Math 120 HW 4 Solutions

Xiaoyu He, with Questions 5A/5B/5C by Prof. Church

April 27, 2018

4.2.8

Prove that if H has finite index n then there is a normal subgroup K of G with $K \leq H$ and $|G : K| \leq n!$.

Consider the action of G by left-multiplication on left cosets of H, which corresponds to a homomorphism $\alpha\colon G \to \mathrm{Perm}(G//H)$. We claim that $K = \ker(\alpha)$ is the desired subgroup. First, we check that $K \leq H$. An element $k \in G$ belongs to K if and only if

$$k \cdot (gH) = gH \qquad \text{for all } gH \in G//H.$$

In particular, if $k \in K$ then $k \cdot H = H$. This is true if and only if $k \in H$, so this shows $k \in K \implies k \in H$, i.e. $K \leq H$.

It remains to show that $[G : K] \leq n!$. By the first isomorphism theorem, we have $G/\ker(\alpha) \cong \mathrm{im}(\alpha)$, which is a subgroup of $\mathrm{Perm}(G//H)$. Since $[G : H] = n$, the coset space $G//H$ is a set with n elements, so $\mathrm{Perm}(G//H)$ has cardinality $n!$. Therefore

$$[G : K] = |G/K| = |G/\ker(\alpha)| \leq |\mathrm{Perm}(G//H)| = n!.$$

4.3.10*

Let σ be the 5-cycle (12345) in S_5. In each of (a) to (c) find and explicit element $\tau \in S_5$which accomplishes the specified conjugation:

(a) $\tau\sigma\tau^{-1} = \sigma^2$**.**
$\tau = 13524$.
(b) $\tau\sigma\tau^{-1} = \sigma^{-1}$**.**
$\tau = 15432$.
(c) $\tau\sigma\tau^{-1} = \sigma^{-2}$**.**
$\tau = 14253$.

4.3.11*

In each of (a)-(d) determine whether σ_1 and σ_2 are conjugate. If they are, give an explicit permutation τsuch that $\tau\sigma_1\tau^{-1} = \sigma_2$.

(a) $\sigma_1 = (12)(345)$ **and** $\sigma_2 = (123)(45)$**.**
$\tau = 45123$.
(b) $\sigma_1 = (15)(372)(10\ 6\ 8\ 11)$ **and** $\sigma_2 = (3\ 7\ 5\ 10)(49)(13\ 11\ 2)$**.**
$\tau = 4\ 2\ 13\ 1\ 9\ 7\ 11\ 5\ 6\ 3\ 10\ 12\ 8$.
(c) $\sigma_1 = (15)(372)(10\ 6\ 8\ 11)$ **and** $\sigma_2 = \sigma_1^3$**.**
Not conjugate.
(d) $\sigma_1 = (13)(246)$ **and** $\sigma_2 = (35)(24)(16)$**.**
Not conjugate.

5.1.1*

Show that the center of a direct product is the direct product of centers:

$$Z(G_1 \times G_2 \times \cdots \times G_n) = Z(G_1) \times Z(G_2) \times \cdots \times Z(G_n).$$

Deduce that a direct product of groups is abelian if and only if each of the factors is abelian.

Check that if $x, y \in G_1 \times \cdots \times G_n$, then $xy = yx$ iff $x_i y_i = y_i x_i$ in each coordinate $i \leq n$. Thus, a fixed x lies in $Z(G_1 \times \cdots \times G_n)$ iff each coordinate x_i lies in the center $Z(G_i)$.

A group G is abelian iff $Z(G) = G$, so this implies the second claim.

5.1.5*

Exhibit a normal subgroup of $Q_8 \times Z_4$ (note that every subgroup of each factor is normal).

Any product $H \times K$ of subgroups $H \leq Q_8$ and $K \leq Z_4$ is normal, for example $Q_8 \times Z_2$.

Question 1

Recall that F_n denotes a free group on n elements.

***In at most two sentences,* prove that F_2 is not isomorphic to F_3.**

The number of homomorphisms from $F_2 \to Z_2$ is 4 but the number of homomorphisms from $F_3 \to Z_2$ is 8, so F_2 and F_3 can't be isomorphic.

Remarks: in general, the number of homomorphisms $n(F_n, G)$ from the free group on n elements to any finite group G will be $|G|^n$, by exactly the same argument as in HW3, Problem 6. Also, we used the following fact [you did not need to write any of this]:

Fact 1. *If $A \cong B$ are two isomorphic groups, then $n(A, G) = n(B, G)$ for any third group G.*

Proof. Let $\phi : A \to B$ be an isomorphism from A to B. Then, ϕ induces a bijection between $\mathrm{Hom}(A, G)$ and $\mathrm{Hom}(B, G)$. Given any $f : B \to G$, precomposition with ϕ gives $f \circ \phi : A \to G$. Similarly, in the reverse direction any $f' : A \to G$ can be turned into a map $f' \circ \phi^{-1} : B \to G$. It is easy to check that this is a bijection. □

Question 2

Given a group G, the *center* of G is the subgroup

$$Z(G) = \{z \in G | zg = gz \textbf{ for all } g \in G\}$$

of elements that commute with every element of G. The center $Z(G)$ is an abelian subgroup of G. (You may assume this without proof, but you should understand why it is true.)

(a) Prove that $Z(G)$ is a normal subgroup of G.

If $h \in Z(G)$ then for any $g \in G$, $gh = hg$, so

$$\begin{aligned} ghg^{-1} &= hgg^{-1} \\ &= h, \end{aligned}$$

whence $ghg^{-1} \in Z(G)$ as well. Thus $Z(G)$ is a normal subgroup.

(b) Prove that if the quotient group $G/Z(G)$ is cyclic, then G is abelian.

Let $xZ(G)$ be a generator for the cyclic group $G/Z(G)$, and suppose $g, h \in G$ are any two elements. Then, g, h each lie in some coset $x^i Z(G)$, so there exist $i, j \in \mathbb{N}$ for which $g \in x^i Z(G)$ and $h \in x^j Z(G)$. Write $g = x^i g'$ and $h = x^j h'$.

Since $g', h' \in Z(G)$, they commute with everything. Thus,

$$\begin{aligned} gh &= x^i g' x^j h' \\ &= x^{i+j} g' h' \\ &= x^{j+i} h' g' \\ &= x^j h' x^i g' \\ &= hg. \end{aligned}$$

Thus, every pair of elements $g, h \in G$ commute, and G is abelian.

(c) (Optional) Is it true that if the quotient group $G/Z(G)$ is abelian, then G is abelian? Prove or give a counterexample.

It's false. If $G = Q_8$ is the quaternion group, the center is $Z(G) = \{\pm 1\}$ and the quotient $G/Z(G)$ is isomorphic to $Z_2 \times Z_2$. The dihedral group D_8 is another counterexample. So is the group of matrices of the form $\begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{bmatrix}$.

Question 3

Give a presentation for $G = \mathbb{Z} \times \mathbb{Z}$; i.e. fill in the blanks:

$$\mathbb{Z} \times \mathbb{Z} \cong \langle \_\_\_\_ | \_\_\_\_\_\_\_\_ \rangle.$$

(You do not have to prove your presentation is correct.)

One presentation is $\mathbb{Z} \times \mathbb{Z} = \langle x, y | xyx^{-1}y^{-1} = 1 \rangle$. Here's how one proves the presentation is correct.

First, there is a homomorphism $\phi : F(x, y) \to \mathbb{Z} \times \mathbb{Z}$ sending x to $(1, 0)$ and y to $(0, 1)$. Because $\mathbb{Z} \times \mathbb{Z}$ is abelian, $\phi(xyx^{-1}y^{-1}) = (0, 0)$, so ϕ vanishes on $xyx^{-1}y^{-1}$.

This means that $\ker \phi \ni xyx^{-1}y^{-1}$, so in particular since $\ker \phi$ is a normal subgroup of $F(x, y)$, it contains the normal closure N of $\langle xyx^{-1}y^{-1} \rangle$. This means $\phi(N) = 0$, and ϕ is constant on every coset of N. Define the induced map $\phi^* : F(x, y)/N \to \mathbb{Z} \times \mathbb{Z}$ to be the map whose value on a coset is just the value of ϕ on any representative: $\phi^*(aN) = \phi(a)$. Since ϕ is constant on each coset, this map is well-defined. (Note: what we did is essentially prove the universal property of group quotients).

Note that $\langle x, y | xyx^{-1}y^{-1} = 1 \rangle = F(x, y)/N$ by definition, so it suffices to show ϕ^* is an isomorphism. Because $\phi^*(x^m y^n) = (m, n)$, ϕ^* is surjective. Meanwhile, every single element of $\langle x, y | xyx^{-1}y^{-1} = 1 \rangle$ is a product of powers of x and y, which can be commuted through each other because $xy = yx$. Thus, every single element of $\langle x, y | xyx^{-1}y^{-1} = 1 \rangle$ has a representation as $x^m y^n$. Such an element is sent to the identity $(0, 0)$ of $\mathbb{Z} \times \mathbb{Z}$ iff $m = n = 0$, so ϕ^* is injective. Putting it all together that ϕ^* is an isomorphism.

Question 4

(Optional.) Prove that S_3 has the presentation

$$S_3 \cong \langle a, b | a^2 = 1, b^2 = 1, ababab = 1 \rangle.$$

In Example 6.3.1 it is shown that to show that $\langle S, R \rangle$ is a presentation for a given finite group G, it is sufficient to check that (i) S is a generating set of G, and (ii) any group generated by S satisfying the relations in R has order $\leq |G|$.

Let $G = \langle a, b | a^2 = 1, b^2 = 1, ababab = 1 \rangle$. We want to define a homomorphism $\varphi \colon G \to S_3$. To do this, define the homomorphism $f \colon F_2 \to S_3$ sending $a \mapsto (1\ 2)$ and $b \mapsto (2\ 3)$. Since $G = F_2 / \langle\langle a^2, b^2, ababab \rangle\rangle$, as we saw in class, to check that this defines a homomorphism $\varphi \colon G \to S_3$ it suffices to check that $\langle\langle a^2, b^2, ababab \rangle\rangle \subset \ker(f)$. Since this is the normal closure of the three elements, to do this it suffices to check that $f(a^2) = 1$, $f(b^2) = 1$, and $f(ababab) = 1$. We can do this directly, noting in the latter case that $f(ab) = (1\ 2\ 3)$.

Moreover, since (1 2) and (2 3) generate S_3 [as you can easily check by hand], we know that $\varphi\colon G \to S_3$ is *surjective*. To finish the proof, we will show that $|G| \le 6$; since it surjects to a set of size 6, this will show that actually $|G| = 6$ and that φ is an isomorphism.

To show that $|G| \le 6$, we check that every element of G is equal to one of the six elements $1, a, b, ab, ba, aba$. Indeed, since $a^2 = 1$ and $b^2 = 1$ in G, and also $a = a^{-1}$ and $b = b^{-1}$, any word of G can be reduced to an alternating product consisting only of a's and b's: $abababab\cdots$ or $bababab\cdots$. Right-multiplying $ababab = 1$ by ba, we find that $abab = ba$. Similarly, $baba = ab$. This shows that every alternating word of length at least 4 can be reduced further in length. It follows that all distinct words of G are one of $1, a, b, ab, ba, aba, bab$. But we have one duplicate here: $aba = bab$ (from right-multiplying the last relation by bab). Thus, we have the desired conclusion that any group generated by S satisfying $a^2 = 1, b^2 = 1$ and $ababab = 1$ has order at most 6.

(Double Optional.) What additional relation(s) do you need to add to the following to get a presentation of S_4?

$$S_4 \cong \langle a, b, c | a^2 = 1, b^2 = 1, c^2 = 1, ababab = 1, bcbcbc = 1, \_\_\_\_ \rangle$$

One way is to add the relation $acac = 1$.

Question 5A

a) The homomorphism $\varphi\colon F_2 \to Z_2 \times Z_2$ takes a word $w = a^{n_1}b^{m_1}\dots a^{n_k}b^{m_k}$ in F_2 to $(x^{\sum_i n_i}, x^{\sum_i m_i}) \in Z_2 \times Z_2$. Hence we conclude that the kernel is

$$K = \left\{ \text{words } w = a^{n_1}b^{m_1}\dots a^{n_k}b^{m_k} \text{ such that } \sum_{i=1}^{k} n_i, \sum_{i=1}^{k} m_i \text{ are both even} \right\}.$$

b) We start by choosing a set of coset representatives for K in F_2:

$$C = \{1, a, b, ab\}.$$

We define the function $f\colon F_2 \to C$ so that $\varphi(g) = \varphi(f(g))$; in other words, $f(g)$ is the representative of the coset Kg. Since $f(g)$ and g map to the same element of $Z_2 \times Z_2$, the element $g \cdot f(g)^{-1}$ lies in the kernel K for any $g \in F_2$.

Lemma 1. The set

$$X = \left\{ cl \cdot f(cl)^{-1} \,|\, c \in C,\, l \in \{a, b, a^{-1}, b^{-1}\} \right\}$$

is a set of generators for K.

Proof. Let $w = l_1 \cdots l_n$ be a word in F_2, with $l_i \in \{a, b, a^{-1}, b^{-1}\}$. Define $c_0 = 1$ and $c_k = f(l_1 \cdots l_k)$ for $1 \le k \le n$. We can write $f(w)$ as a telescoping product

$$f(w) = wc_n^{-1} = \left(c_0 l_1 c_1^{-1}\right)\left(c_1 l_2 c_2^{-1}\right)\cdots\left(c_{n-1} l_n c_n^{-1}\right).$$

Now notice that by definition,

$$c_k = f(l_1 \cdots l_k) = f\big(f(l_1 \cdots l_{k-1}) l_k\big) = f(c_{k-1} l_k) \in C.$$

This means that each of the parenthesized terms

$$c_{k-1} l_k c_k^{-1} = c_{k-1} l_k f\left(c_{k-1} l_k\right)^{-1}$$

is one of the generators in X!
This shows that for *any* w in F_2, the element wc_n^{-1} is in the subgroup $\langle X \rangle$ generated by X. In particular, if $w \in K$, then $c_n = f(l_1 \cdots l_n) = f(w) = 1$, so the previous sentence says that $w \in \langle X \rangle$. This concludes the proof of the lemma. □

Lemma 2. The subset

$$S = \left\{ cl \cdot f(cl)^{-1} \,|\, c \in C,\, l \in \{a, b\} \right\}$$

still generates K.

Proof. We will show that each generator $x = cl \cdot f(cl)^{-1} \in X$ for $c \in C$ and $l \in \{a^{-1}, b^{-1}\}$ is the inverse of some generator in S. This shows that $\langle S \rangle = \langle X \rangle$, and the previous lemma showed that $\langle X \rangle = K$.

Suppose for instance $l = a^{-1}$ and consider $g = f(ca^{-1}) \in C$. The generator $ca^{-1} \cdot f(ca^{-1})^{-1} \in X$ is equal to $ca^{-1} \cdot g^{-1} \in K$, so call this element k. We obtain

$$a^{-1}g^{-1} = c^{-1}k$$

and inverting both sides

$$ga = k^{-1}c.$$

Taking the coset representative yields then

$$f(ga) = f(k^{-1}c) = f(c) = c$$

since $c \in C$. In particular we get

$$x = ca^{-1} \cdot f(ca^{-1})^{-1} = ca^{-1}g^{-1} = \left(gac^{-1}\right)^{-1} = \left(ga \cdot f(ga)^{-1}\right)^{-1}.$$

Therefore x^{-1} is the generator $ga \cdot f(ga)^{-1}$ in S, as claimed. The proof is identical for b^{-1} in place of a^{-1}. ☐

The preceding gives an **upper bound** of 5 generators.

We now illustrate how you could show that the nontrivial elements of S generate K freely (to get a **lower bound** and show that K can't be generated by fewer than 5 elements). Let's start by making these elements explicit: in the following two tables we put the c's into rows, and the letters l's into columns. We first describe $f(cl)^{-1}$:

	a	b
1	a^{-1}	b^{-1}
a	1	$(ab)^{-1}$
b	$(ab)^{-1}$	1
ab	b^{-1}	a^{-1}

and then the corresponding generator $cl \cdot f(cl)^{-1}$:

	a	b
1	1	1
a	a^2	1
b	$ba(ab)^{-1}$	b^2
ab	$abab^{-1}$	$abba^{-1}$

Getting rid of the trivial elements, we obtain that S consists of the five elements

$$S = \left\{a^2, ba(ab)^{-1}, b^2, aba^{-1}, abba^{-1}\right\}.$$

Notice in particular that all these generators are already in reduced form as written, and that no two generators are each other's inverses.

We saw in the proof of the second lemma that X consists of the elements of S together with their inverses; therefore

$$X = \left\{a^2, ba(ab)^{-1}, b^2, aba^{-1}, abba^{-1}, a^{-2}, b^{-2}, ab^{-1}(ab)^{-1}, ba^{-1}(ab)^{-1}, aba^{-1}b^{-1}\right\}.$$

Since the extended generators obtained are all distinct, each of the element in X uniquely pinpoints the pair $(c, l) \in C \times \left\{a^{\pm 1}, b^{\pm 1}\right\}$ which gave rise to it. For each element $s \in X$, we call the element l in this pair the "l-factor of s".

Let now $s_1, s_2 \in X$. Suppose that some cancellation between the l-factors of s_1 and s_2 happens. Then it can be checked (either formally or by a case-by-case analysis) that $s_1 = s_2^{-1}$: in particular, one and only one among s_1 and s_2 belongs to S (the other belongs to $X \backslash S$).

Therefore, when we consider a word $w = s_1 \ldots s_n$ over the generators of S, we cannot have any cancellation of the relevant l-factors: in particular, none of these words cancel trivially to 1. This proves that S is a set of free generators.

Summing up, $K \cong F_5$ is a free group on five generators, and a choice of generators is given by

$$S = \left\{a^2, ba(ab)^{-1}, b^2, aba^{-1}, abba^{-1}\right\}.$$

Question 5B

Denote as usual by a and b the free generators of F_2. Consider the homomorphism $\varphi\colon F_2 \to \mathbb{Z}\times\mathbb{Z}$ sending $a \mapsto (x,1)$ and $b \mapsto (1,x)$, where we write $\mathbb{Z} = \{\ldots, x^{-2}, x^{-1}, 1, x, x^2, \ldots\}$.

Lemma 3. The kernel of φ is the commutator subgroup N.

Proof. Since $\mathbb{Z}\times\mathbb{Z}$ is abelian, we know that $\varphi(x^{-1}y^{-1}xy) = \varphi(x)^{-1}\varphi(y)^{-1}\varphi(x)\varphi(y) = 1 \in \mathbb{Z}\times\mathbb{Z}$ for any $x, y \in F_2$. This shows that every commutator is contained in $\ker\varphi$, so $N \subset \ker\varphi$. For the other direction, you could look to Question 3, where we showed that $\ker\varphi = \langle\langle a^{-1}b^{-1}ab$, so to show that $\ker\varphi \subset N$ we just need to check that $a^{-1}b^{-1}ab \in N$ which is true by definition. Or, you can argue as follows (which is essentially just repeating the proof of Question 3).

Temporarily denote $A = F_2/N$. Since $N \subset \ker\varphi$, the map $\varphi\colon F_2 \to \mathbb{Z}^2$ descends to a map $\overline{\varphi}\colon A \to \mathbb{Z}^2$ sending $\overline{a} \mapsto (x,1)$ and $\overline{b} \mapsto (1,x)$. The kernel of $\overline{\varphi}$ is isomorphic to $\ker\varphi/N$, so our goal is to prove that $\overline{\varphi}$ is injective. Assume for a contradiction that $\ker\overline{\varphi}$ contains a nontrivial element k. The group A is abelian since $\overline{x}\overline{y}\overline{x}^{-1}\overline{y}^{-1} = \overline{xyx^{-1}y^{-1}} = \overline{1}$ for any $\overline{x}, \overline{y} \in A$. Since A is generated by $\overline{x}$ and $\overline{y}$, we can write $k = \overline{x}^a\overline{y}^b$ where either a or b is nonzero. But then $\overline{\varphi}(k) = \varphi(x^a y^b) = (a,b) \in \mathbb{Z}\times\mathbb{Z}$ which is not the identity. This contradiction shows that $\ker\varphi = N$, as claimed. $\square$

Consider the function Etch-A-Sketch from words $w \in F_2$ to *finite subsets* of $\mathbb{Z}^2$, defined as follows. Given a reduced word $w \in F_2$, write it as $w = l_1 \cdots l_n$ where $l_i \in \{a, b, a^{-1}, b^{-1}\}$. For each $k = 1, \ldots, n$, consider the element $\varphi(l_1 \cdots l_k) \in \mathbb{Z}^2$. Finally, the subset Etch-A-Sketch(w) is defined to be the set of all such elements:

$$\text{Etch-A-Sketch}(w) = \big\{\varphi(l_1 \cdots l_k)\big\}_{k=1}^{n}$$

The key observation is that

$$x \in N \text{ and } w = xy \in F_2 \implies \text{Etch-A-Sketch}(w) \subset \text{Etch-A-Sketch}(x) \cup \text{Etch-A-Sketch}(y). \qquad (*)$$

To see this, first write $x = x_1 \cdots x_n$ and $y = y_1 \cdots y_m$. If we did *not* reduce the concatenation xy, then the terms appearing would be

$$\varphi(x_1), \varphi(x_1x_2), \ldots, \varphi(x_1x_2\cdots x_n), \varphi(x_1x_2\cdots x_ny_1), \ldots, \varphi(x_1x_2\cdots x_ny_1\cdots y_m).$$

However, we have assumed that $x \in N$, so $\varphi(x_1x_2\cdots x_n)$ is the identity. Therefore the terms appearing can be rewritten as

$$\varphi(x_1), \varphi(x_1x_2), \ldots, \varphi(x_1x_2\cdots x_n), \varphi(y_1), \ldots, \varphi(y_1\cdots y_m),$$

or in other words Etch-A-Sketch$(x) \cup$ Etch-A-Sketch(y). To obtain w from the concatenation xy we simply cancel adjacent terms; this has the effect of possibly removing some terms from Etch-A-Sketch(w), which is why we have Etch-A-Sketch$(w) \subset$ Etch-A-Sketch$(x) \cup$ Etch-A-Sketch(y).

Now assume for a contradiction that N is generated by a finite set $S = \{g_1, \ldots, g_k\}$. Let X be the finite set $X = \text{Etch-A-Sketch}(g_1) \cup \cdots \cup \text{Etch-A-Sketch}(g_k)$. The generators $g = g_1, \ldots, g_k$ all have the property that Etch-A-Sketch$(g) \subset X$ by definition. But the key observation $(*)$ shows that this property is *preserved under multiplication*. Therefore every element g of the subgroup $\langle S\rangle$ generated by S has Etch-A-Sketch$(g) \subset X$.

Therefore to obtain a contradiction, we just need to show that for every finite subset $X \subset \mathbb{Z}^2$, there are elements $g \in N$ for which Etch-A-Sketch$(g) \not\subset X$. This is pretty easy. For example, let M be the cardinality of X and choose $g = a^M b a^{-M} b^{-1}$. The finite subset Etch-A-Sketch(g) is equal to

$$\{(1,1), (x,1), \ldots, (x^M,1), (x^M,x) \ldots, (x,x), (1,x)\}$$

which has cardinality $2M$, so it cannot be contained in X. $\square$

Question 5C

There are various fancy ways to prove elements generate a free group, but in this case there is a concrete argument that you could have discovered by experimentation: when you multiply these generators together, the entries of the matrices always get larger! So there's no way to end up back at the identity matrix. That's the argument we use in the proof below.

By construction, G is the image of a surjective homomorphism $F_2 \to G$ from the free group on the two generators; we need to prove that that this homomorphism is injective, so it is an isomorphism $F_2 \simeq G$. It suffices to prove that the kernel is trivial; in other words, that any nontrivial reduced word in the generators x and y represents a matrix that is not the identity matrix.

Assume for a contradiction that there is a nontrivial word w in the kernel. We can assume that w starts with x or x^{-1} (because if it does not, we can consider instead the word xwx^{-1}, which is still in the kernel and does start with x).

Therefore we can write $w = x^{n_1}y^{n_2}\cdots x^{n_{m-1}}y^{n_m}$ or $w = x^{n_1}y^{n_2}\cdots x^{n_m}$ where $n_i \in \mathbb{Z}$ and all $n_i \neq 0$ (depending on whether w ends with $y^{\pm}$ or $x^{\pm}$) for some $m \geq 1$.

Let the matrices $M_0, M_1, M_2, \dots, M_m$ be the running products

$$M_0 = I,\ M_1 = x^{n_1},\ M_2 = x^{n_1}y^{n_2}, \dots, \quad M_i = x^{n_1}y^{n_2}\cdots z^{n_i}$$

(where z would be x or y depending on whether i is even or odd). Write $M_i = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}$.

When i is even we have $M_{i+1} = M_i x^{n_{i+1}}$, so

$$\begin{bmatrix} a_{i+1} & b_{i+1} \\ c_{i+1} & d_{i+1} \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \begin{bmatrix} 1 & 2n_{i+1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_i & 2n_{i+1}\cdot a_i + b_i \\ c_i & 2n_{i+1}\cdot c_i + d_i \end{bmatrix}.$$

When i is odd we have $M_{i+1} = M_i y^{n_{i+1}}$, so

$$\begin{bmatrix} a_{i+1} & b_{i+1} \\ c_{i+1} & d_{i+1} \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2n_{i+1} & 1 \end{bmatrix} = \begin{bmatrix} 2n_{i+1}\cdot b_i + a_i & b_i \\ 2n_{i+1}\cdot d_i + c_i & d_i \end{bmatrix}$$

In other words, there are two sequences α_i and β_i such that

$$M_i = \begin{bmatrix} \alpha_{i-1} & \alpha_i \\ \beta_{i-1} & \beta_i \end{bmatrix} \text{ when } i \text{ is odd} \qquad M_i = \begin{bmatrix} \alpha_i & \alpha_{i-1} \\ \beta_i & \beta_{i-1} \end{bmatrix} \text{ when } i \text{ is even}$$

Notice that the matrix formulas above show that the sequence α_i satisfies the recursion

$$\alpha_{i+1} = 2n_{i+1}\alpha_i + \alpha_{i-1}$$

(starting with $\alpha_{-1} = 0$ and $\alpha_0 = 1$, since $M_0 = I$).

We now prove by induction that $|\alpha_i|$ is a strictly increasing sequence: $|\alpha_{i+1}| > |\alpha_i|$ for all $i \geq 0$. This is certainly true for $i = -1$, since $|\alpha_1| = 1 > |\alpha_0| = 0$. Applying the triangle inequality to the recursion above shows that

$$|\alpha_{i+1}| \geq |2n_{i+1}\alpha_i| - |\alpha_{i-1}|.$$

Since $n_{i+1} \neq 0$, we know that $|2n_{i+1}\alpha_i| \geq |2\alpha_i| = 2|\alpha_i|$. And by induction, we can assume that $|\alpha_{i-1}| < |\alpha_i|$. Therefore we conclude that

$$|\alpha_{i+1}| \geq |2n_{i+1}\alpha_i| - |\alpha_{i-1}| \geq 2|\alpha_i| - |\alpha_{i-1}| > 2|\alpha_i| - |\alpha_i| = |\alpha_i|.$$

Therefore $|\alpha_{i+1}| > |\alpha_i|$, which is precisely what we needed to continue the induction.

The end of the proof is now easy. The word w represents the matrix M_m, whose top row has entries α_{m-1} and α_m. Since the sequence $|\alpha_i|$ is strictly increasing and $m \geq 1$, we know that $|\alpha_m| > |\alpha_0| = 1$. Therefore the top row of M_m contains an entry with absolute value bigger than 1, so $w = M_m$ cannot be the identity.