, $\ell N = N$, i.e. $\ell \in N$. Since $\bigcap_{j=1}^k N_j$ is an intersection of some of G's maximal subgroups, it follows that $\Phi(G) \leq \bigcap_{j=1}^k N_j \leq N$, which completes the proof. $\square$

4. Suppose that $g\Phi(G) = h\Phi(G)$. We want $f(g)\Phi(G') = f(h)\Phi(G')$, equivalently, we want $f(gh^{-1}) \in \Phi(G')$. Since $gh^{-1} \in \Phi(G)$, we have $gh^{-1} = x^p[y, z]$ for some $x, y, z \in G$. Now, $f(x^p[y, z]) = f(x)^p[f(y), f(z)] \in \Phi(G')$, so we have shown the induced map is well-defined. The forward direction is clear, so we shall focus on the backwards direction. Suppose that $\overline{f}$ is surjective and let $g \in G'$. We have the following commutative diagram:

$$\begin{tikzcd}
G \arrow[r, "f"] \arrow[d] & G' \arrow[d] \\
G/\Phi(G) \arrow[r, two heads, "\overline{f}"] & G'/\Phi(G')
\end{tikzcd}$$

We proceed by proving the following lemma:

Lemma 3. *Suppose that $\varphi : G \to H$ is a group homomorphism, and let $N \trianglelefteq H$. Then $[G : \varphi^{-1}(N)] = [N\varphi(G) : N] = [\varphi(G) : \varphi(G) \cap N]$. In particular, if $\varphi(G)$ is surjective, taking preimages preserves the index.*

Proof. We first have that $N\varphi(G) \leq H$, and since N is normal in H, it is normal in $N\varphi(G)$. Now let $\psi : G \to N\varphi(G)/N$ be defined by $\psi(g) = \varphi(g)N$. Firstly, this is a group homomorphism since φ is a group homomorphism. Next, ψ is surjective: notice that each $n\varphi(g)N \in N\varphi(G)/N$ is equal to $\varphi(g)n_1N = \varphi(g)N$ because N is normal, so $\psi(g) = \varphi(g)N = n\varphi(g)N$. Finally, $\ker \psi = \varphi^{-1}(N)$. By the first isomorphism theorem, we have shown that

$$G/\varphi^{-1}(N) \cong N\varphi(G)/N$$

Next, by the second isomorphism theorem, $N\varphi(G)/N \cong \varphi(G)/(S \cap \varphi(G))$. We have concluded $[G : \varphi^{-1}(N)] = [N\varphi(G) : N] = [\varphi(G) : \varphi(G) \cap N]$ by Langrange's theorem. Finally, if $\varphi(G)$ is surjective, we have $[G : \varphi^{-1}(N)] = [H : N]$, which completes the proof. $\square$

Now suppose that $H \leq G$ is maximal, and that $\overline{f}(H) = f(H)/\Phi(G')$ were not maximal. Then we would a maximal $L/\Phi(G')$ such that have $f(H)/\Phi(G') < L/\Phi(G') < G'/\Phi(G')$. Now, since $G'/\Phi(G')$ is a p-group, $L/\Phi(G')$ has index p. Since $\overline{f}$ is onto, $\overline{f}^{-1}(L/\Phi(G'))$ also has index p, strictly containing $H/\Phi(G)$. The containment is strict otherwise $\overline{f}(H/\Phi(G)) = \overline{f}(\overline{f}^{-1}(L/\Phi(G'))) = L/\Phi(G')$ since $\overline{f}$ is surjective. But this is a contradiction–for $H/\Phi(G)$ has minimal index, as

$$[G/\Phi(G) : H/\Phi(G)] = \frac{|G|/|\Phi(G)|}{|H|/|\Phi(G)|} = \frac{|G|}{|H|} = p$$

Since $f(H)/\Phi(G')$ is maximal, we must have $f(H)$ maximal too by the above claims. Let $h : G/\ker f \to G'$ be the induced map. Notice that h is injective. Now, we have that

$$h\left(\bigcap_{H<G \text{ maximal}} H/\ker f\right) = \bigcap_{H<G \text{ maximal}} h(H/\ker f) \geq \Phi(G'),$$

since $\bigcap_{H<G \text{ maximal}} h(H/\ker f)$ has only some of the maximal subgroups in the intersection. Since the induced map is surjective so is f, so $f(G) \geq \Phi(G')$. The hard part is now over. Let $g' \in G'$. Find $g \in G$ such that $f(g)\Phi(G) = g'\Phi(G')$. Thus, $f(g) = g' \cdot \eta$ for some $\eta \in \Phi(G')$. By the above find $\zeta \in G$ such that $f(\zeta) = \eta$. Now, $f(g\zeta^{-1}) = g' \cdot \eta \cdot \eta^{-1} = g'$, and we are done. ■

5. (a) We shall instead proceed by showing that $\Phi(G)$ is the set of nongenerators of G. Suppose that $\langle H \rangle < G$. Then $\langle H \rangle < N$ for some maximal subgroup N. Then for any $x \in \Phi(G)$, $\langle x, H \rangle < N$ too since $x \in N$. Now let $x \in G$ and suppose that, for any $H \subset G$, if $\langle H \rangle < G$ then $\langle x, H \rangle < G$. Letting N be an arbitrary maximal subgroup of G, we see that $\langle N \rangle = N$, and $N \leq \langle x, N \rangle < G$, so $\langle x, N \rangle = N$ as well. This tells us that $x \in N$. Since this was true for an arbitrary maximal subgroup, we must have $x \in \Phi(G)$. Now, suppose that $\{ \overline{x_1}, \dots, \overline{x_n} \}$ is a generating set for $G/\Phi(G)$. Notice that $\{ x_1, \dots, x_n \}$ will generate G iff $\langle x_1, \dots, x_n \rangle \geq \Phi(G)$. The forward direction is obvious. Let $x \in G$ and write $x\Phi(G) = \prod x_i \Phi(G)$. Then $x\varphi = \prod x_i \cdot \eta$ for some $\varphi, \eta \in \Phi(G)$. Then $x = \prod x_i \eta \varphi^{-1}$, and since $\langle x_1, \dots, x_n \rangle \geq \Phi(G)$, we can write $\eta\varphi^{-1} = \prod x_j$, so $x = \prod x_i \prod x_j \in \langle x_1, \dots, x_n \rangle$. Now, if $\langle x_1, \dots, x_n \rangle$ did not generate G, then we could add the remaining elements of $\Phi(G)$ to make it generate all of G. But this can't be–adding elements of $\Phi(G)$ to our generating set cannot make it generate all of G by the above, a contradiction.

   (b) Write $G/\Phi(G) = (\mathbb{Z}/p)^n$. Every minimal system of generators of this group has n elements, since this is an n dimensional vector space over $\mathbb{Z}/p$, and each minimal system of generators is a basis. We can pull one of these back to get a generating set for G. If G had a generating set with less than n elements, then $(\mathbb{Z}/p)^n$ would have a generating set with less than n elements, a contradiction. We have won.