GENERATING SETS

KEITH CONRAD

1. Introduction

In $\mathbf{R}^n$, every vector can be written as a (unique) linear combination of the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$. A notion weaker than a basis is a spanning set: a set of vectors in $\mathbf{R}^n$ is a spanning set if its linear combinations fill up the whole space. The difference between a spanning set and a basis is that a spanning set may contain more vectors than necessary to span the space. For instance, $\{(1,0),(3,1),(2,1)\}$ is a spanning set for $\mathbf{R}^2$ but is not a basis. Omitting a vector from this spanning set will give a basis of $\mathbf{R}^2$. A basis is a minimal spanning set. All bases of $\mathbf{R}^n$ have size n.

In a group, the analogue of a spanning set is called a generating set.

Definition 1.1. In a group G, a subset $X \subset G$ is a *generating set* for G if every $g \in G$ can be written as a product of powers of elements taken from X:

$$g = x_1^{a_1} x_2^{a_2} \cdots x_r^{a_r}, \tag{1.1}$$

where $x_i \in X$ and $a_i \in \mathbf{Z}$. We also say that X *generates* G and write $G = \langle X \rangle$. If G has a finite generating set, we say G is a *finitely generated* group.

Instead of using exponents in the definition of a generating set, we could write powers as repeated copies of the same factor or its inverse (so $g^3h^{-2} = gggh^{-1}h^{-1}$). Therefore a subset X of G is a generating set for G when every element of G is a product of elements from X and inverses of elements from X. This is closer to what the idea of "generating set" sounds like: through repeated use of the group operations (multiplication and inversion) we can produce all elements of G from X.

Example 1.2. Every permutation in S_n is a product of cycles, so the cycles in S_n are a generating set of S_n.

Example 1.3. The group $\mathbf{Z}/(m) \times \mathbf{Z}/(n)$ is generated by $(1,0)$ and $(0,1)$, since $(\overline{a}, \overline{b}) = a(\overline{1}, \overline{0}) + b(\overline{0}, \overline{1})$

Example 1.4. A group has a one-element generating set exactly when it is a cyclic group. For instance, $\mathbf{Z}$ has the one-element generating sets $\{1\}$ and $\{-1\}$.

Example 1.5. Dihedral groups have two generators: $D_n = \langle r, s \rangle$ and every element is r^i or $r^i s$. For a general group with two generators x and y, we usually can't write elements in the condensed form $x^m y^n$ for some m and n in $\mathbf{Z}$, *e.g.*, $yxyx^2$ is not x^3y^2 (or y^2x^3). We'll see an example in Section 4.

Example 1.6. The infinite nonabelian matrix group $\{\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) : a = \pm 1, b \in \mathbf{Z}\}$ is finitely generated. Since $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^b$ and $\left(\begin{smallmatrix} -1 & b \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^b\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, this group has generating set $\{\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\}$.

Example 1.7. The group $\mathbf{Q}$ is not finitely generated: a finite set of rational numbers has a common denominator, say N, and the subgroup of $\mathbf{Q}$ generated by these rational numbers (their integral multiples and sums thereof) will only give rise to rational numbers with denominators dividing N. Not all rationals have such denominators (try $1/(N+1)$), so $\mathbf{Q}$ doesn't have a finite set of generators as an additive group.

Example 1.8. A finitely generated group is at most countable, so an uncountable group is not finitely generated.

In this handout, we look at generating sets of the groups S_n, A_n, $\mathrm{SL}_2(\mathbf{Z})$, $\mathrm{GL}_n(\mathbf{R})$, and $\mathrm{SL}_n(\mathbf{R})$. The following table summarizes some the generating sets we will obtain for various groups, and indicates where the proofs are found. At the end we discuss minimal generating sets, which have some counterintuitive properties in nonabelian groups.

Group	Generating Set	Size	Where
S_n, $n \geq 2$	(ij)'s	$\frac{n(n-1)}{2}$	Theorem 2.1
	$(12), (13), \ldots, (1n)$	$n-1$	Theorem 2.2
	$(12), (23), \ldots, (n-1\ n)$	$n-1$	Theorem 2.3
	$(12), (12\ldots n)$ if $n \geq 3$	2	Theorem 2.5
	$(12), (23\ldots n)$ if $n \geq 3$	2	Corollary 2.6
	$(ab), (12\ldots n)$ if $(b-a, n) = 1$	2	Theorem 2.8
A_n, $n \geq 3$	3-cycles	$\frac{n(n-1)(n-2)}{3}$	Lemma 3.1
	$(1ij)$'s	$(n-1)(n-2)$	Theorem 3.2
	$(12i)$'s	$n-2$	Theorem 3.3
	$(i\ i+1\ i+2)$'s	$n-2$	Theorem 3.4
	$(123), (12\ldots n)$ if $n \geq 4$ odd	2	Theorem 3.5
	$(123), (23\ldots n)$ if $n \geq 4$ even	2	Theorem 3.5
$\mathrm{SL}_2(\mathbf{Z})$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	2	Theorem 4.1
$\mathrm{GL}_n(\mathbf{R}), \mathrm{SL}_n(\mathbf{R})$	Elementary Matrices	Infinite	Theorem 5.2

2. Generators for S_n

The group S_n is generated by its cycles. The following theorem shows the 2-cycles (the transpositions) are enough to generate S_n.

Theorem 2.1. *For $n \geq 2$, S_n is generated by its transpositions.*

Proof. This is clear for $n = 1$ and 2. For $n \geq 3$, we note $(1) = (12)^2$ and every cycle of length > 2 is a product of transpositions:

$$(i_1 i_2 \ldots i_k) = (i_1 i_2)(i_2 i_3) \cdots (i_{k-1} i_k).$$

For example,

$$(13526) = (13)(35)(52)(26).$$

Since the cycles generate S_n, and products of transpositions give us all cycles, the transpositions generate S_n. □

Since transpositions have order 2 (though *not* every element of order 2 is a transposition, *e.g.*, $(12)(34)$), Theorem 2.1 tells us S_n is generated by elements of order 2.

The total number of transpositions in S_n is $\binom{n}{2} = \frac{n(n-1)}{2}$, so Theorem 2.1 provides us with a generating set of $\approx n^2/2$ transpositions. The next theorem shows we can get a generating set for S_n containing just $n-1$ transpositions.

Theorem 2.2. *For $n \geq 2$, S_n is generated by the $n-1$ transpositions*

$$(12), (13), \ldots, (1n).$$

Proof. The theorem is obvious for $n = 2$, so we take $n \geq 3$.

By Theorem 2.1, it suffices to write each transposition in S_n as a product of the transpositions involving a 1. For a transposition (ij) where i and j are not 1, check that

$$(ij) = (1i)(1j)(1i). \tag{2.1}$$

□

Here is a different generating set of $n-1$ transpositions.

Theorem 2.3. *For $n \geq 2$, S_n is generated by the $n-1$ transpositions*

$$(1\ 2), (2\ 3), \ldots, (n-1\ n).$$

Proof. By Theorem 2.1, it suffices to show each transposition (ab) in S_n is a product of transpositions of the form $(i\ i+1)$ where $i < n$.

Since $(ab) = (ba)$, without loss of generality $a < b$. We will argue by induction on $b-a$ that (ab) is a product of transpositions $(i\ i+1)$. This is obvious when $b - a = 1$, since $(ab) = (a\ a+1)$ is one of the transpositions we want in the desired generating set. Now suppose $b - a = k > 1$ and the theorem is settled for all transpositions moving a pair of integers whose difference is less than k.

Consider the formula

$$(a\ b) = (a\ a+1)(a+1\ b)(a\ a+1).$$

The first and third transpositions on the right side lie in our desired generating set. The middle transposition permutes a pair of integers with difference $b-(a+1) = k-1 < k$. By induction, $(a+1\ b)$ is a product of transpositions $(i\ i+1)$, so $(a\ b)$ is as well. □

Now we are ready to cut down the size of a generating set for S_n to *two*.

Lemma 2.4. *For a k-cycle $(i_1 i_2 \ldots i_k)$ in S_n and an arbitrary $\sigma \in S_n$,*

$$\sigma(i_1 i_2 \ldots i_k)\sigma^{-1} = (\sigma(i_1)\sigma(i_2)\ldots\sigma(i_k)). \tag{2.2}$$

Proof. To verify (2.2) in general, we will check both sides have the same effect on all integers from 1 to n. (Permutations are equal when they have the same values at the same numbers.)

For an integer from 1 to n of the form $\sigma(i_r)$, where $1 \leq r < k$, the left side of (2.2) has the effect $\sigma(i_r) \mapsto i_r \mapsto i_{r+1} \mapsto \sigma(i_{r+1})$, and the right side also sends $\sigma(i_r)$ to $\sigma(i_{r+1})$. The same reasoning shows both sides of (2.2) send $\sigma(i_k)$ to $\sigma(i_1)$.

For an integer ℓ from 1 to n that is *not* among $\{\sigma(i_1), \ldots, \sigma(i_k)\}$, it is not part of the cycle on the right side of (2.2), so the right side fixes ℓ. What does the left side of (2.2) do to ℓ? We want to compute

$$(\sigma(i_1 i_2 \ldots i_k)\sigma^{-1})(\ell).$$

The permutation σ^{-1} sends ℓ to $\sigma^{-1}(\ell)$. Since $\ell \notin \{\sigma(i_1), \ldots, \sigma(i_k)\}$, $\sigma^{-1}(\ell)$ is not among $\{i_1, \ldots, i_k\}$. Therefore the cycle $(i_1 i_2 \ldots i_k)$ fixes $\sigma^{-1}(\ell)$, and then σ sends $\sigma^{-1}(\ell)$ to ℓ. Thus the left side of (2.2) fixes ℓ, just like the right side. □

Theorem 2.5. *For $n \geq 2$, S_n is generated by the transposition* (12) *and the n-cycle* $(12\ldots n)$.

Proof. By Theorem 2.3, it suffices to show products of the permutations (12) and $(12\ldots n)$ yield all transpositions of the form $(i\ i+1)$. We may take $n \geq 3$.

Set $\sigma = (12\ldots n)$. Then

$$\sigma(12)\sigma^{-1} = (\sigma(1)\ \sigma(2)) = (23),$$

and more generally for $k = 1, 2, \ldots, n-2$,

$$\sigma^k(12)\sigma^{-k} = (\sigma^k(1)\ \sigma^k(2)) = (k+1\ k+2).$$

□

Corollary 2.6. *For $n \geq 3$, S_n is generated by* (12) *and the $(n-1)$-cycle* $(23\ldots n)$.

Proof. This is immediate from Theorem 2.5 since $(12\ldots n) = (12)(23\ldots n)$. □

We've found a pair of generators for S_n consisting of a transposition and n-cycle, and then a transposition and $(n-1)$-cycle. These have orders 2 and n, and then 2 and $n-1$. How small can the orders of a *pair* of generators of S_n be?

Theorem 2.7. *For $n \geq 3$ except for $n = 5, 6, 8$, S_n is generated by an element of order* 2 *and an element of order* 3.

We omit the proof of this theorem. It was first proved by G. A. Miller [4] in 1901, and his proof relied on a choice of a prime between n and $n/2$, so it was not explicit in terms of n. An explicit set of generators of order 2 and 3 was given by Dey and Wiegold [1] in 1971, unaware of Miller's earlier work. As an example, S_9 is generated by

$$(14)(28)(59) \text{ and } (123)(456)(789).$$

While S_n is generated by the *particular* transposition (12) and n-cycle $(12\ldots n)$, it is usually not true that S_n is generated by an *arbitrary* transposition and n-cycle. For instance, (13) and (1234) do not generate S_4. The reason is that these two permutations preserve congruences mod 2 (if two numbers in $1, 2, 3, 4$ are both even or both odd, applying either permutation to them returns values that are both even or both odd), so the subgroup they generate in S_4 has this property while S_4 does not have this property.

Theorem 2.8. *For $1 \leq a < b \leq n$, the transposition* (ab) *and n-cycle* $(12\ldots n)$ *generate S_n if and only if $(b-a, n) = 1$.*

Here the transposition (ab) is general, but the n-cycle is the standard one.

Proof. Let $d = (b-a, n)$. We will show every $g \in \langle (ab), (12\ldots n)\rangle$ preserves mod d congruences among $\{1, 2, \ldots, n\}$:

$$i \equiv j \bmod d \Longrightarrow g(i) \equiv g(j) \bmod d. \tag{2.3}$$

It suffices to check this when $g = (ab)$ and when $g = (12\ldots n)$. For $i \neq a$ or b, $(ab)(i) = i$. Also $(ab)(a) = b \equiv a \bmod d$ and $(ab)(b) = a \equiv b \bmod d$, so $(ab)(i) \equiv i \bmod d$ for all i. Thus (2.3) holds for $g = (ab)$. As for $g = (12\ldots n)$, we have $(12\ldots n)(i) \equiv i+1 \bmod n$, so also $(12\ldots n)(i) \equiv i+1 \bmod d$ since $d \mid n$. Therefore

$$i \equiv j \bmod d \Longrightarrow i+1 \equiv j+1 \bmod d,$$

so (2.3) holds for $g = (12\ldots n)$.

For $d > 1$, the group S_n does not preserve mod d congruences: pick $i \not\equiv j \bmod d$ and consider the transposition (ij). So if $\langle (ab), (12\ldots n)\rangle = S_n$ then we must have $d = 1$.

Now we prove the converse direction: if $(b-a, n) = 1$ then $\langle (ab), (12\ldots n)\rangle = S_n$. Let $\sigma = (12\ldots n)$, so $\sigma^i(a) \equiv a + i \bmod n$. Therefore $\sigma^{b-a}(a) \equiv b \bmod n$, and both sides of the congruence are between 1 and n, so $\sigma^{b-a}(a) = b$. Since $(b-a, n) = 1$, $\langle \sigma \rangle = \langle \sigma^{b-a}\rangle$ and σ^{b-a} is an n-cycle sending a to b, so $\sigma^{b-a} = (ab\ldots)$ where the dots are other numbers in the range $\{1, 2\ldots, n\}$. Then

$$\langle (ab), \sigma\rangle = \langle (ab), \sigma^{b-a}\rangle = \langle (ab), (ab\ldots)\rangle.$$

A suitable relabeling of the numbers $1, 2\ldots, n$ (that is, making an overall conjugation on S_n) turns (ab) into (12) and $(ab\ldots)$ into $(12\ldots n)$, so $\langle (ab), \sigma\rangle$ is conjugate to $\langle (12), (12\ldots n)\rangle$, which is S_n by Theorem 2.5. Therefore $\langle (ab), \sigma\rangle = S_n$. □

Corollary 2.9. *For each transposition $\tau = (ab)$ and n-cycle σ in S_n, $\langle \tau, \sigma\rangle = S_n$ if and only if $(k, n) = 1$, where $\sigma^k(a) = b$.*

Proof. Exercise. □

Corollary 2.10. *For a prime number p, S_p is generated by an arbitrary transposition and p-cycle.*

Proof. An arbitrary p-cycle in S_p can be written as $(12\ldots p)$ by relabeling the objects being permuted (that means by applying an overall conjugation on S_p), so to show an arbitrary transposition and p-cycle generate S_p it suffices to show each transposition and the specific p-cycle $(12\ldots p)$ generate S_p. For a transposition (ab) where $1 \le a < b \le p$, we have $(b-a, p) = 1$, so $\langle (ab), (12\ldots p)\rangle = S_p$ by Theorem 2.8. □

3. Generators for A_n

We now look for generating sets of A_n, where $n \ge 3$. (For $n = 1$ and 2, A_n is trivial.) Our development will parallel in large part what we did for S_n. In particular, we will see that A_n can be generated by two permutations.

Within S_n, every element of A_n is a product of transpositions, but the transpositions themselves do *not* lie in A_n. The smallest cycles in A_n (excluding the trivial 1-cycle) are the 3-cycles. Do these generate A_n? Yes.

Lemma 3.1. *For $n \ge 3$, each element of A_n is a product of 3-cycles. Therefore the 3-cycles generate A_n.*

Proof. The identity (1) is (123)(132), which is a product of 3-cycles. Now pick a non-identity element of A_n, say σ. Write it as a product of transpositions in S_n:

$$\sigma = \tau_1\tau_2\cdots\tau_r.$$

The left side has sign 1 and the right side has sign $(-1)^r$, so r is even. Therefore we can collect the products on the right into successive transpositions $\tau_i\tau_{i+1}$, where $i = 1, 3, \ldots$ is odd. We will now show every product of two transpositions in S_n is a product of two 3-cycles, so σ is a product of 3-cycles.

Case 1: τ_i and τ_{i+1} are equal. Then $\tau_i\tau_{i+1} = (1) = (123)(132)$, so we can replace $\tau_i\tau_{i+1}$ with a product of two 3-cycles.

Case 2: τ_i and τ_{i+1} have exactly one element in common. Let the common element be a, so we can write $\tau_i = (ab)$ and $\tau_{i+1} = (ac)$, where $b \ne c$. Then

$$\tau_i\tau_{i+1} = (ab)(ac) = (acb) = (abc)(abc),$$

so we can replace $\tau_i\tau_{i+1}$ with a product of two 3-cycles.

Case 3: τ_i and τ_{i+1} have no elements in common. This means τ_i and τ_{i+1} are disjoint, so we can write $\tau_i = (ab)$ and $\tau_{i+1} = (cd)$ where a, b, c, d are distinct (so $n \geq 4$). Then

$$\tau_i\tau_{i+1} = (ab)(cd) = (ab)(bc)(bc)(cd) = (bca)(cdb) = (abc)(bcd),$$

so we can replace $\tau_i\tau_{i+1}$ with a product of two 3-cycles. □

Let's reduce the number of 3-cycles needed.

Theorem 3.2. *For $n \geq 3$, the group A_n is generated by 3-cycles of the form $(1ij)$.*

Proof. For each 3-cycle (abc) not containing a 1, we have $(abc) = (1ab)(1bc)$. Now use Lemma 3.1. □

Theorem 3.2 is an analogue for alternating groups (on at least 3 letters) of Theorem 2.2 for symmetric groups (on at least 2 letters): each theorem says there is a generating set of cycles that contain 1 and have the smallest length consistent with the parity allowed in the group. The 3-cycles containing a 1 leave two undetermined terms, so our generating set for A_n in Theorem 3.2 is quite a bit larger than the generating set for S_n in Theorem 2.2. Here is a sharper analogue of Theorem 2.2 for A_n, where we only allow one undetermined entry in the 3-cycles.

Theorem 3.3. *For $n \geq 3$, the group A_n is generated by 3-cycles of the form $(12i)$.*

Proof. For $n = 3$, the only such 3-cycle is (123), and we know $A_3 = \{(1), (123), (132)\}$ is generated by (123). We now take $n \geq 4$.

Since $(12i)^{-1} = (1i2)$, each 3-cycle in A_n containing 1 and 2 is generated by the 3-cycles of the form $(12i)$. For a 3-cycle containing 1 but not 2, say $(1ij)$, check

$$(1ij) = (12j)(12j)(12i)(12j).$$

By Theorem 3.2, we're done. □

We can describe Theorem 3.3 in a "coordinate-free" manner as follows: A_n is generated by the 3-cycles moving a common pair of terms.

Here is an A_n-analogue of Theorem 2.3.

Theorem 3.4. *For $n \geq 3$, the consecutive 3-cycles $(i\ i+1\ i+2)$, with $1 \leq i \leq n-2$, generate A_n.*

Proof. This is true for $n = 3$ since $A_3 = \{(1), (123), (132)\}$ is cyclic with generator (123).

We know by Theorem 3.3 that A_4 is generated by (123) and (124). Since

$$(124) = (123)(123)(234)(123), \tag{3.1}$$

we see (123) and (234) also generate A_4.

Now take $n \geq 5$. Theorem 3.3 says a generating set for A_n is the set of 3-cycles $(12i)$. We argue by induction on i that these particular 3-cycles can be produced from products of consecutive 3-cycles. This is obvious for $i = 3$, and is shown in (3.1) for $i = 4$. For $i \geq 5$, assume $(12j)$ is a product of consecutive 3-cycles for $3 \leq j < i$. Then the equation

$$(1\ 2\ i) = (1\ 2\ i-2)(1\ 2\ i-1)(i-2\ i-1\ i)(1\ 2\ i-2)(1\ 2\ i-1)$$

and the inductive hypothesis show $(12i)$ is a product of consecutive 3-cycles. □

Here is an analogue of Theorem 2.5.

Theorem 3.5. *For $n \geq 4$, A_n is generated by two elements:*

$$(123) \text{ and } \begin{cases} (12\dots n), & \text{if } n \text{ is odd,} \\ (23\dots n), & \text{if } n \text{ is even.} \end{cases}$$

The theorem is true in a redundant sense when $n = 3$.

Proof. It suffices, by Theorem 3.4, to obtain all $(i\ i{+}1\ i{+}2)$ from the indicated permutations.

First suppose n is odd. Let $\sigma = (12\dots n)$, so $\sigma \in A_n$. Then, for $1 \leq k \leq n-3$.

$$\sigma^k(123)\sigma^{-k} = (\sigma^k(1)\sigma^k(2)\sigma^k(3)) = (k+1\ k+2\ k+3).$$

Now suppose n is even. Let $\sigma = (23\dots n)$, so $\sigma \in A_n$. Then, for $1 \leq k \leq n-3$,

$$\sigma^k(123)\sigma^{-k} = (\sigma^k(1)\sigma^k(2)\sigma^k(3)) = (1\ k+2\ k+3).$$

This does not give us the consecutive 3-cycles right away, but we can obtain them from what we now have, since

$$(k\ k+1\ k+2) = (1\ k\ k+1)(1\ k+1\ k+2).$$

□

There is an analogue of Theorem 2.7 for alternating groups, with a slightly different set of exceptions.

Theorem 3.6. *For $n \geq 3$ except for $n = 6, 7, 8$, A_n is generated by an element of order 2 and an element of order 3.*

We omit the proof. As an example, A_9 is generated by

$$(14)(29)(37)(56) \text{ and } (123)(456)(789).$$

4. $\mathrm{SL}_2(\mathbf{Z})$

The group $\mathrm{SL}_2(\mathbf{Z})$ consists of 2×2 integer matrices with determinant 1 (under multiplication). For instance, $\left(\begin{smallmatrix} 26 & 7 \\ 11 & 3 \end{smallmatrix}\right)$ is in $\mathrm{SL}_2(\mathbf{Z})$. There are two important matrices in this group:

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Since $S^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $S^4 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, the matrix S has order 4. Since $T^k = \left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)$ for all $k \in \mathbf{Z}$, so T has infinite order.

Theorem 4.1. *The group $\mathrm{SL}_2(\mathbf{Z})$ is generated by S and T.*

Proof. As the proof will reveal, this theorem is essentially the Euclidean algorithm in disguise. A worked example follows the proof.

First we check how S and a power of T change the entries in a matrix. Verify that

$$S\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix} \quad \text{and} \quad T^k\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+ck & b+dk \\ c & d \end{pmatrix}.$$

Thus, up to a sign change, multiplying by S on the left interchanges the rows. Multiplying by a power of T on the left adds a multiple of the second row to the first row and does not change the second row. Given a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z})$, we can carry out the Euclidean algorithm on a and c by using left multiplication by S and powers of T. We use the power of T to carry out the division (if $a = cq + r$, use $k = -q$) and use S to interchange the roles of a and c to guarantee that the larger of the two numbers (in absolute value) is in the upper-left corner. Multiplication by S will cause a sign change in the upper row, but this has no serious effect on the algorithm.

Since $ad - bc = 1$, a and c are relatively prime, so the last step of Euclid's algorithm will have a remainder of 1. This means, after suitable multiplication by S's and T's, we will have transformed the matrix $(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix})$ into one with first column $\binom{\pm 1}{0}$ or $\binom{0}{\pm 1}$. Left-multiplying by S interchanges the rows up to a sign, so we can suppose the first column is $\binom{\pm 1}{0}$. Each matrix of the form $(\begin{smallmatrix} 1 & x \\ 0 & y \end{smallmatrix})$ in $\mathrm{SL}_2(\mathbf{Z})$ must have $y = 1$ (the determinant is 1), and then it is $(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}) = T^x$. A matrix $(\begin{smallmatrix} -1 & x \\ 0 & y \end{smallmatrix})$ in $\mathrm{SL}_2(\mathbf{Z})$ must have $y = -1$, so the matrix is $(\begin{smallmatrix} -1 & x \\ 0 & -1 \end{smallmatrix}) = -T^{-x} = (\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix})T^x$. Since $(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}) = S^2$, we can finally unwind and express our original matrix in terms of S's and T's. □

Example 4.2. Take $A = (\begin{smallmatrix} 26 & 7 \\ 11 & 3 \end{smallmatrix})$. Since $26 = 11 \cdot 2 + 4$, we want to subtract $11 \cdot 2$ from 26:

$$T^{-2}A = \begin{pmatrix} 4 & 1 \\ 11 & 3 \end{pmatrix}.$$

Now we want to switch the roles of 4 and 11. Multiply by S:

$$ST^{-2}A = \begin{pmatrix} -11 & -3 \\ 4 & 1 \end{pmatrix}.$$

Divide -11 by 4: $-11 = 4 \cdot (-3) + 1$, so we want to add $4 \cdot 3$ to -11. Multiply by T^3:

$$T^3ST^{-2}A = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}.$$

Once again, multiply by S to switch the entries of the first column (up to sign):

$$ST^3ST^{-2}A = \begin{pmatrix} -4 & -1 \\ 1 & 0 \end{pmatrix}.$$

Our final division is: $-4 = 1(-4) + 0$. We want to add 4 to -4, so multiply by T^4:

$$T^4ST^3ST^{-2}A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = S.$$

Thus, left-multiplying by the inverses of all the S's and T's on the left side, we obtain

$$A = T^2S^{-1}T^{-3}S^{-1}T^{-4}S.$$

Since S has order 4, we can write S^{-1} as S^3 if we wish to use a positive exponent on S. However, a similar idea does not apply to the negative powers of T.

In this example, we wrote $(\begin{smallmatrix} 26 & 7 \\ 11 & 3 \end{smallmatrix})$ in terms of S and T, but not in the form S^aT^b or T^bS^a. The next theorem shows that is impossible.

Theorem 4.3. *Each matrix of the form S^aT^b or T^bS^a has 0 in one of its entries.*

Proof. Since S has order 4, each matrix S^aT^b or T^bS^a has the form $S^a(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix})$ and $(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix})S^a$ where $0 \leq a \leq 3$. Since $S^0 = I_2$ and $S^2 = -I_2$, this gives us 6 matrix products:

$$T^b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad ST^b = \begin{pmatrix} 0 & -1 \\ 1 & b \end{pmatrix}, \quad S^2T^b = T^bS^2 = \begin{pmatrix} -1 & -b \\ 0 & -1 \end{pmatrix},$$

$$S^3T^b = \begin{pmatrix} 0 & 1 \\ -1 & -b \end{pmatrix}, \quad T^bS = \begin{pmatrix} b & -1 \\ 1 & 0 \end{pmatrix}, \quad T^bS^3 = \begin{pmatrix} -b & 1 \\ -1 & 0 \end{pmatrix}.$$

In all cases, the product has 0 in one of its entries. □

Since $\mathrm{SL}_2(\mathbf{Z}) = \langle S, T\rangle$, we can also say $\mathrm{SL}_2(\mathbf{Z}) = \langle S, ST\rangle$. This is interesting because S has order 4 and $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ has order 6. In other words, $\mathrm{SL}_2(\mathbf{Z})$ is an infinite group that can be generated by two elements of finite order. Moreover, in the quotient group[1] $\mathrm{PSL}_2(\mathbf{Z}) := \mathrm{SL}_2(\mathbf{Z})/\{\pm I_2\} = \langle \overline{S}, \overline{ST}\rangle$, $\overline{S}$ has order 2 and $\overline{ST}$ has order 3. This quotient group is "universal" for the property of being generated by elements of order dividing 2 and 3: if $G = \langle x, y\rangle$ is a group generated by elements x and y where $x^2 = e$ and $y^3 = e$, there is a unique group homomorphism $\mathrm{PSL}_2(\mathbf{Z}) \to G$ sending $\overline{S}$ to x and $\overline{ST}$ to y, so G is isomorphic to a quotient group of $\mathrm{PSL}_2(\mathbf{Z})$. For example, A_n and S_n are generated by elements of order 2 and 3 when $n \geq 9$, so these groups are quotient groups of $\mathrm{PSL}_2(\mathbf{Z})$.

5. $\mathrm{GL}_n(\mathbf{R})$ and $\mathrm{SL}_n(\mathbf{R})$

In linear algebra, matrices get into row echelon form by elementary row operations:

- Multiply a row by a nonzero scalar.
- Add a multiple of one row to another.
- Exchange two rows.

Elementary row operations on an $m \times n$ matrix can be expressed using left multiplication by an $m \times m$ matrix having a particularly simple form: just one entry might not be 0 or 1. We will illustrate this on a 2×3 matrix $\left(\begin{smallmatrix} a & b & c \\ d & e & f \end{smallmatrix}\right)$.

(1) Multiplying a row by λ is left multiplication by $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right)$, for instance

$$\begin{pmatrix} \lambda a & \lambda b & \lambda c \\ d & e & f \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

(2) Adding λ times one row to another row is left multiplication by $\left(\begin{smallmatrix} 1 & \lambda \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & 0 \\ \lambda & 1 \end{smallmatrix}\right)$, for instance

$$\begin{pmatrix} a + \lambda d & b + \lambda e & c + \lambda f \\ d & e & f \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

(3) Exchanging rows is left multiplication by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$:

$$\begin{pmatrix} d & e & f \\ a & b & c \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

An *elementary matrix* is one that corresponds to an elementary row operation. They come in three flavors:

- $D_i(\lambda)$ has 1's on the main diagonal except for a $\lambda \neq 0, 1$ in the ith slot along the main diagonal.
- $I_n + \lambda E_{ij}$ with $i \neq j$ is a matrix with 1's on the main diagonal and a nonzero λ in position (i, j).
- R_{ij} with $i \neq j$ has 1's in positions (i, j) and (j, i) and also in every diagonal position except the ith and jth, and 0's everywhere else. Note $R_{ij} = R_{ji}$. (This is a matrix analogue of the transposition (ij).)

The inverse of such a matrix is of the same type: $D_i(\lambda)^{-1} = D_i(1/\lambda)$, $(I_n + \lambda E_{ij})^{-1} = I_n - \lambda E_{ij}$, and $R_{ij}^{-1} = R_{ij}$.

Theorem 5.1. *The group* $\mathrm{GL}_n(\mathbf{R})$ *is generated by elementary matrices.*

[1] The group $\{\pm I_2\}$ is the center of $\mathrm{SL}_2(\mathbf{Z})$.

Proof. Pick $A \in \mathrm{GL}_n(\mathbf{R})$. Its reduced row echelon form is I_n, the only $n \times n$ invertible reduced row echelon matrix. Elementary row operations correspond to left multiplication by an elementary matrix, so there are elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_1 E_2 \ldots E_k A = I_n.$$

Therefore $A = E_k^{-1} \cdots E_2^{-1} E_1^{-1}$. The inverse of an elementary matrix is an elementary matrix, so A is a product of elementary matrices. □

Is the subgroup $\mathrm{SL}_n(\mathbf{R})$ generated by the elementary matrices $I_n + \lambda E_{ij}$? They have determinant 1 and the other elementary matrices do not: $\det(D_i(\lambda)) = \lambda$ and $\det(R_{ij}) = -1$. The proof of Theorem 5.1 does not work in $\mathrm{SL}_n(\mathbf{R})$ since row reduction can involve row exchanges or scaling a row to make an entry 1, and these correspond to the two types of elementary matrices not in $\mathrm{SL}_n(\mathbf{R})$. Nevertheless, it turns out that the guess is correct:

Theorem 5.2. *For $n \geq 2$, the matrices $I_n + \lambda E_{ij}$ with $i \neq j$ and $\lambda \in F^\times$ generate $\mathrm{SL}_n(\mathbf{R})$.*

Proof. This involves tedious computations. First we treat the case $n = 2$. Pick $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{R})$. To show it is a product of matrices $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)$, first suppose $b \neq 0$. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (d-1)/b & 1 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ (a-1)/b & 1 \end{pmatrix}.$$

If $c \neq 0$ then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & (a-1)/c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & (d-1)/c \\ 0 & 1 \end{pmatrix}.$$

If $b = 0$ and $c = 0$ then the matrix is $\left(\begin{smallmatrix} a & 0 \\ 0 & 1/a \end{smallmatrix}\right)$, and

$$\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (1-a)/a & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/a \\ 0 & 1 \end{pmatrix}.$$

To handle $n > 2$, we induct on n. Pick $A \in \mathrm{SL}_n(\mathbf{R})$. We will show multiplying A on the left *or right* by matrices $I_n + \lambda E_{ij}$ leads to a block matrix $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A' \end{smallmatrix}\right)$. Since this is in $\mathrm{SL}_n(\mathbf{R})$ we have $\det A' = 1$, so $A' \in \mathrm{SL}_{n-1}(\mathbf{R})$. By induction A' is a product of elementary matrices $I_{n-1} + \lambda E_{ij}$, so $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A' \end{smallmatrix}\right)$ is a product of block matrices $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & I_{n-1}+\lambda E_{ij} \end{smallmatrix}\right) = I_n + \lambda E_{i+1\ j+1}$. Thus

$$(\text{product of some } I_n + \lambda E_{ij}) A (\text{product of some } I_n + \lambda E_{ij}) = \text{product of some } I_n + \lambda E_{ij},$$

and we can solve for A to see that it is a product of matrices $I_n + \lambda E_{ij}$.

Multiplying A by $I_n + \lambda E_{ij}$ on the left or right is an elementary row or column operation:

$$(I_n + \lambda E_{ij})A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} + \lambda a_{j1} & \cdots & a_{in} + \lambda a_{jn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

where the ith row equals ith row of $A + \lambda(j$th row of A, and

$$A(I_n + \lambda E_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1j} + \lambda a_{1i} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} + \lambda a_{ni} & \cdots & a_{nn} \end{pmatrix},$$

where the jth column equals the jth column of $A + \lambda(i$th column of $A)$.

Looking along the first column of A, some entry is not 0 (since $\det A = 1$). If $a_{k1} \neq 0$ where $k > 1$ then

$$\left(I_n + \frac{1-a_{11}}{a_{k1}}E_{1k}\right) A = \begin{pmatrix} 1 & \cdots \\ \vdots & \ddots \end{pmatrix}. \tag{5.1}$$

If $a_{21}, \dots, a_{n1}$ are all 0, then $a_{11} \neq 0$ and

$$\left(I_n + \frac{1}{a_{11}}E_{21}\right) A = \begin{pmatrix} a_{11} & \cdots \\ 1 & \cdots \\ \vdots & \ddots \end{pmatrix}.$$

Then by (5.1) with $k = 2$,

$$(I_n + (1-a_{11})E_{12})\left(I_n + \frac{1}{a_{11}}E_{21}\right) A = \begin{pmatrix} 1 & \cdots \\ \vdots & \ddots \end{pmatrix},$$

Once we have a matrix with upper left entry 1, multiplying it on the left by $I_n + \lambda E_{i1}$ for $i \neq 1$ will add λ to the $(i,1)$-entry, so with a suitable λ we can make the $(i,1)$-entry of the matrix 0. Thus multiplication on the left by suitable matrices of the form $I_n + \lambda E_{ij}$ produces a block matrix $\left(\begin{smallmatrix} 1 & * \\ \mathbf{0} & B \end{smallmatrix}\right)$ whose first column is all 0's except for the upper left entry, which is 1. Multiplying this matrix on the right by $I_n + \lambda E_{1j}$ for $j \neq 1$ adds λ to the $(1,j)$-entry without changing a column other than the jth column. With a suitable choice of λ we can make the $(1,j)$-entry equal to 0, and carrying this out for $j = 2, \dots, n$ leads to a block matrix $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A' \end{smallmatrix}\right)$, which is what we need to conclude the proof by induction. □

6. Minimal Generating Sets

Since a basis of $\mathbf{R}^n$ is the same thing as a minimal spanning set, perhaps the analogue of a basis of $\mathbf{R}^n$ in the context of groups is a *minimal generating set*: a generating set that is no longer a generating set when some element is removed. In contrast to bases for $\mathbf{R}^n$, different minimal generating sets of a nonabelian group may not have the same size.

Example 6.1. Two minimal generating sets for S_4 are $\{(12), (23), (34)\}$ and $\{(12), (1234)\}$. That these sets each generate S_4 is a special case of Theorems 2.3 and 2.5.

Why is the first generating set minimal? Any two transpositions in it have a common fixed point or commute. These properties are preserved under multiplication, and S_4 has neither property, so two permutations in the first set do not generate S_4. In the second set, removing a permutation leaves just one, which does not generate S_4 since S_4 is not cyclic.

Example 6.2. The generating sets in Theorems 2.2 and 2.3 are both minimal. Why? Well, removing a transposition $(1i)$ from the generating set in Theorem 2.2 will leave a set of transpositions with a common fixed point i, and thus they can't generate all of S_n. To explain why the generating set in Theorem 2.3 is minimal is a bit more subtle. It we take out (12) or $(n-1\ n)$, then at most we will be able to generate permutations having a common fixed point (1 or n), and thus we do not get all of S_n. If we take out $(i\ i+1)$ for some i from 2 to $n-1$, then we are left with

$$(12), (23), \dots, (i-1\ i), (i+1\ i+2), \dots, (n-1\ n).$$

Convince yourself that these transpositions can never produce the transposition that swaps 1 and n. The "bridge" linking 1 to n through the generating set has been broken when $(i\ i+1)$ is taken out.

Theorem 6.3. *A set of transpositions that generates S_n must have at least $n-1$ transpositions, so a set of $n-1$ transpositions in S_n that generates S_n is a minimal generating set of S_n.*

Proof. Let $\tau_1, \dots, \tau_m$ be distinct transpositions in S_n, and assume they generate S_n. We want to show $m \geq n-1$. We will use graph theory.

Create a graph whose vertices are $\{1, 2, \dots, n\}$ and make an edge between i and j if (ij) is one of the transpositions $\tau_1, \dots, \tau_m$. So our graph has n vertices and m edges. For two vertices i and j in the graph, to say there is an element of the group $\langle \tau_1, \dots, \tau_m \rangle$ that sends i to j is equivalent to saying there is a path in our graph starting at i and ending at j. Therefore when $\langle \tau_1, \dots, \tau_m \rangle = S_n$, this graph is connected. A connected graph with n vertices must have at least $n-1$ edges, so $m \geq n-1$. □

Remark 6.4. Your prior experience with bases in linear algebra might suggest that 2-element generating sets for S_n, being as small as possible (for $n \geq 3$), should be the most useful generating sets. In fact, the generating set of transpositions $(i\ i+1)$ in Theorem 2.3 turns out to be the most important one.

Because minimal generating sets in a non-abelian group need not have a common size, there is not a theory of "bases" for finitely generated groups that permits the same degree of elegance as bases for finite-dimensional vector spaces in linear algebra. In fact, finitely generated groups can fail to satisfy natural analogues of theorems about finite-dimensional vector spaces:

- a subgroup of a finitely generated group need not be finitely generated, (!)
- even if a subgroup of a finitely generated group is finitely generated, it might require more generators than the original group. (!)

For a finite group G, let $d(G)$ denote the smallest possible size of a generating set for G. (With analogies to linear algebra in mind, we might naively think of $d(G)$ as a "dimension," hence the notation.) For instance, $d(G) = 1$ precisely when G is a cyclic group. Theorems 2.5 and 3.5 tell us $d(S_n)$ and $d(A_n)$ equal 2 when these groups are non-abelian (that is, except for very small n). Consider the following questions about $d(G)$ for a general finite group G, motivated by an analogy with bases in $\mathbf{R}^n$:

- Do all minimal generating sets for G have $d(G)$ terms?
- Does every generating set for G contain a generating set of $d(G)$ elements?
- If $H \subset G$ is a subgroup, is $d(H) \leq d(G)$?

The answer to these questions, in the context of all finite groups, is *no*. Indeed, the symmetric groups S_n for $n \geq 4$ answer the first two questions in the negative, using the minimal generating set $(12), (13), \dots, (1n)$ from Theorem 2.2. The symmetric groups also provide examples that negate the third question, as follows. Let H be the subgroup of S_n generated by the transpositions $(12), (34), \dots, (2j-1\ 2j), \dots$. While $d(S_n) = 2$, it can be shown that $d(H) = [n/2]$. In particular, $d(H) > d(S_n)$ for $n \geq 6$.

Despite this bad news in general, if we restrict our attention to finite groups G of prime-power order then the answers to the first two questions above are *yes*: all minimal generating sets of G have the same size and every generating set of G contains a generating set with that minimal size. The main result leading to this is called the Burnside Basis Theorem, which involves the Frattini subgroup.

Definition 6.5. For a nontrivial finite group G, its *Frattini subgroup* $\Phi(G)$ is the intersection of the maximal ($\Rightarrow$ proper) subgroups of G:

$$\Phi(G) = \bigcap_{\text{max.subgp.}} H$$

If G is trivial, define $\Phi(G)$ to be trivial too.

Always $\Phi(G) \lhd G$, since conjugates of maximal subgroups are also maximal subgroups.

When G is a finite p-group, its maximal subgroups are its subgroups of index p and they are all normal subgroups. For a maximal subgroup M, G/M has order p, so $[G, G] \subset M$ from G/M being abelian. Thus $[G, G] \subset \Phi(G)$, so $G/\Phi(G)$ is an abelian p-group. Since p-th powers in each G/M are trivial, we have $\{g^p : g \in G\} \subset \Phi(G)$, so all nontrivial elements of $G/\Phi(G)$ have order p. Thus $G/\Phi(G) \cong (\mathbf{Z}/(p))^r$ for some $r \geq 1$. We can think of r as a dimension: $G/\Phi(G)$ is a vector space over $\mathbf{Z}/(p)$ (scale by exponentiation: $(a \bmod p)g = g^a$) and $r = \dim_{\mathbf{Z}/(p)}(G/\Phi(G))$.

Example 6.6. If $G = \langle g \rangle$ is nontrivial cyclic of p-power order then it has only one maximal proper subgroup, $\langle g^p \rangle$, so $\Phi(G) = \langle g^p \rangle$ and $G/\Phi(G) = \langle g \rangle / \langle g^p \rangle$ is cyclic of order p.

Example 6.7. The quaternion group Q_8 has maximal subgroups $\langle i \rangle$, $\langle j \rangle$, $\langle k \rangle$, which intersect in $\{\pm 1\}$, so $\Phi(Q_8) = \{\pm 1\}$ and $Q_8/\Phi(Q_8) = \langle \overline{i}, \overline{j} \rangle \cong (\mathbf{Z}/(2))^2$.

Example 6.8. If

$$G = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in \mathbf{Z}/(p) \right\}$$

for a prime p then

$$\Phi(G) = \left\{ \begin{pmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : b \in \mathbf{Z}/(p) \right\}$$

and $G/\Phi(G) \cong (\mathbf{Z}/(p))^2$.

Theorem 6.9 (Burnside)**.** *Let G be a finite p-group. For each subset $\{g_1, \dots, g_k\}$ of G,*

$$\langle g_1, \dots, g_k \rangle = G \iff \langle \overline{g}_1, \dots, \overline{g}_k \rangle = G/\Phi(G).$$

A proof of Theorem 6.9 can be found in specialized texts on group theory. The theorem reduces many questions about generating sets of a finite p-group G to questions about spanning sets of a finite-dimensional vector space. In particular, all minimal generating sets of G have common size equal to the size of a basis of $G/\Phi(G)$ as a vector space over $\mathbf{Z}/(p)$ and every generating set for G must contain a generating set of $d(G)$ elements because every spanning set of a vector space of dimension d has at least d elements.

The implication $H \subset G \Rightarrow d(H) \leq d(G)$ has counterexamples among finite p-groups. (There wouldn't be counterexamples if we always had $H \cap \Phi(G) = \Phi(H)$, but such a formula isn't even true if H is a nontrivial subgroup of a cyclic p-group.[2]) Before describing a family of counterexamples, we describe a situation where the implication is valid for p-groups.

Let G^n be the subgroup of a finite group G generated by its nth powers (this may include elements that are not nth powers, just as the commutator subgroup is generated by commutators but its elements may not all be commutators). Then $G^n \lhd G$, and a finite

[2] If $G = \langle g \rangle$, set $H = \Phi(G) = \langle g^p \rangle$. Then $\Phi(H) = \langle g^{p^2} \rangle$, so $H \cap \Phi(G) = \Phi(G) \neq \Phi(H)$ if $|G| > p$.

p-group is called *powerful* if G/G^p is abelian for $p \neq 2$ or G/G^4 is abelian for $p = 2$. All abelian p-groups are powerful, and powerful p-groups share some features of abelian groups, *e.g.*, if G is a powerful p-group then $G^p = \{g^p : g \in G\}$ [2, Prop. 2.6]. It turns out that for powerful p-groups G, $H \subset G \Rightarrow d(H) \leq d(G)$ [2, Theorem 2.9].

Here is a construction of a finite p-group G where $d(H) > d(G)$ for a subgroup H.

Example 6.10. For a prime p, let A be the $p \times p$ matrix over $\mathbf{Z}/(p)$ that permutes the standard basis cyclically: $A(e_1) = e_2, A(e_2) = e_3, \ldots, A(e_p) = e_1$ and A is extended to $(\mathbf{Z}/(p))^p$ linearly:

$$A = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}.$$

This matrix has order p. Set $V := (\mathbf{Z}/(p))^p$ and $G := V \rtimes \langle A \rangle$ (a semi-direct product), with powers of A acting on V by matrix-vector multiplication. The group law is

$$(v, A^i)(w, A^j) = (v + A^i w, A^{i+j}).$$

Then $|G| = p^{p+1}$.

Concretely, we can think of G as certain affine transformations of V: $(v, A^i) \in G$ is the mapping $V \to V$ where $\mathbf{x} \mapsto A^i\mathbf{x} + v$. For $v \in V$, the product vA in G is the mapping $\mathbf{x} \mapsto A\mathbf{x} + v$ while the product Av in G is the mapping $\mathbf{x} \mapsto A(\mathbf{x} + v) = A\mathbf{x} + Av$. Therefore A and v don't commute in G if $Av \neq v$, so A does not commute in G with members of the standard basis of V.

Since powers of A send e_1 to the entire standard basis of V and repeated addition sends each e_i to its $\mathbf{Z}/(p)$-multiples, from A and e_1 we can get all of V and thus A and e_1 generate the group G. Neither A nor e_1 generates G by itself (G is not abelian), so $d(G) = 2$. The subgroup $H := V$ of G is a p-dimensional vector space over $\mathbf{Z}/(p)$, so $d(H) = p$. Thus when $p \neq 2$ we have constructed a p-group G and a subgroup H where $d(H) > d(G)$.

An example where $d(G) > d(H)$ for 2-groups can also be constructed, but this is omitted here.

References

[1] I. M. S. Dey and J. Wiegold, "Generators for alternating and symmetric groups," *J. Australian Math. Soc.* **12** (1971), 63–68.

[2] J. D. Dixon, M. P. F. du Sautoy, A. Mann, D. Segal, *Analytic Pro-p Groups*, 2nd ed., Cambridge Univ. Press, Cambridge, 1999.

[3] A. Lubotzky, A. Mann, "Powerful p-groups I. Finite Groups," *J. Algebra* **105** (1987), 484–505.

[4] G. A. Miller, "On the groups generated by 2 operators," *Bull. Amer. Math. Soc.* **7** (1901), 14–32.