Generating Sequences of Finite Groups

Dan Collins
Cornell University Class of 2010

Advised by R. Keith Dennis
Department of Mathematics, Cornell University

December 3, 2009

Introduction

Generators and relations have always been an important tool for studying groups, and if we are given a group G it is useful to study generating sets of G. If G is a finite group (or even a finitely generated group), we let a "generating sequence" be any finite n-tuple $(g_1, \ldots, g_n)$ so that $\langle g_1, \ldots, g_n \rangle = G$. Since generating sequences are ordered, and allow elements to occur multiple times, it is often useful to work with them as opposed to finite generating sets. This thesis studies generating sequences as well as some group-theoretic ideas that use generating sequences in a central way.

The main motivation for studying this topic comes from an algorithm in computational group theory, the *Product Replacement Algorithm*. This algorithm generates a random element of a given finite group G, in the following manner. We start with a generating sequence $(g_1, \ldots, g_n)$ of G. Then, we randomly pick a coordinate g_i, and replace it by a product $g_j g_i$, $g_j^{-1} g_i$, $g_i g_j$, or $g_i g_j^{-1}$ (with $j \neq i$ chosen at random, and the product in question chosen at random). This results in another generating sequence; we can then repeat this random "product replacement" a fixed number of times. This leaves us with a generating sequence $(g_1', \ldots, g_n')$, and the algorithm returns g_1' as our "randomly-generated element".

This algorithm was first introduced in the early 1990's. It seemed to successfully generate uniformly-distributed randomly elements, and worked faster than other techniques for randomly generating elements (see [CLGM$^+$95]). Some progress was made towards a theoretical understanding of this success. A landmark paper of Diaconis and Saloff-Coste [DSC98] gave bounds on the number of product replacements necessary for the algorithm to work well. Pak's paper "What do we know about the product replacement algorithm?" [Pak99] summarizes some of the initial progress made towards understanding this algorithm, as well as the important questions that still remain.

One of the most important questions towards understanding the product replacement algorithm is "Given a group G and an integer n, can a series of product replacements connect any two length n generating sequences?". This is a question in pure group theory, and it had been studied previously in the group theory literature (under the name of "T-systems"). However, there is much still to do, and a number of other interesting related questions. This thesis discusses some of these. I have also tried to collect many of the important results related to these questions, with full proofs.

Sections 1 and 2 discusses some general theoretical questions about generating sequences. We mention various techniques for determining how many generating sequences there are of a fixed length n for a fixed group G, and for working with the set of these sequences. Some of the most important results are due to a paper of Gaschütz [Gas55], which deal with generating sequences in quotient groups and direct products.

In Section 3, we develop some of the theory to study when we can connect two length n generating sequences by product replacements. Our approach to this is to construct a group of "elementary operations," generated by the product replacement operations. This group then acts on the set of all length n generating sequences, and two sequences can be connected by product replacements if and only if they are in the same orbit. So, any pair of sequences can be connected with each other if and only if the action is transitive. A result of Dunwoody [Dun70] shows that if G is solvable and n is longer than the shortest possible length of a generating sequence of G, then the action is transitive. Another result of Diaconis and Graham [DG99] shows that if A is abelian and n equals the shortest possible length, then the orbits of the action are parametrized by a certain "determinant function." This result suggests a method for defining an algebraic K-theory for finite groups.

Section 4 discusses "homogeneous groups," which satisfy a certain uniformity property related to generating sequences. In particular, we are interested in the "homogeneous cover" of a group G, which is a homogeneous group with a distinguished generating sequence that can map to any generating sequence of G. The concept of a homogeneous group was introduced by Gaschütz in [Gas55]. Most of the results in this section are due to Keith Dennis and some of his colleagues, in particular Ken Brown, Steve Chase, and Laurent Saloff-Coste.

Finally, Section 5 discusses some of the ideas and computations that I have worked on with Keith Dennis as an undergraduate research project. They are mainly based on trying to extend the previously-mentioned result of Diaconis and Graham, to understand the action of "elementary operations" on the set of all generating sequences of a group G of the shortest possible length. One

approach we worked on involved constructing more general determinant functions. We also made many computations, in particular for when G is a p-group, to better understand the general behavior of the elementary operations.

The results discussed in this thesis come from many sources: some are well-known or part of the "folklore," others have to my knowledge not previously appeared anywhere, and a few are new from this project. I have tried to give proper credit where possible. Many of the results, especially those in the first three sections, had been previously collected by my advisor Keith Dennis in an unpublished paper [Den09].

The necessary background for this thesis is an introduction to group theory, such as the first six chapters of *Abstract Algebra* by Dummit and Foote [DF04]. Beyond this background, this thesis should be self-contained; I have given full proofs of most of the results we need (though a few results are cited that are far beyond our scope).

Finally, we mention some possibly nonstandard notation and conventions that we use. We use $f[X]$ and $f^{-1}[Y]$ to denote an image and a preimage under a function f. We define the commutator $[g_1, g_2]$ as $g_1^{-1}g_2^{-1}g_1g_2$. We use Z_n to denote a cyclic group of order n, and write this group multiplicatively (if we want to write a cyclic group additively, we denote it $\mathbb{Z}/n\mathbb{Z}$).

1 Generating Sequences

The fundamental objects we're interested in are generating sequences of finite groups:

Definition 1.1. A *generating sequence* (of length n) of a finite group G is a finite sequence $(g_1, \ldots, g_n)$ of elements of G that generate G.

By definition of the subgroup generated by a set, a sequence $(g_1, \ldots, g_n)$ generates G if and only if every element of G can be written as a product of the elements g_i and their inverses.

Definition 1.2. We let $\Gamma_n(G)$ denote the set of length n generating sequences; it is a subset of G^n, the set of all length n sequences in G.

Definition 1.3. We let $\varphi_n(G)$ denote the number of length n generating sequences (i.e. $\varphi_n(G) = |\Gamma_n(G)|$). Following [Hal36], we call this the *n-th Eulerian function.*

Definition 1.4. We define $r(G)$ as the smallest integer n so that G has a generating sequence of length n.

We remark that there does not seem to be a standard notation for this quantity, and that various symbols have been used for it in the literature (most commonly $d(G)$ or $m(G)$). Note that if $n < r(G)$, $\Gamma_n(G)$ is empty by definition. If $n = r(G)$, then $\Gamma_n(G)$ is nonempty, as we can take a length $r(G)$ generating sequence and append any sequence of $n - r(G)$ elements to the end.

Example 1.5. By definition, we have $r(G) = 1$ if and only if G is generated by a single element, if and only if G is a cyclic group. If we let $G = Z_p$ be a cyclic group of prime order, any non-identity element of Z_p generates it. Therefore, any sequence of n elements in Z_p other than $(1, \ldots, 1)$ is a generating sequence. This means that

$$\Gamma_n(Z_p) = Z_p^n \setminus \{(1, \ldots, 1)\},$$

and in particular

$$\varphi_n(Z_p) = p^n - 1.$$

We can extend this computation to an arbitrary finite cyclic group Z_m. In particular, if m has prime decomposition $p_1^{a_1} \cdots p_k^{a_k}$ (where the p_i are distinct primes and the a_i are positive), we have

$$\varphi_n(Z_m) = p_1^{n(a_1-1)}(p_1^n - 1) \cdots p_k^{n(a_k-1)}(p_k^n - 1). \tag{1}$$

It is possible to prove this directly. Alternatively, it will follow easily from the computation of $\varphi_n(Z_p)$ given some theoretical results we will prove in Sections 1.3 and 2.1; see Corollary 2.10.

While it is easy to understand groups with $r(G) = 1$, there is no easy description of groups satisfying $r(G) = n$ for any fixed $n > 1$. Even groups with $r(G) = 2$ can be surprisingly complicated. For instance, all nonabelian finite simple groups satisfy $r(S) = 2$ (the proof of this uses the Classification Theorem of Finite Simple Groups). Also, if we let A_5 denote the alternating group on 5 symbols, and A_5^n denote the direct product of n copies of A_5, we have $r(A_5^{19}) = 2$ but $r(A_5^{20}) = 3$. This is a consequence of Theorem 2.22 proved in Section 2.3.

1.1 Basic Techniques

The Eulerian function of a group was introduced by Philip Hall in [Hal36]. In this paper, Hall also gave a method for computing $\varphi_n(G)$. He introduced the "Möbius function of a finite group," which is related to the classical Möbius function from number theory. In fact, both of these concepts can be generalized to define a Möbius function for a finite partially ordered set (see for instance Chapter 25 of [vLW03]).

Definition 1.6. Given a finite group G, define the *Möbius function* μ_G as an integer-valued function on the set of all subgroups of G. In particular, we define $\mu_G(G) = 1$, and for $H < G$ we define recursively

$$\mu_G(H) = -\sum_{K:H<K\leq G} \mu_G(K)$$

We can restate this by saying that if $H < G$, we have

$$\sum_{H\leq K\leq G} \mu_G(K) = 0.$$

This allows us to prove a *Möbius inversion theorem*, analogous to classical Möbius inversion:

Theorem 1.7. *Let $f(H)$ be a function on the subgroups of G, and let $F(H)$ be the summation function defined by $F(H) = \sum_{K:K\leq H} f(K)$. Then, we have*

$$f(G) = \sum_{H\leq G} \mu_G(H)F(H).$$

Proof. Expanding out the sum defining $F(H)$, we have

$$\sum_{H\leq G} \mu_G(H)F(H) = \sum_{H\leq G} \sum_{1\leq K\leq H} \mu_G(H)f(K).$$

We can then switch the order of the sums, giving

$$\sum_{H\leq G} \mu_G(H)F(H) = \sum_{K\leq G} \left(\sum_{K\leq H\leq G} \mu_G(H) \right) f(K).$$

By definition of the Möbius function, this reduces to $\mu_G(G) \cdot f(G) = f(G)$. □

To apply this to our situation, let $f(H) = \varphi_n(H) = |\Gamma_n(H)|$. Since there are $|G|^n$ length n sequences in G, and each one generates a subgroup $H \leq G$,

$$\sum_{1\leq H\leq G} \varphi_n(G) = \sum_{1\leq H\leq G} |\Gamma_n(H)| = |G|^n.$$

Applying Möbius inversion gives Hall's formula for $\varphi_n(G)$:

Corollary 1.8.

$$\varphi_n(G) = \sum_{H\leq G} \mu_G(H)|H|^n.$$

This is an effective method for computing $\varphi_n(G)$. However, it doesn't help much with a theoretical understanding of $\varphi_n(G)$ - we can't even tell whether $\varphi_n(G)$ is zero or not without going through the full computation! Ideally, we would like a more explicit formula for $\varphi_n(G)$, such as the one we stated above for $\varphi_n(Z_m)$. We would also like if the formula was associated to a description of the generating sequences $\Gamma_n(G)$.

We could hope to find such a formula that is valid for a single group and some range of n, or for some particular class of G and a single value of n. One such example where this is possible is for two-element generating sequences of dihedral groups:

Example 1.9. Let D_{2n} denote the dihedral group of order $2n$. We know D_{2n} is generated by two elements R ("rotation") and F ("flip") that satisfy $R^n = F^2 = 1$ and $RF = FR^{-1}$. Since D_{2n} is not cyclic, this means $r(D_{2n}) = 2$.

We claim that $\varphi_2(D_{2n}) = 3n\varphi(n)$, where $\varphi(n)$ is Euler's phi function. In particular, we claim that the elements of $\Gamma_2(D_{2n})$ fall into three classes:

1. Sequences (R^iF, R^j) with $1 \le i, j \le n$ and $(j, n) = 1$.
2. Sequences (R^i, R^jF) with $1 \le i, j \le n$ and $(i, n) = 1$.
3. Sequences (R^iF, R^jF) with $1 \le i, j \le n$ and $(j - i, n) = 1$.

By definition, there there are $\varphi(n)$ integers between 1 and n that are coprime to n, so there are $n\varphi(n)$ sequences in each of these three classes, and thus $\varphi_2(D_{2n}) = 3n\varphi(n)$.

Note that any element of D_{2n} is of the form R^iF^j for $0 \le i < n$ and $0 \le j < 2$. Therefore, any generating sequence is of the form (R^iF^j, R^kF^ℓ), and we can't have $j = \ell = 0$ because then the sequence would generate a subgroup of $\langle R\rangle$. So, it remains to check that the other three cases for k, ℓ correspond to the classes (1), (2), (3) listed above.

First, consider a sequence of the form (R^iF, R^j). If $(j, n) = 1$ then R^j generates $\langle R\rangle$, so in particular generates R. Then it generates R^{-i} and hence $F = R^{-i}R^iF$, so the sequence generates D_{2n}. Conversely, if (R^iF, R^j) generates, note that $(R^iF)^{-1} = R^iF$ and $(R^iF)R^j = R^iR^{-j}F = (R^j)^{-1}(R^iF)$. Therefore, any product of R^iF, R^j, and their inverses can be written in the form $(R^j)^k(FR^i)^\ell$ for $0 \le k < n$ and $0 \le \ell < 2$, and moreover this element is in the coset $F^\ell\langle R\rangle$ of the subgroup $\langle R\rangle$. So, if $\langle R^iF, R^j\rangle = D_{2n}$, we can write $R = (R^j)^k(FR^i)^\ell$. Since $R \in \langle R\rangle$, we must have that $\ell = 0$, so $\langle R^j\rangle$ generates the cyclic group $\langle R\rangle$. This means $(j, n) = 1$, as desired. So, the sequences of the form (R^iF, R^j) that generate are exactly those in class (1).

An identical argument shows that the sequences of the form (R^i, R^jF) that generate are those with $(i, n) = 1$, so exactly those in class (2). Finally, consider a sequence of the form (R^iF, R^jF). If this generates, then so does (R^iF, R^{j-i}) because $R^jF = R^{j-i}R^iF$. By the above, $(j - i, n) = 1$. Conversely, if $(j - i, n) = 1$ then (R^iF, R^{j-i}) generates. Since $R^{j-i} = (R^jF)(R^iF)^{-1}$, this means (R^iF, R^jF) generates, as desired. So class (3) exactly describes the generating sequences of this form.

Another way to study the set $\Gamma_n(G)$ is to work with group actions on it. For instance, the symmetric group S_n acts (on the right) on $\Gamma_n(G)$ by permuting the coordinates of a generating sequence, i.e. by

$$(g_1, \dots, g_n) \cdot \sigma = (g_{\sigma(1)}, \dots, g_{\sigma(n)})$$

Another group that acts on $\Gamma_n(G)$ is $\mathrm{Aut}(G)$, the group of automorphisms of G, by applying each automorphism coordinatewise:

$$\alpha \cdot (g_1, \dots, g_n) = (\alpha(g_1), \dots, \alpha(g_n)).$$

The action of the automorphism group is particularly useful:

Proposition 1.10. *The action of* $\mathrm{Aut}(G)$ *on* $\Gamma_n(G)$ *is free (i.e. if* $\alpha \in \mathrm{Aut}(G)$ *satisfies* $\alpha \cdot s = s$ *for some* $s \in \Gamma_n(G)$*, then* $\alpha = \mathrm{id}$*). Therefore, the order* $|\mathrm{Aut}(G)|$ *divides the cardinality* $|\Gamma_n(G)|$*.*

Proof. Let $s = (g_1, \ldots, g_n)$ be a generating sequence so that $\alpha \cdot s = s$. Then $\alpha(g_i) = g_i$ for each i. Since α is an automorphism, we know $\alpha(g^{-1}) = \alpha(g)^{-1}$ and $\alpha(gh) = \alpha(g)\alpha(h)$, so $\alpha(g) = g$ for any g that can be written as a product of the g_i and their inverses. Since $(g_1, \ldots, g_n)$ is a generating sequence, this holds for all $g \in G$, and α is the identity function.

By the orbit-stabilizer theorem, each orbit of $\Gamma_n(G)$ under $\mathrm{Aut}(G)$ has size $|\mathrm{Aut}(G)|$. Since $\Gamma_n(G)$ is a disjoint union of its orbits, $|\Gamma_n(G)|$ is a multiple of $|\mathrm{Aut}(G)|$. □

Definition 1.11. We let $h_n(G) = |\Gamma_n(G)|/|\mathrm{Aut}(G)|$. This is called the *reduced Eulerian function*. It is an integer by the above proposition, and is equal to the number of orbits of the action of $\mathrm{Aut}(G)$ on $\Gamma_n(G)$.

The reduced Eulerian function will be important to us later on. For now, we remark that when Hall defined this function in his paper [Hal36], he used his Möbius function formula to compute $h_2(A_5) = 19$. He used this to prove the fact we mentioned above, that A_5^{19} can be generated by two elements but A_5^{20} cannot; see Section 2.3.

Another method for thinking about generating sequences is in terms of homomorphisms out of a free group. If we let F_n denote the free group with n generators and let $x_1, \ldots, x_n$ be a free basis for F_n, then for any sequence $(g_1, \ldots, g_n)$ there is a unique homomorphism $\pi : F_n \to G$ with $\pi(x_i) = g_i$ for each i. Since the image of F_n is generated by the image of the generating set $\{x_1, \ldots, x_n\}$, π is surjective if and only if $(g_1, \ldots, g_n)$ is a generating sequence. Therefore, there is a bijective correspondence between $\Gamma_n(G)$ and the set of surjective homomorphisms $\pi : F_n \to G$.

Definition 1.12. If $s = (g_1, \ldots, g_n)$ is a generating sequence, we let $\pi_s : F_n \to G$ denote the surjective homomorphism given by $\pi_s(x_i) = g_i$. We let K_s denote the kernel of π_s.

The kernels K_s are closely related to the action of $\mathrm{Aut}(G)$ on $\Gamma_n(G)$:

Proposition 1.13. *Two generating sequences $s, t \in \Gamma_n(G)$ are in the same orbit under $\mathrm{Aut}(G)$ if and only if $K_s = K_t$.*

Proof. Let $s = (s_1, \ldots, s_n)$ and $t = (t_1, \ldots, t_n)$. If s, t are in the same orbit, there is an automorphism α with $\alpha \cdot s = t$. Then,

$$\pi_t(x_i) = t_i = \alpha(s_i) = \alpha(\pi_s(x_i)).$$

Since π_t and $\alpha \circ \pi_s$ agree on a generating set $\{x_1, \ldots, x_n\}$ of F_n, they are equal as functions. Since α is invertible, $\ker \pi_t = \ker(\alpha \circ \pi_s) = \ker \pi_s$.

Conversely, assume $K_s = K_t$. Since $K_s \subseteq \ker \pi_t$, the universal property for quotient groups implies that π_t factors through F_n/K_s, i.e. $\pi_t = \alpha \circ \pi_s$ for some homomorphism $\alpha : G \to G$. Similarly, since $K_t \subseteq \ker \pi_s$, $\pi_s = \beta \circ \pi_t$ for some homomorphism $\beta : G \to G$. Combining these equations, we get $\pi_s = \beta \circ \alpha \circ \pi_s$, and surjectivity of π_s means $\beta \circ \alpha$ is the identity map. Similarly $\alpha \circ \beta$ is the identity; this means α is invertible and hence an automorphism. Therefore, we have

$$t_i = \pi_t(x_i) = \alpha(\pi_s(x_i)) = \alpha(s_i),$$

so $t = \alpha \cdot s$ and hence t, s are in the same orbit. □

1.2 Irredundant Generating Sequences

Irredundant generating sequences are those that contain "just enough" elements to generate:

Definition 1.14. A generating sequence $(g_1, \ldots, g_n)$ of a finite group G is *irredundant* if no proper subsequence generates G (i.e. if there is no element g_i that we can remove from the sequence and still get a generating sequence). A generating sequence that has a proper subsequence that generates is called *redundant*.

Any generating sequence of length $r(G)$ is irredundant, because no sequence of length $r(G) - 1$ generates G. However, irredundant generating sequences do not need to be of length $r(G)$. For instance, $(\overline{2}, \overline{3})$ is an irredundant generating sequence of the cyclic additive group of $\mathbb{Z}/6\mathbb{Z}$.

It is often useful to know whether generating sequences that we are working with are redundant or not. In particular, we would like to know when a generating sequence is "long enough" that it is forced to be redundant. Accordingly, we define:

Definition 1.15. For a finite group G, let $\overline{r}(G)$ be the maximum length of an irredundant generating sequence.

By the pigeonhole principle, $|G|$ is a crude upper bound for $\overline{r}(G)$ (any sequence of length greater than $|G|$ must have a repeated element, so must be redundant). In particular, this shows that $\overline{r}(G)$ exists for every finite group.

We can get a better upper bound on $\overline{r}(G)$ as follows. For a positive integer n, let $\lambda(n)$ be the total number of prime factors of n (so if n has prime factorization $p_1^{a_1} \cdots p_k^{a_k}$, then $\lambda(n) = a_1 + \cdots + a_n$). Then, we have:

Proposition 1.16. $\overline{r}(G) \leq \lambda(|G|)$.

Proof. Suppose $(g_1, \ldots, g_n)$ is an irredundant generating sequence of G. Define a sequence of subgroups $G_i = \langle g_1, \ldots, g_i \rangle$ of G (note $G_n = G$, and take $G_0 = 1$ for convenience). By irredundancy, g_i cannot be generated from $g_1, \ldots, g_{i-1}$, so we have $G_{i-1} \lneq G_i$ for each i. This means each index $[G_i : G_{i-1}]$ is greater than 1, so some prime q_i divides it. We can write the order of G as a product of the indices of these subgroups:

$$|G| = [G_n : G_{n-1}] \cdots [G_2 : G_1][G_1 : G_0].$$

This means that a product of n primes $q_n \cdots q_1$ divides $|G|$, so n cannot be greater than $\lambda(|G|)$. $\square$

This upper bound is sharp in the sense that there are groups with $\overline{r}(G) = \lambda(G)$. Berkovich [Ber86] proves that such groups are exactly the "complemented groups" discussed by Hall [Hal37]. One of the equivalent classifications of such groups is as subgroups of direct products of groups of squarefree order. (Groups of squarefree order have been completely classified as well, by Hölder in 1895. A corollary to this classification is that every group of squarefree order is a semidirect product of two cyclic groups.)

Also, we remark that it is nontrivial to compute $\overline{r}(G)$, even for well-understood groups. It is easy to see that symmetric groups satisfy $\overline{r}(S_n) \geq n - 1$, as the sequence of transpositions $(1\ 2), (1\ 3), \ldots, (1\ n)$ is irredundant. Similarly, alternating groups satisfy $\overline{r}(A_n) \geq n - 2$, by taking a sequence $(1\ 2\ 3), (1\ 2\ 4), \ldots, (1\ 2\ n)$ of 3-cycles. It turns out that equality holds in these cases: we have $\overline{r}(S_n) = n-1$ and $\overline{r}(A_n) = n-2$ (see [Whi00]). However, the proof of this uses the classification theorem of finite simple groups!

Given a group G, we know that the minimum size of an irredundant generating sequence is $r(G)$ and the maximum size of an irredundant generating sequence is $\overline{r}(G)$. It turns out that there is an irredundant sequence of every length between $r(G)$ and $\overline{r}(G)$ as well. This is a consequence of a more general result of Tarski; see [Tar75] or Chapter II.4 of [BS]. We give a proof specialized to our case:

Theorem 1.17. *If $r(G) \leq n \leq \overline{r}(G)$, G has an irredundant generating sequence of length n.*

Proof. We show that if we have an irredundant generating sequence s of length $k \geq r(G) + 1$, then there exists an irredundant sequence of length $k - 1$. This suffices to prove the theorem, as we can start with $k = \overline{r}(G)$ and recursively find irredundant sequences of all sizes down to $r(G)$.

Start with sequence $s = (s_1, \ldots, s_k)$ of length $k \geq r(G) + 1$. For any element $g \in G$, we can write g as a product of some sequence of the s_i and their inverses. We will come up with a metric for comparing "how far apart" s and another sequence t are, and show that the "closest" sequence t of length less than k must have length $k - 1$ and be irredundant.

To formalize this, let $\ell(g)$ denote the length of the shortest word in the s_i and s_i^{-1} that equals g in G. Given a sequence $t = (g_1, \ldots, g_n) \in G^n$, define $\ell(t) = \sum \ell(g_i)$, $m(t) = \max\{\ell(g_i)\}$, and $f(t) = |\{i : \ell(t_i) = m(t)\}|$. The function $\ell(t)$ is our measure of how far t is from s, with $m(t)$ and then $f(t)$ serving as "tiebreakers".

We then take t to be a generating sequence of length less than k that is minimal relative to this measure; i.e. if t' is another generating sequence of less than k, then either $\ell(t) < \ell(t')$, or $\ell(t) = \ell(t')$ and $m(t) < m(t')$, or $\ell(t) = \ell(t')$, $m(t) = m(t')$, and $f(t) \leq f(t')$. Such a t exists because the set of all images $(\ell(t), m(t), f(t)) \in \mathbb{N}^3$ is finite, and therefore has a minimal element with respect to the lexicographic order. We then need to show that this minimality implies t is irredundant and of length $k-1$.

So, assume that t is redundant, and thus a proper subsequence t' of t generates. By rearranging t, we can assume without loss of generality that t' is t with t_1 removed. Then, we must have $\ell(t_1) = 0$ (otherwise $\ell(t') < \ell(t)$ and t can't be minimal), so $t_1 = 1$. Moreover, we must have $\ell(t_i) > 1$ for some i, or else the non-identity elements of t would all be either s_j or s_j^{-1}; since there are fewer than k of them, t would correspond to a proper subsequence of s and thus could not generate. Pick i so $\ell(t_i) = m(t)$, and write $t_i = s_i^{\pm 1} g$ for some g with $\ell(g) = \ell(t_i) - 1$. If we define t'' by replacing $t_1 = 1$ with s_i and replacing t_i with g, then $\ell(t'') = \ell(t)$ (as $\ell(t_1'') = \ell(t_1) + 1$ and $\ell(t_i'') = \ell(t_i) - 1$). Moreover, we must either have $m(t'') < m(t)$ or $f(t'') < f(t)$, because we replaced t_i that satisfied $\ell(t_i) = m(t)$ with t_i'' such that $\ell(t_i'') < m(t)$. Either way, $(\ell(t''), m(t''), f(t'')) < (\ell(t), m(t), f(t))$, contradicting minimality of t. So, t must be irredundant.

Now, let t have length $m < k$. Again we must have $\ell(t_i) > 1$ for some i, and we can assume without loss of generality that $\ell(t_1) = m(t) > 1$. Write $t_1 = s_j^{\pm 1} g$ with $\ell(g) = \ell(t_1) - 1$, and let τ be the generating sequence $(s_j, g, t_2, \ldots, t_m)$ of length $m+1$. Note $\ell(t) = \ell(\tau)$, as we replace t_1 with s_j and g with $\ell(s_j) = 1$ and $\ell(g) = \ell(t_1) - 1$. Again we either have $m(\tau) < m(t)$ or $m(\tau) = m(t)$ and $f(\tau) < f(t)$ because we removed an element of maximum ℓ-value. So, $(\ell(\tau), m(\tau), f(\tau)) < (\ell(t), m(t), f(t))$. By minimality of t, τ must not have length less than k, and thus its length $m+1$ must equal k. This proves that the length of t is $k-1$, as desired. □

So, $r(G)$ and $\overline{r}(G)$ completely determine the set of lengths attained by irredundant sequences in G. On the other hand, there are no restrictions on $r(G)$ or $\overline{r}(G)$ besides $r(G) \leq \overline{r}(G)$. In particular, for any fixed $r \leq \overline{r}$, we can construct a group with $r(G) = r$ and $\overline{r}(G) = \overline{r}$. An example of such a group is $Z_p^r \times Z_m$ where p is a prime number and m is a product of $\overline{r} - r$ primes $p_1, \ldots, p_k$ which are distinct from each other and from p. This is a consequence of some results proven later (in particular, Propositions 2.8, 2.9, 2.10, and 2.20).

An interesting open question is to characterize groups with $r(G) = \overline{r}(G)$, and in particular what groups satisfy $r(G) = \overline{r}(G) = 2$. Proposition 2.20 shows that Z_p^r satisfies satisfies $r(G) = \overline{r}(G) = r$ for p prime. A fact known as the "Burnside basis theorem" implies that $r(G) = \overline{r}(G)$ for any p-group G. (This theorem states that $G/\Phi \cong Z_p^r$, where Φ is the "Frattini subgroup" defined in the next section. Proposition 1.30 in that section implies $r(G) = r(Z_p^r)$ and $\overline{r}(G) = \overline{r}(Z_p^r)$).

1.3 Quotients and Generating Sequences

If $h : G \to H$ is a surjective homomorphism, applying h element-wise to a generating sequence of G gives a generating sequence of H. This proves $r(H) \leq r(G)$, and moreover induces a map $\Gamma_n(G) \to \Gamma_n(H)$:

Definition 1.18. If G, H are finite groups and $h : G \to H$ is a surjective homomorphism, define $\overline{h} : \Gamma_n(G) \to \Gamma_n(H)$ by $\overline{h}(g_1, \ldots, g_n) = (h(g_1), \ldots, h(g_n))$.

By the first isomorphism theorem, it is equivalent to work with quotient groups G/N. While we know $r(G/N) \leq r(G)$, it is less obvious how to relate $\overline{r}(G/N)$ to $\overline{r}(G)$ and how to relate $\varphi_n(G/N)$ and $\varphi_n(G)$. The right way to approach this problem is to start with generating sequences in G/N:

Definition 1.19. If $s = (s_1, \ldots, s_n)$ is a sequence of G/N, a *lift* of s is a sequence $\hat{s} = (g_1, \ldots, g_n)$ that projects to s (so $s_i = g_i N$ for each i).

An obvious question to ask is whether a particular generating sequence $s \in \Gamma_n(G/N)$ has a lift to a generating sequence of $\Gamma_n(G)$. It is clear that $n \geq r(G)$ is a necessary condition. Surprisingly, this is also sufficient; every generating sequence of G/N of length at least $r(G)$ has a lift. This was proven by Gaschütz in [Gas55]. We prove it as a corollary to a stronger result, which is attributed to Roquette (see p.361 of [FJ08]).

Lemma 1.20. *Any two generating sequences s, s' of G/N with the same length have the same number of lifts to generating sequences of G. (This number may be zero.)*

Proof. We prove this by induction on $|G|$. The base case of $|G| = 1$ is trivial, as G/N must be the trivial group, so it only has one generating sequence.

For the inductive step, fix some G and N, and assume that we know the result holds for every group H (and every $N' \trianglelefteq H$) with $|H| < |G|$. Let s be a length k generating sequence of G/N. There are $|N|^k$ lifts of s to G; if $s = (g_1N, \ldots, g_kN)$ then we can take any $\hat{s} = (g_1n_1, \ldots, g_kn_k)$ for $n_i \in N$.

Any such $\hat{s}$ generates some subgroup $H \leq G$. For each subgroup, define $f_H(s)$ to be the number of lifts $\hat{s}$ that generate H. We have the identity

$$|N|^n = \sum_{H \leq G} f_H(s),$$

which we can rearrange to

$$f_G(s) = |N|^n - \sum_{H < G} f_H(s).$$

Now, $f_G(s)$ is the number of lifts of s to generating sequences of G. To show that $f_G(s)$ is constant as a function of $s \in \Gamma_k(G/N)$, it suffices to show that the right hand side is independent of s.

Fix some lift $\hat{s}$ so that $\langle g_1n_1, \ldots, g_kn_k \rangle = H < G$. Any element $gN \in G/N$ can be written as a product of the $g_iN = (g_in_i)N$ and their inverses. Thus, the corresponding product $g' \in G$ of the (g_in_i) and their inverses satisfies $g'N = gN$, so $g = g'n$ for $g' \in \langle g_1n_1, \ldots, g_kn_k \rangle = H$ and $n \in N$. Since g was arbitrary we have $G = HN$. Therefore, we have

$$f_G(s) = |N|^n - \sum_{H < G, HN = G} f_H(s).$$

For fixed H with $HN = G$, we can apply the second isomorphism theorem to get an isomorphism

$$G/N = HN/N \cong H/(H \cap N).$$

Applying this isomorphism to s gives a generating sequence $\widetilde{s}$ of $H/(H \cap N)$. A generating sequence $(h_1, \ldots, h_n)$ of H projects to s in G/N if and only if it projects to $\widetilde{s}$ in $H/(H \cap N)$. So, $f_H(s)$ is exactly the number of lifts of a length k generating sequence in $H/(N \cap H)$ to a generating sequence of H. Since $|H| < |G|$, we know by induction that $f_H(s)$ is independent of s. Therefore, $f_G(s) = |N|^n - \sum_{H<G} f_H(s)$ is independent of s, finishing the inductive step and hence the proof. □

Corollary 1.21 (Gaschütz's Lemma)**.** *If $n \geq r(G)$ and $s \in \Gamma_n(G/N)$, s has a lift to a generating sequence of G.*

Proof. Since $n \geq r(G)$, G has a generating sequence s' of length n, which projects to a generating sequence $\widetilde{s}' \in \Gamma_n(G/N)$. By Lemma 1.20, s and $\widetilde{s}'$ have the same number of lifts to generating sequences of G. Since $\widetilde{s}'$ has at least one, so does s. □

This can easily extend to surjective homomorphisms $h : G \to H$. Given a sequence $(g'_1, \ldots, g'_n)$ of H, we define a lift to be a sequence $(g_1, \ldots, g_n)$ in G so that $h(g_i) = g'_i$. By the first isomorphism theorem, Lemma 1.20 gives that all generating sequences of the same length in H have the same number of lifts to generating sequences of G, and Corollary 1.21 gives that any generating sequence in H of length $n \geq r(G)$ has a lift to a generating sequence of G. Therefore, we have:

Corollary 1.22. *If $h : G \to H$ is a surjective homomorphism and $n \geq r(G)$, the map $\overline{h} : \Gamma_n(G) \to \Gamma_n(H)$ defined in Definition 1.18 is surjective.*

Moreover, we can define:

Definition 1.23. If $\varphi : G \to H$ is a surjective homomorphism, we define the *n-th lifting index* $[\![G : H]\!]_n$ as the number of lifts of a length n generating sequence of H to a generating sequence of G.

Though our notation does not include the homomorphism $h : G \to H$, it is not immediately obvious that $[\![G : H]\!]_n$ is the same for all surjective homomorphisms $h : G \to H$. The following proposition shows this is true, and relates the Eulerian function of G to the Eulerian function of any homomorphic image H:

Proposition 1.24. *If H is a homomorphic image of G, we have $\varphi_n(G) = [\![G : H]\!]_n \varphi_n(H)$. In particular, this gives an alternative definition of the lifting index as $\frac{\varphi_n(G)}{\varphi_n(H)}$, which is independent of the surjective homomorphism $G \to H$.*

Proof. We can write $\Gamma_n(G)$ as the disjoint union of sets Γ_s (for $s \in \Gamma_n(H)$), where Γ_s consists of the sequences projecting to s (under some fixed surjective homomorphism $\varphi : G \to H$). Since each set Γ_s has cardinality $[\![G : H]\!]_n$, we have

$$\varphi_n(G) = |\Gamma_n(G)| = [\![G : H]\!]_n |\Gamma_n(H)| = [\![G : H]\!]_n \varphi_n(H),$$

as desired. □

Corollary 1.25. *Lifting indices are multiplicative; if H is a homomorphic image of G and K is a homomorphic image of H, then*

$$[\![G : K]\!]_n = [\![G : H]\!]_n [\![H : K]\!]_n.$$

Proof.

$$[\![G : K]\!]_n = \frac{\varphi_n(G)}{\varphi_n(K)} = \frac{\varphi_n(G)}{\varphi_n(H)} \frac{\varphi_n(H)}{\varphi_n(K)} = [\![G : H]\!]_n [\![H : K]\!]_n.$$ □

Finally, as a corollary to the proof of Lemma 1.20, we can give a formula for $[\![G : H]\!]_n$ using the Möbius function from Definition 1.6.

Proposition 1.26. *If N is the kernel of a surjective homomorphism $G \to H$, then*

$$[\![G : H]\!]_n = \sum_{K \leq G : KN = G} \mu_G(K) |K \cap N|^n.$$

Proof. In the proof of Lemma 1.20, we wrote

$$|N|^n = \sum_{K \leq G} f_K(s),$$

where $f_K(s)$ was the number of lifts of s to a generating sequence of K. Moreover, we showed that $f_K(s)$ was only nonzero if $KN = G$, and that in that case we proved $f_K(s)$ was the number of lifts of a generating sequence in $K/K \cap N$ to K, i.e. $f_K(s) = [\![K : K \cap N]\!]_n$. Thus,

$$|N|^n = \sum_{K \leq G : KN = G} [\![K : K \cap N]\!]_n.$$

For any $L \leq G$, we can apply the same argument to the set of lifts of some $s \in \Gamma_n(L \cap N)$ to sequences in L, to get

$$|L \cap N|^n = \sum_{K \leq L : K(L \cap N) = L} [\![K : K \cap (L \cap N)]\!]_n = \sum_{K \leq L : K(L \cap N) = L} [\![K : K \cap N]\!]_n. \tag{2}$$

We want to apply Möbius inversion (Theorem 1.7) to obtain the desired formula. This requires a little care, as the theorem requires a summatory function $F(L) = \sum_{K \leq L} f(K)$ that sums over all $K \leq L$. To do this, define an "indicator function" for subgroups (as a function of K):

$$1_{KN=G} = \begin{cases} 1 & KN = G \\ 0 & KN \neq G \end{cases}.$$

We claim that we have

$$1_{LN=G}|L \cap N|^n = \sum_{K \leq L} 1_{KN=G} [\![K : K \cap N]\!]_n. \tag{3}$$

If $LN \neq G$, then $KN \neq G$ for any $K \leq L$, so both sides of equation 3 are zero and hence the equation holds. If $LN = G$, we claim $KN = G$ holds if and only if $K(N \cap L) = L$. To see this, note that if $KN = G$ then any element $\ell \in L$ is a product of an element $k \in K$ and $n \in N$, and moreover $n = k^{-1}\ell \in L$ means $\ell = kn \in K(N \cap L)$. Conversely, $K(N \cap L) = L$ implies $KN = K(N \cap L)N = LN = G$. Thus, if $LN = G$ we can use equation 2 to prove equation 3:

$$1_{LN=G}|L \cap N|^n = |L \cap N|^n = \sum_{K \leq L : K(N \cap L) = L} [\![K : K \cap N]\!]_n = \sum_{K \leq L} 1_{KN=G} [\![K : K \cap N]\!]_n.$$

We can then apply Möbius inversion to equation 3, which gives

$$[\![G : N]\!]_n = \sum_{K \leq G} \mu_G(K) 1_{KN=G} |K \cap N|^n = \sum_{K \leq G : KN = G} \mu_G(K) |K \cap N|^n. \qquad \square$$

Another corollary to Gaschütz's lemma is the behavior of $\overline{r}(G)$ under quotients:

Proposition 1.27. *If G is a finite group and N a normal subgroup, $\overline{r}(G/N) \leq \overline{r}(G)$.*

Proof. If $\overline{r}(G/N) < r(G)$, this is trivial. Otherwise, Let s be an irredundant generating sequence of length $\overline{r}(G/N)$ in G/N. By Gaschütz's lemma, s has a lift to a generating sequence $\hat{s}$ of G. Then $\hat{s}$ is irredundant; if a proper subseuqence generated G, its projection to G/N would be a proper subsequence of s that generated G/N. $\square$

We end this section by discussing the Frattini subgroup $\Phi = \Phi(G)$ for a finite group G. The Frattini subgroup consists of all elements that are "non-generators" of G. The relation between generating sequences of G and G/Φ is very simple, and we can use it in computations.

Recall that a maximal subgroup of G is a proper subgroup $M \leq G$ that is not properly contained in any proper subgroup (so if $M \leq H \leq G$, either $H = M$ or $H = G$). Then, we define:

Definition 1.28. The *Frattini subgroup* $\Phi = \Phi(G)$ is the intersection $\bigcap M$ of all maximal subgroups of G.

Proposition 1.29. *The Frattini subgroup is exactly the set of non-generators of G: elements $g \in G$ so that if $X \cup \{g\}$ generates G, then X generates G. Moreover, the Frattini subgroup is characteristic (for any automorphism α of G, the image $\alpha[\Phi]$ is equal to Φ), and therefore is normal in G.*

Proof. First, let x be a non-generator. If M is a maximal subgroup, the set M does not generate G, and by definition of x being a non-generator $\{x\} \cup M$ does not generate G. By maximality, $\langle M, x \rangle = M$ and hence $x \in M$. Since M was arbitrary, $x \in \bigcap M = \Phi$.

Conversely, let $x \in \Phi$, so $x \in M$ for every maximal subgroup M. Assume $X \cup \{x\}$ generates G, but X does not. Since G is finite, the lattice of subgroups is finite, so $\langle X \rangle$ is contained in some maximal subgroup M. This means $X \subseteq M$, and we also have $x \in M$ because $x \in \Phi$. Therefore $\langle X, x \rangle \subseteq M$, which is a contradiction. So if $X \cup \{x\}$ generates G then so must X, which proves x is a non-generator.

Finally, let α be an automorphism of G and $M \leq G$ a maximal subgroup. Then the image $\alpha[M]$ is maximal because α^{-1} takes any subgroup between $\alpha[H]$ and G to a subgroup between H and G. Therefore α permutes the set of maximal subgroups (as α^{-1} induces its inverse on this set), giving

$$\alpha[\Phi] = \alpha\left[\bigcap M\right] = \bigcap \alpha[M] = \bigcap M = \Phi.$$

Thus Φ is a characteristic subgroup and hence a normal subgroup. $\square$

Given these basic properties of the Frattini subgroup, we can prove:

Proposition 1.30. *Let G be a finite group and $N \trianglelefteq G$ be contained in the Frattini subgroup Φ. Then for any $s \in \Gamma_n(G/N)$, any lift $\hat{s}$ of s to G is a generating sequence of G. Therefore, $\varphi_k(G) = |N|^n \varphi_k(G/N)$, $r(G) = r(G/N)$, and $\overline{r}(G) = \overline{r}(G/N)$.*

Proof. Let $(g_1, \ldots, g_n)$ be such that $(g_1N, \ldots, g_kN)$ generates G/N. Any $gN \in G/N$ can be written as a product of the g_iN and their inverses, and therefore any $g \in G$ can be written as a product of the g_i and their inverses along with some $n \in N$. This means that the set $\{g_1, \ldots, g_k, n_1, \ldots, n_\ell\}$ generates G, where the $n_1, \ldots, n_\ell$ are the elements of N. However, each n_i is in Φ so is a non-generator. Then we can thus remove n_ℓ and still have a generating set, then remove $n_{\ell-1}$ from this new set, and so on. Removing all of the elements n_i in this way, we get that $\{g_1, \ldots, g_k\}$ generates G.

Since there are $|N|^k$ lifts of s, this immediately implies that $[\![G : G/N]\!]_k = |N|^k$, and hence

$$\varphi_k(G) = |N|^n \varphi_k(G/N).$$

Also, since every generating sequence of G/N lifts, in particular the generating sequences of length $r(G/N)$ lift, so $r(G) \leq r(G/N)$ (and the reverse inequality holds in general). If s is an irredundant generating sequence in G, then consider the projection $\overline{s}$ in G/N. This is also irredundant, as if a proper subsequence of $\overline{s}$ generates then its lift to a proper subsequence of s also generates, which is a contradiction. This proves $\overline{r}(G) \leq \overline{r}(G/N)$, and the reverse inequality holds in general. □

As an easy application of this, we can compute $\overline{r}$ and φ_n for a cyclic group of prime power order Z_{p^k}. In particular:

Corollary 1.31. *If p is a prime number and $k \geq 1$, $\overline{r}(Z_{p^k}) = 1$ and $\varphi_n(Z_{p^k}) = (p^n - 1)p^{n(k-1)}$. In particular, if we let x be a generator of Z_{p^k}, a sequence $(g_1, \ldots, g_n)$ generates if and only if some g_i is not a power of x^p.*

Proof. As a cyclic group, we know that the subgroups of Z_{p^k} are exactly those of the form $\langle x^d \rangle$ for divisors d of p^k. Therefore, the subgroups are those of the form $\langle x^{p^i} \rangle$ for $0 \leq i \leq k$. Since these are nested subgroups, we can see that $\langle x^p \rangle$ contains all proper subgroups, and hence is the unique maximal subgroup. Therefore the Frattini subgroup Φ is $\langle x^p \rangle$. Moreover, the quotient Z_{p^k}/Φ is isomorphic to Z_p.

By example 1.5, we know that every sequence in Z_p that contains some non-identity element generates, and that $\varphi_n(Z_p) = p^n - 1$. Applying the previous proposition gives that

$$\varphi_n(Z_{p^k}) = |\langle x^p \rangle|^k \varphi_n(Z_p) = (p^{k-1})^n (p^n - 1),$$

and that every sequence in Z_{p^k} that is not contained in Φ generates. Finally, note that it is trivial that $\overline{r}(Z_p) = 1$ (any generating sequence contains a one-element generating subsequence), so $\overline{r}(Z_{p^k}) = 1$ as well. □

2 Generating Sequences in Direct Products

If we understand the behavior of generating sequences for two groups G, H, we can ask what happens for the direct product $G \times H$. If $s = (g_1, \ldots, g_n)$ is a sequence in G and $t = (h_1, \ldots, h_n)$ is a sequence in H, we can form a sequence

$$((g_1, h_1), \ldots, (g_n, h_n))$$

in $G \times H$, which we denote (s, t) (using the natural identification of $(G \times H)^n$ with $G^n \times H^n$).

Since the coordinate projections $G \times H \to G$ and $G \times H \to H$ are surjective homomorphisms, by the previous section we know that if (s, t) is a generating sequence of $G \times H$ then s is a generating sequence of G and t is a generating sequence of H. Therefore, we can identify the generating sequences of $G \times H$ with a subset of $\Gamma_n(G) \times \Gamma_n(H)$. Some easy properties are:

Proposition 2.1. *If G and H are finite groups, $\max\{r(G), r(H)\} \leq r(G \times H) \leq r(G) + r(H)$, $\overline{r}(G \times H) \geq \overline{r}(G) + \overline{r}(H)$, and $\varphi_n(G \times H) \leq \varphi_n(G) \cdot \varphi_n(H)$.*

Proof. The statement about φ_n is immediate from the identification of $\Gamma_n(G \times H)$ as a subset of $\Gamma_n(G) \times \Gamma_n(H)$. For r, note that since G, H are projections of $G \times H$, we have $r(G), r(H) \leq r(G \times H)$. Moreover, if $(g_1, \ldots, g_n)$ is a length $r(G)$ generating sequence of G and $(h_1, \ldots, h_m)$ is a length $r(H)$ generating sequence of H, then

$$((g_1, 1), \ldots, (g_n, 1), (1, h_1), \ldots, (1, h_m))$$

is a generating sequence of $G \times H$, proving $r(G \times H) \leq r(G) + r(H)$.

Finally, if $(g_1, \ldots, g_n)$ is a length $\overline{r}(G)$ irredundant generating sequence in G and $(h_1, \ldots, h_m)$ is a length $\overline{r}(H)$ irredundant generating sequence in H, then

$$((g_1, 1), \ldots, (g_n, 1), (1, h_1), \ldots, (1, h_m))$$

is an irredundant generating sequence of $G \times H$, as if we remove any element the projection of the sequence to one of the coordinates does not generate. Thus $\overline{r}(G \times H) \geq \overline{r}(G) + \overline{r}(H)$. □

2.1 Relatively Prime Sequences and Groups

We can ask when a given $(s, t) \in \Gamma_n(G) \times \Gamma_n(H)$ gives a generating sequence of $G \times H$. First, we give a name to such sequences:

Definition 2.2. Two sequences $s \in \Gamma_n(G)$ and $t \in \Gamma_n(H)$ are said to be *relatively prime* if (s, t) generates $G \times H$.

Recall that if $s = (s_1, \ldots, s_n)$ is a generating sequence of G, we define $\pi_s : F_n \to G$ as the surjective homomorphism with $\pi_s(x_i) = s_i$. We can give an equivalent formulation that can be easier to check, using the functions π_s:

Proposition 2.3. *Two sequences $s \in \Gamma_n(G)$ and $t \in \Gamma_n(H)$ are relatively prime if and only if for every $g \in G$ there is some $u \in F_n$ with $\pi_s(u) = g$ and $\pi_t(u) = 1$ and for every $h \in H$ there is $v \in F_n$ with $\pi_t(v) = h$ and $\pi_s(v) = 1$.*

Proof. Note that $\pi_{(s,t)}$ is the product of the maps π_s and π_t. If s and t are relatively prime then $\pi_{(s,t)}$ is surjective, so for any $g \in G$ there is u so that

$$(\pi_s(u), \pi_t(u)) = \pi_{(s,t)}(u) = (g, 1),$$

and similarly for any $h \in H$ there is v with $(\pi_s(v), \pi_t(v)) = (1, h)$.

Conversely, assume that s and t satisfy the specified condition. To prove (s, t) generates, it suffices to show $\pi_{(s,t)} = \pi_s \times \pi_t$ is surjective. To see this, fix $(g, h) \in G \times H$, and pick u, v that satisfy $\pi_s(u) = g$, $\pi_t(u) = 1$, $\pi_s(v) = 1$, and $\pi_t(v) = h$. Then,

$$\pi_{(s,t)}(uv) = (\pi_s(u), \pi_t(u))(\pi_s(v), \pi_t(v)) = (g, 1)(1, h) = (g, h).$$ □

Further recall that we defined K_s to be $\ker \pi_s$. We can give an third equivalent formulation of our definition in terms of the K_s. In fact, we can prove a more general lemma, which can be thought of as the "Chinese remainder theorem for groups":

Lemma 2.4. *Let G, G_1, G_2 be groups, and $h_1 : G \to G_1$ and $h_2 : G \to G_2$ be surjective homomorphisms with kernels K_1 and K_2, respectively. Then $h : G \to G_1 \times G_2$ given by $h(g) = (h_1(g), h_2(g))$ has kernel $K_1 \cap K_2$, and h is surjective if and only if $K_1 K_2 = G$.*

Proof. It is clear that $h(g) = (h_1(g), h_2(g)) = (1, 1)$ if and only if $g \in K_1$ and $g \in K_2$, so $\ker h = K_1 \cap K_2$. Moreover, assume that h is surjective, and fix some $x \in G$. Then $h(x) = (g_1, g_2) = (1, g_2)(g_1, 1)$, and by surjectivity there are x_1, x_2 with $h(x_1) = (g_1, 1)$ and $h(x_2) = (1, g_2)$. This means $x_1 \in K_1$, $x_2 \in K_2$, and we get $h(x) = h(x_1)h(x_2) = h(x_1 x_2)$. Therefore, $h(x^{-1} x_1 x_2) = 1$, so $y = x(x_1 x_2)^{-1} \in K_1 \cap K_2$, and thus

$$x = (yx_1)x_2 \in K_1 K_2.$$

Since x was arbitrary, $K_1K_2 = G$.

Conversely, assume $K_1K_2 = G$ (which means $K_2K_1 = G$ as well), and fix an element $(a_1, a_2) \in G_1 \times G_2$. By surjectivity of h_1, h_2, we have $a_1 = h_1(b_1)$ and $a_2 = h_2(b_2)$ for $b_i \in G$. As $G = K_2K_1$, we have $b_1 = k_2k_1$ for $k_i \in K_i$. Also, we have $k_2^{-1}b_2 = m_1m_2$ for $m_i \in K_i$. Then, consider $g = k_2m_1$. Note that

$$h_1(k_2m_1) = h_1(k_1k_2) = h_1(b_1) = a_1$$

because $m_1, k_1 \in K_1$, and similarly

$$h_2(k_2m_1) = h_2(k_2m_1m_2) = h_2(k_2k_2^{-1}b_2) = a_2,$$

so $h(g) = (a_1, a_2)$. As (a_1, a_2) was arbitrary, h is surjective. □

Corollary 2.5. *Two sequences $s \in \Gamma_n(G)$ and $t \in \Gamma_n(H)$ are relatively prime if and only if $K_sK_t = F_n$.*

Now, we consider a sufficient condition for all pairs of generating sequences $(s, t) \in \Gamma_n(G) \times \Gamma_n(H)$ to be relatively prime:

Definition 2.6. We say two groups G and H are *relatively prime* if they have no nontrivial common quotient groups (i.e. if $G/N \cong H/N'$ for $N \trianglelefteq G$ and $N' \trianglelefteq H$, then G/N and H/N' are the trivial group 1).

Proposition 2.7. *If G and H are relatively prime groups, then any $s \in \Gamma_n(G)$ and $t \in \Gamma_n(H)$ are relatively prime sequences.*

Proof. Recall that we let $K_s = \ker \pi_s$ and $K_t = \ker \pi_t$. Since π_s is a surjective homomorphism, it preserves normal subgroups, so $\pi_s[K_t]$ is normal in G. Let $\pi : G \to G/\pi_s[K_t]$ be the canonical projection. Then $\pi \circ \pi_s$ takes K_t to zero by definition, so by the universal property of quotient groups, it factors through $F_n/K_t \cong H$.

This means that there is a surjective homomorphism $\pi' : H \to G/\pi_s[K_t]$, so $G/\pi_s[K_t]$ is a quotient group of both G and H. Since G and H are relatively prime, this group must be trivial, so $\pi_s[K_t] = G$. Therefore, for any $g \in G$, there is $u \in K_t$ with $\pi_s(u) = g$, and moreover by definition of K_t we have $\pi_t(u) = 1$. By a parallel argument, for any $h \in H$ there is $v \in F_n$ with $\pi_s(v) = 1$ and $\pi_t(v) = h$. Thus s and t are relatively prime by Proposition 2.3. □

Corollary 2.8. *If G and H are relatively prime groups, we have $r(G \times H) = \max\{r(G), r(H)\}$, $\overline{r}(G \times H) = \overline{r}(G) + \overline{r}(H)$, and $\varphi_n(G \times H) = \varphi_n(G)\varphi_n(H)$.*

Proof. Since every $s \in \Gamma_n(G)$ and $t \in \Gamma_n(H)$ are relatively prime, $\Gamma_n(G \times H)$ is in bijective correspondence with $\Gamma_n(G) \times \Gamma_n(H)$, and hence $\varphi_n(G \times H) = \varphi_n(G)\varphi_n(H)$. Moreover, if $n = \max\{r(G), r(H)\}$ we can find length n generating sequences of G and H and combine them to get a length n generating sequence of $G \times H$, proving $r(G \times H) \leq \max\{r(G), r(H)\}$ (and the reverse inequality holds in general).

To show $\overline{r}(G \times H) = \overline{r}(G) + \overline{r}(H)$, it suffices to show that if $n > \overline{r}(G) + \overline{r}(H)$ then a sequence s of length n in $G \times H$ is necessarily redundant. To see this, note that there is a subsequence s_1 of length at most $\overline{r}(G)$ such that it projection to G is a generating sequence. Similarly there is a subsequence s_2 of length at most $\overline{r}(H)$ with a projection that generates H. Then, let $s_1 \cup s_2$ denote the subsequence of s containing all coordinates in s_1 or s_2. This is a proper subsequence of s by length considerations. Moreover, if we project to each coordinate, we get generating sequences of G and H respectively, and since G and H are relatively prime the combination of these sequences (which is $s_1 \cup s_2$) generates. Hence s has a proper subsequence that generates, so is redundant. □

The following proposition is the motivation for the terminology "relatively prime groups," which is in turn the motivation for the terminology "relatively prime sequences." Also, as a corollary, we can prove Equation 1, the formula for $\varphi_n(Z_m)$ that we stated earlier.

Proposition 2.9. *If G and H are finite groups with $|G|$ and $|H|$ relatively prime integers, then G and H are relatively prime groups.*

Proof. If Q is a common quotient group of G and H, then we know $|Q|$ divides both $|G|$ and $|H|$. Since these numbers are relatively prime, we must have $|Q| = 1$, so the only common quotient of G and H is the trivial group. □

Corollary 2.10. *Letting Z_m be the cyclic group of order $m = p_1^{a_1} \cdots p_k^{a_k}$ (with the p_i distinct primes and the a_i positive), we have $r(Z_m) = 1$, $\overline{r}(Z_m) = k$, and*

$$\varphi_n(Z_m) = p_1^{n(a_1-1)}(p_1^n - 1) \cdots p_k^{n(a_k-1)}(p_k^n - 1).$$

In particular, a sequence $(g_1, \ldots, g_n)$ in Z_m generates if and only if its projection to $Z_{p_i^{a_i}}$ generates for each i.

Proof. We know that $Z_m \cong Z_{p_1^{a_1}} \times \cdots \times Z_{p_k^{a_k}}$. We then prove the result by induction on k. The base case was established in Corollary 1.31, where we showed $r(Z_{p^i}) = \overline{r}(Z_{p^i}) = 1$ and $\varphi_n(Z_{p^i}) = p^{n(i-1)}(p^n - 1)$. For the inductive step, the proposition implies $Z_{p_1^{a_1}} \times \cdots \times Z_{p_{k-1}^{a_{k-1}}}$ and $Z_{p_k^{a_k}}$ are relatively prime, and the desired claims follow from Proposition 2.7. □

2.2 Gaschütz's Theorem and Cosocles

If G and H are relatively prime groups, the propositions of the previous section are sufficient to determine much of the behavior of generating sequences of $G \times H$. If not, we need some more theoretical machinery. This is provided by another theorem of Gaschütz [Gas55]. We start with two new definitions:

Definition 2.11. Let G be a finite group. A *maximal normal subgroup* of G is a proper subgroup that is maximal with respect to being normal (i.e a group $M \trianglelefteq G$ so that if $M \leq M' \trianglelefteq G$ then $M' = M$ or $M' = G$). Note that this is *not* the same as being M being both a maximal subgroup and a normal subgroup.

One trivial property of maximal normal subgroups of finite groups is that if $N \trianglelefteq G$ then $N \leq M$ for a maximal normal subgroup M. We then define:

Definition 2.12. Let G be a finite group. Define $N \trianglelefteq G$ to be the intersection of all maximal normal subgroups of G. Define the *cosocle* $\mathrm{CoSoc}(G)$ as the quotient group G/N.

We note that calling G/N the "cosocle" is not standard, and the subgroup N is sometimes itself called the cosocle. The "socle" of a group, the subgroup generated by all minimal normal subgroups, is well-studied in the group theory literature. The best "dual" construction to this is the quotient by the intersection of all maximal normal subgroups, which is why we call this the cosocle. The cosocle and socle turn out to share a number of important properties (for instance, the socle and the cosocle are both always direct products of finite simple groups; we prove this for the cosocle in Proposition 2.14).

We start our study of the cosocle with a lemma about the subgroup N used to define it. One can think of this as the normal-subgroup analogue to the Frattini subgroup. It turns out to have its own version of a "non-generator property":

Lemma 2.13. *Let G be a finite group and N the intersection of all maximal normal subgroups of G. Then, if $M \trianglelefteq G$ satisfies $MN = G$ we have $M = G$.*

Proof. Suppose that $M \trianglelefteq G$ is a proper normal subgroup with $MN = G$. Then, M is contained in some maximal normal subgroup M'. By construction of N, we have $N \leq M'$, and hence $G = MN \leq M'$. This contradicts the fact that M' is a proper subgroup; so we must have that $M = G$. □

Note that the "non-generator property" for the Frattini subgroup can be expressed in the same way: if $H \leq G$ satisfies $H\Phi = G$, then $H = G$.

Proposition 2.14. *If G is a finite group, $\mathrm{CoSoc}(G)$ is isomorphic to a direct product of finite simple groups.*

Proof. Recursively define a sequence of maximal normal subgroups of G by taking M_i to be a maximal normal subgroup so that $M_1 \cap \cdots \cap M_{i-1} \not\subseteq M_i$. Since G is finite, we eventually get a sequence $M_1, \ldots, M_k$ so that $M_1 \cap \cdots \cap M_k$ is contained in every maximal normal subgroup. We then have $M_1 \cap \cdots \cap M_k$ is equal to the intersection N of all maximal normal subgroups, so $\mathrm{CoSoc}(G)$ is $G/(M_1 \cap \cdots \cap M_k)$ by definition.

Now, we claim that for each i, $G/(M_1 \cap \cdots \cap M_i) \cong G/M_1 \times \cdots \times G/M_i$. We prove this by induction on i; the base case of $i = 1$ is trivial. For the inductive step, note that $M_1 \cap \cdots \cap M_{i-1} \not\subseteq M_i$ means that the product $(M_1 \cap \cdots \cap M_{i-1})M_i$ is a normal subgroup properly containing M_i. Since M_i is maximal normal, we have $G = (M_1 \cap \cdots \cap M_{i-1})M_i$. Then, apply Lemma 2.4 (the Chinese Remainder Theorem for groups), using the canonical projections $h_1 : G \to G/(M_1 \cap \cdots \cap M_{i-1})$ and $h_2 : G \to G/M_i$. Since the product of the kernels $(M_1 \cap \cdots \cap M_{i-1})M_i$ is G, the lemma implies that the product map $h : G \to G/(M_1 \cap \cdots \cap M_{i-1}) \times G/M_i$ is surjective, and moreover $\ker h = M_1 \cap \cdots \cap M_{i-1} \cap M_i$. Thus, by the first isomorphism theorem and the inductive hypothesis,

$$\frac{G}{M_1 \cap \cdots \cap M_i} \cong \frac{G}{M_1 \cap \cdots \cap M_{i-1}} \times \frac{G}{M_i} \cong \frac{G}{M_1} \times \cdots \times \frac{G}{M_i}.$$

Thus, we have

$$\mathrm{CoSoc}(G) = \frac{G}{M_1 \cap \cdots \cap M_k} \cong \frac{G}{M_1} \times \cdots \times \frac{G}{M_k}.$$

Since each M_i is maximal normal, each quotient group G/M_i must be simple, as if $N/M_i \trianglelefteq G/M_i$ then $M_i \leq N \trianglelefteq G$ means $N = M_i$ or $N = G$. Thus we have proven $\mathrm{CoSoc}(G)$ is a direct product of finite simple groups. □

Proposition 2.15. *If G, H are finite groups,* $\mathrm{CoSoc}(G \times H) \cong \mathrm{CoSoc}(G) \times \mathrm{CoSoc}(H)$.

The bulk of this proof comes from the following lemma abut maximal normal subgroups of direct products:

Lemma 2.16. *Let G_1 and G_2 be finite groups, and $M \subseteq G_1 \times G_2$ a maximal normal subgroup. Define $M_1 = \{g : (g, 1) \in M\}$ and $M_2 = \{g : (1, g) \in M\}$. Then each M_i is either equal to G_i or a maximal normal subgroup of G_i. Moreover $M_1 \times M_2 \leq M$.*

Proof. Start by noting that $M_i \trianglelefteq G_i$. If $g, h \in M_1$ then $(g, 1)$ and $(h, 1)$ are in M, so $(g, 1)(h, 1)^{-1} = (gh^{-1}, 1) \in M$ means $gh^{-1} \in M_1$, and thus M_1 is a subgroup. Moreover, if $g \in M_1$ and $\alpha \in G_1$ then

$$(\alpha, 1)(g, 1)(\alpha, 1)^{-1} = (\alpha g \alpha^{-1}, 1) \in M$$

(because $M \trianglelefteq G$) means $\alpha g \alpha^{-1} \in M_1$. Thus $M_1 \trianglelefteq G_1$, and similarly $M_2 \trianglelefteq G_2$; this also implies $M_1 \times M_2 \trianglelefteq G_1 \times G_2$. Also, it is clear by construction that $M_1 \times M_2 \leq M$.

If $M_1 = G_1$, then we claim $M = G_1 \times M_2$ and that M_2 is maximal normal in G_2. We have $G_1 \times M_2 \subseteq M$ by construction, and for the converse note that if $(g, h) \in M$ then $G_1 \times 1 \subseteq M$ implies $(1, h) = (g, h)(g^{-1}, 1) \in M$. Then, $M_2 \neq G_2$ (because otherwise $M = G_1 \times G_2$ is not maximal normal). If $M_2 \leq N \trianglelefteq G$ then $M = G_1 \times M_2 \leq G_1 \times N \trianglelefteq G_1 \times G_2$, and maximal normality of M implies that N is either M_2 or G. Thus, if $M_1 = G_1$, M_2 is maximal normal in G_2. Similarly, if $M_2 = G_2$, then M_1 is maximal normal in G_1.

This leaves the case where M_1 and M_2 are proper normal subgroups of G_1 and G_2, respectively. Since $M_1 \times M_2 \leq M$, we have

$$\frac{M}{M_1 \times M_2} \trianglelefteq \frac{G_1 \times G_2}{M_1 \times M_2} \cong \frac{G_1}{M_1} \times \frac{G_2}{M_2}.$$

So, $M/(M_1 \times M_2)$ corresponds to a normal subgroup M' of $G_1/M_1 \times G_2/M_2$. Using the canonical isomorphism between $(G_1 \times G_2)/(M_1 \times M_2)$ and $G_1/M_1 \times G_2/M_1$, we can compute

$$M' = \{(gM_1, hM_2) : (g, h) \in M\}.$$

By the third isomorphism theorem,

$$\frac{G_1/M_1 \times G_2/M_2}{M'} \cong \frac{(G_1 \times G_2)/(M_1 \times M_2)}{M/(M_1 \times M_2)} \cong \frac{G}{M},$$

which we know is simple; hence M' is maximal normal.

Now, note that M' satisfies the following property: If $(M_1, gM_2) \in M'$ then $g \in M_2$ and if $(gM_1, M_2) \in M'$ then $g \in M_1$. To see this, note that $(M_1, gM_2) \in M'$ means $(m_1, gm_2) \in M$ for $(m_1, m_2) \in M_1 \times M_2$. Since $(m_1, m_2) \in M$ as well, $(1, g) = (m_1, gm_2)(m_1, m_2)^{-1} \in M$ implies $g \in M_1$ by definition. An identical proof works if $(gM_1, M_2) \in M'$.

If we define the projections $\pi_i : M' \to G_i/M_i$, M' is contained in $\pi_1[M'] \times G_2/M_2$ and $G_1/M_1 \times \pi_2[M']$. Moreover, containment is proper. To see this, assume that $M' = \pi_1[M'] \times G_2/M_2$. Since $G_2 \neq M_2$, we can pick $g \notin M_2$, and then $(M_1, gM_2) \in \pi_1[M'] \times G_2/M_2 = M'$. By the previous paragraph, this means $g \in M_2$, a contradiction. Thus we have a chain of inclusions

$$M' \lneq \pi_1[M'] \times G_2/M_2 \trianglelefteq G_1/M_1 \times G_2/M_2;$$

since M' is maximal, we must have $\pi_1[M'] = G_1/M_1$. Similarly, we can prove $\pi_2[M'] = G_2/M_2$.

Then, define a map $f : G/M_1 \to G/M_2$ by $f(g) = h$, where $(g, h) \in M'$. Note that this is well-defined; for any $g \in G/M_1$, $(g, h) \in M'$ for some h (by the fact that $\pi_1[M'] = G_1$), and this h is unique (if $(g, h') \in M'$, then $(g, h')(g, h)^{-1} = (1, h'h^{-1}) \in M'$ forces $h' = h$ by the fact proven two paragraphs ago). Furthermore, $f(g_1g_2) = f(g_1)f(g_2)$, as if $(g_1, h_1), (g_2, h_2) \in M'$, then $(g_1g_2, h_1h_2) \in M'$. So f is a homomorphism. We can define an analogous homomorphism $f' : G/M_2 \to G/M_1$ by $f(h) = g$, where $(g, h) \in M'$. It is clear that f and f' are inverses. Therefore, f is an isomorphism between G/M_1 and G/M_2. If we let $F : G_1/M_1 \times G_2/M_2 \to (G_1/M_1)^2$ be the isomorphism given by $F(g, h) = (g, f^{-1}(h))$, we get $F[M'] = \{(g, g) : g \in G_1/M_1\}$.

Then, we compute $(G_1/M_1)^2/F[M'] \cong G_1/M_1$. Since $F[M']$ is maximal normal, this quotient must be simple, so G_1/M_1 (and hence G_2/M_2) must be simple. Therefore, M_1 and M_2 are maximal normal subgroups of G_1 and G_2, respectively, proving the lemma. We remark that in this case, G_1/M_1 must be cyclic of prime order, as the diagonal subgroup $\{(g, g) : g \in G\}$ is normal in $G \times G$ if and only if G is abelian. □

Proof of Proposition 2.15. Let N_i be the intersection of all maximal normal subgroups of G_i, and N be the intersection of all maximal normal subgroups of $G_1 \times G_2$. If M_1 is maximal normal in G_1, then $M_1 \times G_2$ is maximal normal in $G_1 \times G_2$, and similarly for $G_1 \times M_2$ if M_2 is maximal normal in G_2. Therefore,

$$N \leq \bigcap_{M_1} (M_1 \times G_2) \cap \bigcap_{M_2} (G_1 \times M_2) = (N_1 \times G_2) \cap (G_1 \times N_2) = N_1 \times N_2.$$

Conversely, by Lemma 2.16, if M is any maximal normal subgroup then $M_1 \times M_2 \leq M$ where each M_i is either equal to G_i or maximal normal in G_i. Therefore,

$$N_1 \times N_2 \leq M_1 \times M_2 \leq M.$$

Since N is the intersection of all such M, we get $N_1 \times N_2 \leq N$, proving equality of these two subgroups. Therefore:

$$\mathrm{CoSoc}(G_1 \times G_2) = \frac{G_1 \times G_2}{N_1 \times N_2} \cong \frac{G_1}{N_1} \times \frac{G_2}{N_2} = \mathrm{CoSoc}(G_1) \times \mathrm{CoSoc}(G_2). \qquad \square$$

Finally, we can prove the main theorem of this section, that two sequences are relatively prime if and only if they are relatively prime when after we pass to cosocles:

Theorem 2.17 (Gaschütz)**.** *Let G and H be finite groups, and let $s \in \Gamma_n(G)$ and $t \in \Gamma_n(G)$ be generating sequences. Let $\overline{s}$ and $\overline{t}$ be the projections of s and t to generating sequences of* $\mathrm{CoSoc}(G)$ *and* $\mathrm{CoSoc}(H)$*. Then s and t are relatively prime if and only if $\overline{s}$ and $\overline{t}$ are relatively prime.*

Proof. If s and t are relatively prime, then (s,t) generates $G \times H$ and therefore $(\overline{s},\overline{t})$ generates $\mathrm{CoSoc}(G) \times \mathrm{CoSoc}(H) \cong \mathrm{CoSoc}(G \times H)$, which is a quotient of $G \times H$.

So, suppose $\overline{s}$ and $\overline{t}$ are relatively prime, so $(\overline{s},\overline{t})$ generates $\mathrm{CoSoc}(G) \times \mathrm{CoSoc}(H) \cong \mathrm{CoSoc}(G \times H)$. In particular, we have a surjective homomorphism $\theta : F_n \to \mathrm{CoSoc}(G \times H)$ taking $(x_1, \ldots, x_n)$ to $(\overline{s},\overline{t})$. Let $\theta_1 : F_n \to \mathrm{CoSoc}(G)$ and $\theta_2 : F_n \to \mathrm{CoSoc}(H)$ be the surjective homomorphisms taking $(x_1, \ldots, x_n)$ to $\overline{s}$ and $\overline{t}$, respectively; thus we have $\theta = (\theta_1, \theta_2)$. Let $K_1 = \ker\theta_1$ and $K_2 = \ker\theta_2$. By Lemma 2.4, $\ker\theta = K_1 \cap K_2$ and $K_1K_2 = F_n$ (as θ is surjective).

Similarly, let $\pi : F_n \to G \times H$ be the homomorphism taking $(x_1, \ldots, x_n)$ to (s,t). Let $\pi_1 : F_n \to G$ and $\pi_2 : F_n \to H$ be the coordinate projections, so $\pi = (\pi_1, \pi_2)$, and let $L_1 = \ker\pi_1$ and $L_2 = \ker\pi_2$. By Lemma 2.4, $\ker\pi = L_1 \cap L_2$. Moreover, if we can show $L_1L_2 = F_n$ then that lemma implies that π is surjective and hence s and t are relatively prime.

Now, for $i = 1, 2$, let N_i be the intersection of all maximal normal subgroups in G_i, so θ_i is π_i composed with the canonical projection $p : G_i \to G_i/N_i$. We claim that $\pi_i[K_i] = N_i$. To see this, note that if $x \in K_i$, then $\theta_i(x) = p(\pi_i(x)) = 1$, which means $\pi_i(x) \in N_i$ by definition of the projection p, and hence $\pi_i[K_i] \subseteq N_i$. Conversely, if $n \in N$, then $n = \pi_i(y)$ for some $x \in F_n$ by surjectivity, and by definition we have

$$1 = p(n) = p(\pi_i(y)) = \theta_i(y),$$

so $y \in K_i$ and hence $N_i \subseteq \pi_i[K_i]$.

Since we have established $K_1K_2 = G$, we have

$$N_1\pi_1[K_2] = \pi_1[K_1]\pi_1[K_2] = \pi_1[K_1K_2] = \pi_1[F_n] = G_1.$$

As $K_2 \trianglelefteq F_n$ and π_1 is a surjective homomorphism, $\pi_1[K_2] \trianglelefteq G_1$. Then, we can apply Lemma 2.13, which gives $\pi_1[K_2] = G_1$.

Next, consider the subgroup $L_1K_2 \leq F_n$. If $k_1 \in K_1$ then there exists $k_2 \in K_2$ with $\pi_1(k_1) = \pi_1(k_2)$ (as we just showed $\pi_1[K_2] = G_1$). This means $k_1k_2^{-1} = \ell_1 \in \ker\pi_1 = L_1$, so $k_1 = \ell_1k_2 \in L_1K_2$ and hence $K_1 \leq L_1K_2$. Since $K_2 \leq L_1K_2$ trivially, we have $F_n = K_1K_2 \leq L_1K_2$, so $L_1K_2 = F_n$.

We can now re-apply the argument of the previous two paragraphs. Since $L_1K_2 = G$, we have

$$\pi_2[L_1]N_1 = \pi_2[L_1]\pi_2[K_2] = \pi_2[L_1K_2] = \pi_2[F_n] = G_2.$$

Lemma 2.13 gives $\pi_2[L_1] = G_2$. Then, if $k_2 \in K_2$, there is $\ell_1 \in L_1$ with $\pi_2(\ell_1) = \pi_2(k_2)$. This means $\ell_1^{-1}k_2 = \ell_2 \in L_2$, so $k_2 = \ell_1\ell_2 \in L_1L_2$. Therefore $K_2 \leq L_1L_2$, and hence $F_n = L_1K_2 \leq L_1L_2$. This proves $L_1L_2 = F_n$, and hence that s and t are relatively prime. □

Corollary 2.18. *Let G and H be finite groups. We have*

$$r(G \times H) = \max\{r(G), r(H), r(\mathrm{CoSoc}(G \times H))\}.$$

If $n \geq r(G \times H)$, then

$$\varphi_n(G \times H) = [\![G : \mathrm{CoSoc}(G)]\!]_n [\![H : \mathrm{CoSoc}(H)]\!]_n \cdot \varphi_n(\mathrm{CoSoc}(G \times H))$$

Proof. Since G, H, and $\mathrm{CoSoc}(G \times H)$ are all quotient groups of $G \times H$, it is clear that we have

$$r(G \times H) \geq m = \max\{r(G), r(H), r(\mathrm{CoSoc}(G \times H))\}.$$

On the other hand, since $m \geq r(\mathrm{CoSoc}(G \times H))$, then there is a generating sequence $(\overline{s},\overline{t})$ of $\mathrm{CoSoc}(G) \times \mathrm{CoSoc}(H) \cong \mathrm{CoSoc}(G \times H)$. Moreover, since $m \geq r(G)$, Gaschütz's Lemma (Corollary 1.21) implies that $\overline{s}$ has a lift s to a generating sequence of G. Similarly, since $m \geq r(H)$, $\overline{t}$ has a lift s to a generating sequence of G. Then Theorem 2.17 implies (s,t) generates $G \times H$, so $r(G \times H) \leq m$.

Now, suppose $n \geq r(G \times H)$, and consider $\Gamma_n(G \times H)$. We can partition this set into subsets $\Gamma_{(\overline{s},\overline{t})}$ consisting of all sequences $(s,t) \in \Gamma_n(G \times H)$ that project to a specific sequence $(\overline{s},\overline{t})$ in $\Gamma_n(\mathrm{CoSoc}(G \times H))$. By definition of the lifting index $[\![G : \mathrm{CoSoc}(G)]\!]_n$, there are this many sequences $s \in \Gamma_n(G)$ that project to $\overline{s}$. Similarly, there are $[\![H : \mathrm{CoSoc}(H)]\!]_n$ sequences in $\Gamma_n(H)$ that project

to $\bar{t}$. Therefore, there are $[\![G : \mathrm{CoSoc}(G)]\!]_n [\![H : \mathrm{CoSoc}(H)]\!]_n$ pairs $(s,t) \in \Gamma_n(G) \times \Gamma_n(H)$ that project onto $(\bar{s}, \bar{t})$. By Theorem 2.17, each of these sequences generates $G \times H$, so this set of sequences equals $\Gamma_{(s,t)}$. Therefore,

$$\varphi_n(G \times H) = \sum_{(\bar{s},\bar{t})} |\Gamma_{(s,t)}| = \sum_{(\bar{s},\bar{t})} [\![G : \mathrm{CoSoc}(G)]\!]_n [\![H : \mathrm{CoSoc}(H)]\!]_n$$

$$= [\![G : \mathrm{CoSoc}(G)]\!]_n [\![H : \mathrm{CoSoc}(H)]\!]_n \cdot \varphi_n(\mathrm{CoSoc}(G \times H)). \qquad \square$$

We remark that two groups G and H are relatively prime if and only if $\mathrm{CoSoc}(G)$ and $\mathrm{CoSoc}(H)$ are relatively prime. (Moreover, the two cosocles are relatively prime if and only if they share no common simple group in their decomposition as a direct product of simple groups). This means that if G and H are relatively prime, if $s \in \Gamma_n(G)$ and $t \in \Gamma_n(H)$ then $(\bar{s}, \bar{t})$ automatically generates $\mathrm{CoSoc}(G \times H)$, so the above results reduce to those of the previous section.

2.3 Direct Products of the Same Simple Group

Gaschütz's Theorem 2.17 reduces the question of understanding the generating sequences of $G \times H$ to that of understanding the generating sequences of G, of H, and of $\mathrm{CoSoc}(G \times H)$. By Proposition 2.14, the latter group is a direct product of finite simple groups, so we can write

$$\mathrm{CoSoc}(G \times H) \cong S_1^{n_1} \times \cdots \times S_k^{n_k},$$

where the S_i are distinct finite simple groups. Moreover, if $i \neq j$ (so S_i and S_j are distinct simple groups), $S_i^{n_i}$ and $S_j^{n_j}$ are relatively prime groups. Therefore, the question of understanding generating sequences of $\mathrm{CoSoc}(G \times H)$ reduces to understanding generating sequences of S^n, a direct product of some number of copies of a single finite simple group.

It turns out that there is much we can prove about such groups. In this section, we focus on determining $r(S^n)$ for finite simple groups n. First, we prove a more general result:

Proposition 2.19. *Let G be any nontrivial finite group. If $r(G^n) = m$, then $n \leq h_m(G)$, the reduced Eulerian function $\varphi_m(G)/|\operatorname{Aut}(G)|$.*

Proof. If $r(G^n) = m$, there is a surjective homomorphism $\pi : F_m \to G^n$; projecting coordinate-wise, we get surjective homomorphisms $\pi_i : F_m \to G$ for $1 \leq i \leq n$. Let K_i denote $\ker \pi_i$. Then, the kernels K_i are pairwise distinct. If not, then we would have $K_i = K_j$ for $i \neq j$, which would imply $K_i K_j = K_i \neq F_m$ and hence (by the Chinese Remainder Theorem for Groups, Lemma 2.4) that the product map (π_i, π_j) is not surjective. This would contradict the assumption that the original map π is surjective.

So, there are at least n distinct kernels of surjective homomorphisms $F_m \to G$. Now, any such homomorphism is π_s for some generating sequence $s \in \Gamma_m(G)$. Moreover, by Proposition 1.13, $\ker \pi_s = \ker \pi_t$ if and only if s and t are equivalent under the action of $\operatorname{Aut}(G)$, so the number of distinct kernels is bounded above by the number of orbits of this action. By Proposition 1.10, $\operatorname{Aut}(G)$ acts freely, so there are at most $\varphi_m(G)/|\operatorname{Aut}(G)| = h_m(G)$ distinct kernels. Therefore $n \leq h_m(G)$. $\square$

The main theorem of this section is that if S is a nonabelian simple group then this upper bound is exact; we have $r(S^{h_n(S)}) = n$. First, though, we deal with the case of abelian simple groups; their behavior turns out to be understandable through linear algebra.

Proposition 2.20. *Letting Z_p^k be the direct product of k copies of a cyclic group of prime order Z_p, we have $r(Z_p^k) = \overline{r}(Z_p^k) = k$.*

Proof. Note that Z_p^k is a k-dimensional vector space over the finite field $\mathbb{F}_p$, and generating sequences of Z_p^k are exactly sequences that correspond to spanning sets of the vector space. Therefore, any spanning sequence must be at least k elements long by dimensionality. Moreover, irredundant sequences correspond exactly to minimal spanning sets, which are bases for the vector space; hence every irredundant generating sequence must be of length k. $\square$

We can now focus on nonabelian finite simple groups. The first step is the following lemma:

Lemma 2.21. *Let S be a nonabelian finite simple group. Then, if N is a normal subgroup of S^k, we have $N = N_1 \times \cdots \times N_k$ where each N_i is either 1 or S. In particular, there are exactly k distinct normal subgroups of S^k that have index $|S|$.*

Proof. Let $N \trianglelefteq S^k$. Define $N_i = 1$ if $\pi_i[N] = 1$, and $N_i = S$ otherwise; we claim $N = N_1 \times \cdots \times N_k$ for this list of N_i. We first show that $N_1 \times 1 \times \cdots \times 1 \leq N$. If $N_1 = 1$ then this is trivial, so suppose $N_1 = S$.

Since $N_1 = S$, there must be $n = (n_1, \ldots, n_k) \in N$ with $n_1 \neq 1$. As S is a nonabelian finite simple group, the center $Z(S)$ is trivial, so there is an element $s \in S$ that does not commute with n_1. Then, by normality of N, we have

$$\Big((s, 1, \ldots, 1)^{-1}(n_1, \ldots, n_k)^{-1}(s, 1 \ldots, 1)\Big)(n_1, \ldots, n_k) = ([s, n_1], 1, \ldots, 1) \in N.$$

Since s and n_1 do not commute, $[s, n_1]$ is nontrivial in S. Moreover, by simplicity of S, the normal closure of the element $[s, n_1]$ is S, so every $s' \in S$ can be written as a product of conjugates of $[s, n_1]$ and its inverse. By normality of N, $(s', 1, \ldots, 1) \in N$ where s' is any conjugate of $[s, n_1]$. Since N is a subgroup, $(s', 1, \ldots, 1) \in N$ where s' is any product of such conjugates or their inverses. Therefore $N_1 \times 1 \times \cdots \times 1 \leq N$, as desired.

An identical argument shows that $1 \times \cdots \times 1 \times N_i \times 1 \times \cdots \times 1 \leq N$ for each i. Therefore, the product $N_1 \times \cdots \times N_k$ of these subgroups is contained in N. On the other hand, it is clear that $N \leq N_1 \times \cdots \times N_k$. This proves equality, and hence that every $N \trianglelefteq S^k$ is of the desired form. Moreover, such a normal subgroup $N_1 \times \cdots \times N_k$ is of index $|S|$ if and only if exactly one N_i is equal to 1, with the other N_j equal to S. Since there are k choices of the coordinate i, there are k normal subgroups of index N in S^n. □

We can then prove the main result.

Theorem 2.22. *Let S be a nonabelian simple group. Then $r(S^{h_n(S)}) = n$. More generally, if $h_{n-1}(S) < k \leq h_n(S)$ then $r(S^k) = n$.*

Proof. By Propositions 1.10 and 1.13, there are exactly $m = h_n(S)$ distinct kernels of maps $\pi_s : F_n \to S$. Enumerate these kernels by $K_1, \ldots, K_m$. For each $i \leq m$, define

$$L_i = K_1 \cap \cdots \cap K_i.$$

We want to show that $F_n/L_i \cong S^i$. We will prove this by induction. The base case of $i = 1$ is trivial, as $F_n/L_1 = F_n/K_1 \cong S$ by definition.

Now, suppose that we know $F_m/L_{i-1} \cong S^{i-1}$. Note that if $L_{i-1} \leq K_j$, the subgroup K_j/L_{i-1} of F_m/L_{i-1} has index $|S|$:

$$[\tfrac{F_m}{L_{i-1}} : \tfrac{K_j}{L_{i-1}}] = \frac{[F_m : L_{i-1}]}{[K_j : L_{i-1}]} = [F_m : K_j] = |S|$$

Thus, the subgroups $K_1/L_{i-1}, \ldots, K_{i-1}/L_{i-1}$ are $i - 1$ distinct normal subgroups of index $|S|$ in $F_m/L_{i-1} \cong S^{i-1}$. By Lemma 2.21, these are all such subgroups of F_m/L_{i-1}. Therefore, if $L_{i-1} \leq K_i$, K_i/L_{i-1} would equal some K_j/L_{i-1} with $j < i$, and hence $K_j = K_i$, which is a contradiction. So we must have $L_{i-1} \nleq K_i$. Then $L_{i-1}K_i$ is a normal subgroup of F_n that properly contains K_i, which means $L_{i-1}K_i/K_i$ is a nontrivial normal subgroup of $F_m/K_i \cong S$. Since S is simple, this means $L_{i-1}K_i/K_i = F_m/K_i$ and in particular $L_{i-1}K_i = F_n$.

Now we apply the Chinese Remainder Theorem for groups, Lemma 2.4. Let $h_1 : F_m \to F_m/L_{i-1} \cong S^{i-1}$ and $h_2 : F_m \to F_m/K_i \cong S$ be the canonical projections, with kernels L_{i-1} and K_i, respectively. Since we just showed $L_{i-1}K_i = F_n$, the lemma gives that the product map $h : F_m \to S^{i-1} \times S = S^i$ is surjective and has kernel $L_{i-1} \cap K_i = L_i$. By the first isomorphism theorem, we get $F_n/L_i \cong S^i$, as desired. This finishes the inductive step.

By induction, we get $F_n/L_m \cong S^m$ (where we defined $m = h_n(S)$), and hence $r(S^{h_n(S)}) \leq m$. If k satisfies $h_{n-1}(S) < k \leq h_n(S)$, then S^k is a quotient of $S^{h_n(S)}$, so $r(S^k) \leq r(S^{h_n(S)}) \leq n$. On the other hand, if $r(S^k) \leq n - 1$, then Proposition 2.19 would imply $k \leq h_{n-1}(S)$, a contradiction. Therefore we must have $r(S^k) = n$ in such a situation, and in particular $r(S^{h_n(S)}) = n$. □

3 The Product Replacement Graph

Fix a finite group G and a length n. For any integers $i \neq j$ between 1 and n, define a function $\ell_{ij}^{+} : \Gamma_n(G) \to \Gamma_n(G)$ by

$$\ell_{ij}^{+}(g_1, \ldots, g_n) = (g_1, \ldots, g_{i-1}, g_j g_i, g_{i+1}, \ldots, g_n).$$

Thus, ℓ_{ij}^{+} makes a "local substitution" (by left multiplication of g_j^{+1} in the i-th spot) in a generating sequence by left multiplication, replacing the i-th entry g_i by $g_j g_i$. To see that the resulting sequence still generates, note that it contains $g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_n$ and that we can get g_i from it by multiplying g_j^{-1} and $g_j g_i$.

Moreover, note that ℓ_{ij}^{+} is an invertible function; its inverse can be given by

$$\ell_{ij}^{-}(g_1, \ldots, g_n) = (g_1, \ldots, g_{i-1}, g_j^{-1} g_i, g_{i+1}, \ldots, g_n).$$

This is also called a local substitution (by left multiplication of g_j^{-1} in the i-th spot, in this case). Similarly, we can define functions r_{ij}^{+} and r_{ij}^{-} that are also permutations of $\Gamma_n(G)$ by using right multiplication instead of left multiplication:

$$r_{ij}^{\pm}(g_1, \ldots, g_n) = (g_1, \ldots, g_{i-1}, g_i g_j^{\pm 1}, g_{i+1}, \ldots, g_n).$$

Collectively, we call the functions $\ell_{ij}^{\pm}$ and $r_{ij}^{\pm}$ the "basic elementary operations" on $\Gamma_n(G)$. Note that our notation $\ell_{ij}^{\pm}$ and $r_{ij}^{\pm}$ suppresses the group G. If we need to specify the group, we use the notation $\ell_{ij,G}^{\pm}$ and $r_{ij,G}^{\pm}$.

The product replacement algorithm can then be described as starting with a generating sequence $(g_1, \ldots, g_n)$, applying some random sequence of basic elementary operations, and then returning the first element of the final sequence $(g_1', \ldots, g_n')$. For this algorithm to work properly, it is important that it is possible to get from any generating sequence to any other by this process. To formalize this problem so we can study it, we define:

Definition 3.1. The n-th "product replacement graph" is a graph with $\Gamma_n(G)$ as its set of vertices, and with an edge between two vertices s, s' if there is some basic elementary operation taking s to s'.

Note that the inverse of a basic elementary operation is a basic elementary operation, so there a basic elementary operation taking s to s' if and only if there is one taking s' to s. In general, note that it is possible to get from s to s' through a sequence of basic elementary operations if and only if there is a path between them in this graph. Thus, the desired condition of being able to get from any s to s' is equivalent to the connectedness of the product replacement graph.

To think about this problem in a group-theoretic way, we will want to use an equivalent formulation of the problem in terms of group actions. In particular, for a fixed group G and integer n, we define:

Definition 3.2. Define the n-th group of "elementary operations" on G, denoted $E_n(G)$, as the subgroup of the symmetric group on $\Gamma_n(G)$ generated by the basic elementary operations (so all permutations that can be written as a finite composition of basic elementary operations). Elements of this group are called elementary operations on G.

By definition, two generating sequences s and s' are equvialent under a sequence of basic elementary operations if and only if there is an elementary operation taking s to s'. If we let $E_n(G)$ act on $\Gamma_n(G)$ as permutations, this says that the product replacement graph is connected if and only if $E_n(G)$ acts transitively on $\Gamma_n(G)$. So, it is sufficient to study this group action. In this vein, we define:

Definition 3.3. Let $\mathcal{O}(G, E_n)$ be the set of orbits of $\Gamma_n(G)$ under $E_n(G)$, and let $c(G, E_n)$ be the number of orbits of $\Gamma_n(G)$ under $E_n(G)$.

By definition, $E_n(G)$ acts transitively if and only if $c(G, E_n) = 1$. Also, note that the orbits of the action of $E_n(G)$ are exactly equal to the components of the product replacement graph. Thus $\mathcal{O}(G, E_n)$ and $c(G, E_n)$ are equivalently the set and number of components of this graph, respectively.

We are most interested in finding conditions that guarantee $c(G, E_n) = 1$, but it also turns out to be interesting to work with cases where $c(G, E_n) > 1$ as well. It is also useful to consider the action of other subgroups of the symmetric group on $\Gamma_n(G)$ besides $E_n(G)$. The ones we will use are:

Definition 3.4. Define the n-th group of "left operations" $L_n(G)$ as the subgroup of $\mathcal{S}(\Gamma_n(G))$ generated by the operations ℓ_{ij}^+ and ℓ_{ij}^-. Define the n-th group of "right operations" $R_n(G)$ as the subgroup generated by r_{ij}^+ and r_{ij}^-.

In analogy to Definition 3.3, we define $\mathcal{O}(G, L_n)$ and $\mathcal{O}(G, R_n)$ as the set of orbits of $\Gamma_n(G)$ under the actions of $L_n(G)$ and $R_n(G)$, respectively, and we similarly define $c(G, L_n)$ and $c(G, R_n)$ as the number of orbits. It is clear that any orbit under $L_n(G)$ or $R_n(G)$ is contained in an orbit under $E_n(G)$, and hence that $c(G, L_n), c(G, R_n) \leq c(G, E_n)$.

We remark that there are various other groups $U_n(G) \leq \mathcal{S}(\Gamma_n(G))$ that are of interest, and have been studied in the literature. We can add permutations of $\Gamma_n(G)$ induced by some $\sigma \in S_n$:

$$(g_1, \ldots, g_n) \mapsto (g_{\sigma(1)}, \ldots, g_{\sigma(n)}).$$

We can also use "inversion operations", induced by inversion of some particular coordinate:

$$(g_1, \ldots, g_n) \mapsto (g_1, \ldots, g_i^{-1}, \ldots, g_n).$$

For instance, we can take $U_n(G)$ to be generated by all left operations, right operations, and inversion operations. If we let G be a free group F_n (note that all of the relevant definitions work just as well for infinite groups), connectivity of $U_n(F_n)$ is the Andrews-Curtis conjecture from topology [AC65]. The "finitary Andrews-Curtis conjecture", on connectivity of $U_n(G)$ when G is a finite group, has also recieved attention [BLM05].

3.1 Basic Properties

First, note that left operations and right operations are "dual." This allows us to show that the number of orbits under left operations and under right operations are the same, and hence that we only need to work with one (we usually use L_n):

Proposition 3.5. *For any finite group G and any n, we have $c(G, L_n) = c(G, R_n)$.*

Proof. Recall that for any group G (with binary operation $\cdot$), we can define its "opposite group" G^{op}, with the same underlying set as G, and with binary operation $*$ given by $g * h = h \cdot g$. Moreover, G and G^{op} are canonically isomorphic by the map $g \mapsto g^{-1}$.

The underlying sets of G and G^{op} are equal by definition; we claim moreover that $\Gamma_n(G)$ and $\Gamma_n(G^{op})$ are equal as sets as well. To see this, note that if $(g_1, \ldots, g_n)$ generates G, then by the isomorphism above, $(g_1^{-1}, \ldots, g_n^{-1})$ generates G^{op}. Moreover, since replacing an element by its inverse in a sequence doesn't change what subgroup it generates, we get that $(g_1, \ldots, g_n)$ is a generating sequence of G^{op}. This means $\Gamma_n(G^{op}) \subseteq \Gamma_n(G)$; the reverse inclusion can be proved in an identical manner.

Then, we can compute

$$\ell_{ij,G}^{\pm}(g_1, \ldots, g_n) = (g_1, \ldots, g_j^{\pm 1} \cdot g_i, \ldots, g_n) = (g_1, \ldots, g_i * g_j^{\pm 1}, \ldots, g_n) = r_{ij,G^{op}}^{\pm}(g_1, \ldots, g_n),$$

So, basic left operations for G are equal to the basic right operations for G^{op}. Thus, $L_n(G)$ and $R_n(G^{op})$ are equal as subsets of $\mathcal{S}(\Gamma_n(G))$. Therefore, $c(G, L_n) = c(G^{op}, R_n)$. Since G and G^{op} are isomorphic, we get $c(G, L_n) = c(G, R_n)$, as desired. □

Now, we can relate the number of orbits for L_n or E_n on a group G to the number of orbits for L_n or E_n on a homomorphic image of G:

Proposition 3.6. *Let G be a group and H a homomorphic image of G by a surjective homomorphism $h : G \to H$, and let $\overline{h} : \Gamma_n(G) \to \Gamma_n(H)$ be the induced surjection defined in Definition 1.18. Then, $\overline{h}$ carries orbits of H under $L_n(H)$ or $E_n(H)$ into orbits of G under $L_n(G)$ or $E_n(G)$, respectively, so $c(H, L_n) \leq c(G, L_n)$ and $c(H, E_n) \leq c(G, E_n)$.*

Proof. Using the fact that h is a homomorphism, we can check that $\overline{h} \circ \ell^{\pm}_{ij,G} = \ell^{\pm}_{ij,H} \circ \overline{h}$. Therefore, if $s, s' \in \Gamma_n(G)$ are related by some $\ell^{\pm}_{ij,G}$, then $\overline{h}(s)$ and $\overline{h}(s')$ are related by the corresponding $\ell^{\pm}_{ij,H}$. By taking compositions of such elements, we get that if s and s' are in the same orbit under $L_n(G)$, then $\overline{h}(s)$ and $\overline{h}(s')$ are in the same orbit under $L_n(H)$. So, $\overline{h}$ maps orbits under $L_n(G)$ into orbits under $L_n(H)$. By surjectivity of $\overline{h}$, each orbit of $L_n(H)$ is mapped to, so $c(H, L_n) \leq c(G, L_n)$. An identical argument works for E_n. □

We can also compare the number of orbits for L_n or E_n on groups G and H to the number of orbits on $G \times H$, in the case that the orders $|G|$ and $|H|$ are coprime integers:

Proposition 3.7. *If $|G|$ and $|H|$ are coprime, then $c(G \times H, L_n) = c(G, L_n) \cdot c(H, L_n)$ and $c(G \times H, E_n) = c(G, E_n) \cdot c(H, E_n)$.*

Proof. By Proposition 2.9, G and H are relatively prime groups, so $\Gamma_n(G \times H)$ can be naturally identified with $\Gamma_n(G) \times \Gamma_n(H)$. We claim that the orbits under $L_n(G \times H)$ are the products of orbits under $L_n(G)$ and $L_n(H)$ under this correspondence; this will establish $c(G \times H, L_n) = c(G, L_n) \cdot c(H, L_n)$.

First note that if two sequences (s, t) and (s', t') are equivalent under $L_n(G \times H)$, then there is a sequence of elementary left operations taking (s, t) to (s', t'); looking coordinate-wise, we get a sequence of elementary left operations taking s to s' and a sequence of left operations taking t to t'. Therefore, the orbit of (s, t) is contained in the product of the orbit of s and the orbit of t.

For the converse, first suppose that $\ell^{\pm}_{ij,G}$ takes s to s', and fix $t \in \Gamma_n(H)$. We claim that (s, t) and (s', t) are in the same orbit of $L_n(G \times H)$. To see this, note that since $|G|$ and $|H|$ are coprime, there is some a with $a|H| \equiv 1 \pmod{|G|}$. We can then compute

$$(\ell^{\pm}_{ij,G\times H})^{a|H|}(s, t) = \Big((\ell^{\pm}_{ij,G})^{a|H|}(s), (\ell^{\pm}_{ij,H})^{a|H|}(t)\Big).$$

Now, note that $(\ell^{\pm}_{ij,G})^{|G|}$ is the identity function in $\mathcal{S}(\Gamma_n(G))$, as it takes $(g_1, \ldots, g_n)$ to

$$(g_1, \ldots, g_j^{|G|} g_i, \ldots, g_n) = (g_1, \ldots, g_i, \ldots, g_n)$$

because $|g_j|$ divides $|G|$. Therefore, $(\ell^{\pm}_{ij,H})^{a|H|}(t) = t$, and also $(\ell^{\pm}_{ij,G})^{a|H|}(s) = \ell^{\pm}_{ij,G}(s) = s'$ because $a|H| \equiv 1 \pmod{|G|}$. Thus $\ell^{\pm}_{ij,G\times H}(s, t) = (s', t)$.

So, if s and s' are related by a basic left operation, (s, t) and (s', t) are in the same orbit of $L_n(G \times H)$. Since any left operation is a composition of basic left operations, we get that if s and s' are in the same orbit of $L_n(G)$ then (s, t) and (s', t) are in the same orbit of $L_n(G \times H)$. By a similar argument, we can see that if t, t' are in the same orbit of $L_n(H)$, then (s', t) and (s', t') are in the same orbit of $L_n(G \times H)$, and hence (s, t) and (s', t') are in the same orbit.

This proves that the orbit of (s, t) in $L_n(G \times H)$ is equal to the product of the orbits of s and t in $L_n(G)$ and $L_n(H)$, respectively, which proves the desired equation $c(G \times H, L_n) = c(G, L_n) \cdot c(H, L_n)$. An identical argument shows that the corresponding equation holds for E_n. □

3.2 Results of Dunwoody and Diaconis-Graham

An important first result concerning connectivity of a product replacement graph was due to Dunwoody [Dun70]:

Theorem 3.8 (Dunwoody)**.** *If G is a finite solvable group and $n > r(G)$, then $c(G, E_n) = 1$ (so the product replacement graph is connected).*

Before proving this, recall that a "chief series" of a group G is a series of subgroups

$$1 = G_0 < G_1 < \cdots < G_c = G,$$

where each G_i is a normal subgroup of G that is maximal (there is no $N \trianglelefteq G$ with $G_i < N < G_{i+1}$). It is easy to see that any finite group has a chief series (by starting with $1 < G$ and then adding normal subgroups until no more fit). We also need some lemmas.

Lemma 3.9. *Let $(a + m\mathbb{Z}, b + m\mathbb{Z})$ be a pair of elements in $\mathbb{Z}/m\mathbb{Z}$ (from now on, we will suppress the coset $m\mathbb{Z}$ and write (a, b)). Then a sequence of basic elementary operations takes (a, b) to $(d, 0)$, where d is the GCD of a, b, and m.*

Proof. Let d' be the GCD of a and b, so d is the GCD of d' and m. Note that $a = a'd'$ and $b = b'd'$ for coprime integers a', b'. By Dirichlet's Theorem on primes in arithmetic progressions, there are infinitely many primes of the form $ka' + b'$, so in particular we can pick such a prime p greater than m. By iterating the elementary operation $(\alpha, \beta) \to (\alpha, \beta \pm \alpha)$, we can get from (a, b) to $(a, ka + b) = (a'd', pd')$. Since p is a prime greater than m, p and m are coprime, so p is invertible modulo m. We can then use operations $(\alpha, \beta) \to (\alpha \pm \beta, \beta)$ to add $(1 - a')p^{-1}$ copies of pd' to the first entry $a'd'$, which gets our pair to (d', pd') (since we are working modulo m). We can then go to $(d', 0)$ by some more elementary operations. Finally, note that since d is the GCD of d' and m, $d = xd' + ym$ for some x, y, so $d \equiv xd' \pmod{m}$. So, we can go from $(d', 0)$ to $(d', xd') = (d', d)$, then to (d, d), and finally to $(d, 0)$, as desired. We remark that it is convenient but not necessary to invoke Dirichlet's theorem in this proof. A more elementary, but longer, proof can be given using the Euclidean algorithm. □

Lemma 3.10. *If $n > 1$, then $c(Z_m, L_n) = c(Z_m, E_n) = 1$ for a cyclic group Z_m.*

Proof. Note that since Z_m is abelian, L_n and E_n are equal. For convenience, we will view Z_m as $\mathbb{Z}/m\mathbb{Z}$, so we can work directly with integers. Note that $\gcd(a_1, \ldots, a_n)$ must be coprime to m, as otherwise all elements generated by $(a_1, \ldots, a_n)$ would be multiples of $\gcd(d_1, m)$. This means $\gcd(a_1, \ldots, a_n, m) = 1$.

So, let $(a_1, \ldots, a_n)$ be a generating sequence of Z_m, with $n > 1$. Let d_i denote the GCD of the elements $a_i, \ldots, a_n, m$. Applying the argument of Lemma 3.9 to the last two entries of the sequence, we get that $(a_1, \ldots, a_n)$ is equivalent to $(a_1, \ldots, a_{n-2}, d_{n-1}, 0)$ by a sequence of basic elementary operations. Iterating, we get that $(a_1, \ldots, a_n)$ is equivalent to $(a_1, \ldots, a_{i-1}, d_i, 0, \ldots, 0)$, and finally to $(d_1, 0, \ldots, 0) = (1, 0, \ldots, 0)$. Thus any two sequences are equivalent to $(1, 0, \ldots, 0)$ and hence to each other. □

Lemma 3.11. *Let $1 = G_0 < G_1 < \cdots < G_c = G$ be a chief series of a solvable group G and $(m, g_1, \ldots, g_{n-1})$ be a generating sequence of G so that m is a nontrivial element of G_1. Then, for any $g \in G$, this sequence is equivalent to $(gmg^{-1}, g_1, \ldots, g_{n-1})$ under $E_n(G)$.*

Proof. First, note that since G is a solvable group and G_1 is a minimal normal subgroup of G, we know that G_1 is abelian (see, for instance, [Rot95] Theorem 5.24). Since $(m, g_1, \ldots, g_{n-1})$ is a generating sequence of G, if we set $H = \langle g_1, \ldots, g_{n-1} \rangle$ then $\langle m, H \rangle = G$ and hence $\langle G_1, H \rangle = G$; since G_1 is normal, this implies $HG_1 = G$. Thus, any $g \in G$ can be written as hm' for $h \in H$ and $m' \in M$. Since G_1 is abelian, we have

$$gmg^{-1} = hm'm(m')^{-1}h^{-1} = hmh^{-1}.$$

Since $h \in \langle g_1, \ldots, g_{n-1} \rangle$, we can use left operations to build h on the left side of the first entry of $(m, g_1, \ldots, g_{n-1})$, and h^{-1} on the right side; thus we get that this sequence is equivalent under $E_n(G)$ to

$$(hmh^{-1}, g_1, \ldots, g_{m-1}) = (gmg^{-1}, g_1, \ldots, g_{m-1}),$$

as desired. □

Proof of Theorem 3.8. We prove this theorem by induction on the length c of a chief series for G. For the base case of $c = 1$, the chief series is $1 < G_1$, which means G_1 is simple. The only solvable finite simple groups are cyclic, and Lemma 3.10 shows that if $n > r(Z_m) = 1$ then $c(Z_m, E_n) = 1$.

So, assume that we know the desired statement holds for all groups with chief series of length less than some $c > 1$, and let G have a chief series $1 < G_1 < \cdots < G_c = G$ of length c. Fix some $n > r(G)$, and fix some generating sequence $(h_1, \ldots, h_{n-1})$ of G of length $n-1$. We will show that under E_n, every sequence in $\Gamma_n(G)$ is equivalent to $(1, h_1, \ldots, h_{n-1})$, which will prove $c(G, E_n) = 1$.

Start by noting that G/G_1 has a chief series of length $c-1$ (namely $1 < G_2/G_1 < \cdots < G_c/G_1$), so the inductive hypothesis applies for it. Therefore, if $(g_1, \ldots, g_n)$ is any generating sequence of G, $(g_1G_1, \ldots, g_nG_1)$ and $(G_1, h_1G_1, \ldots, h_{n-1}G_1)$ are both generating sequences of G/G_1, and by induction these sequences are equivalent by an sequence of basic elementary operations on G/G_1. Considering the same sequence of basic elementary operations in G, we get that $(g_1, \ldots, g_n)$ is equivalent to $(m, h_1m_1, \ldots, h_{n-1}m_{n-1})$ for elements $m, m_i \in G_1$. Moreover, we can assume without loss of generality that $m \neq 1$; otherwise, $(h_1m_1, \ldots, h_{n-1}m_{n-1})$ generates G, so we can pick some $m \neq 1$ and write it as a sequence of these elements and their inverses, and then use elementary operations to build it in the leftmost position in our sequence.

Therefore, it suffices to show that any sequence $s = (m, h_1m_1, \ldots, h_{n-1}m_{n-1})$ with $m, m_i \in G_1$ and $m \neq 1$ is equivalent to $(1, h_1, \ldots, h_{n-1})$. To see this, first note that since G_1 is the first term in a chief series, it is a minimal normal subgroup. Therefore, G_1 is generated by the set $\{gmg^{-1} : g \in G\}$ (as this set generates a normal subgroup contained in G_1). So, we can write each m_i as a product of elements gmg^{-1} (and their inverses). For some $m' \in G_1$, let $\ell(m', m)$ denote the shortest length of a product of elements $(gmg^{-1})^{\pm 1}$ that equals m'. Note that $\ell(m', m) = \ell(m', gmg^{-1})$, as conjugates of m are the same as conjugates of gmg^{-1}. Then, given a generating sequence $s = (m, h_1m_1, \ldots, h_{n-1}m_{n-1})$, let $\ell(s)$ denote $\ell(m_1, m) + \cdots + \ell(m_{n-1}, m)$; this is a measure of how far away our sequence is from the desired sequence $(1, h_1, \ldots, h_{n-1})$.

We will prove that each such s is equivalent to $(1, h_1, \ldots, h_{n-1})$ by induction on $\ell(s)$. If $\ell(s) = 0$, then $s = (m, h_1, \ldots, h_{n-1})$; since $(h_1, \ldots, h_{n-1})$ generates G, we can write m as a product of these elements, and use elementary operations to build m^{-1} on the left of m in the first position of the sequence. This proves s is equivalent to $(1, h_1, \ldots, h_{n-1})$, giving the base case of this induction.

For the inductive step, assume we have some s with $\ell(s) > 0$. Then some m_i is nonzero, and in particular can be written as $m_i'(gmg^{-1})^{\pm 1}$ where $\ell(m_i') = \ell(m_i) - 1$. By Lemma 3.11, s is equivalent to

$$(gmg^{-1}, h_1m_1, \ldots, h_{n-1}m_{n-1})$$

and hence to

$$s' = (gmg^{-1}, h_1m_1, \ldots, h_im_i', \ldots, h_{n-1}m_{n-1}).$$

We have $\ell(m_j, m) = \ell(m_j, gmg^{-1})$ by a comment above, and also

$$\ell(m_i', gmg^{-1}) = \ell(m_i', m) < \ell(m_i, m).$$

These imply $\ell(s') < \ell(s)$, and by induction s' (and hence s) is equivalent to $(1, h_1, \ldots, h_{n-1})$. This finishes the induction on $\ell(s')$, proving that all sequences in $\Gamma_n(G)$ are equivalent under $E_n(G)$. Thus, we are also done with our induction on the length of the chief series of G, proving that $c(G, E_n) = 1$ for $n > r(G)$ whenever G is solvable. □

This result leads to the following conjecture:

Conjecture 3.12 (Pak)**.** If G is any finite group and $n > r(G)$, then $c(G, E_n) = 1$.

The following special case of the conjecture is particularly important, and is more widely believed to be true.

Conjecture 3.13 (Wiegold)**.** If S is a finite simple group and $n > 2$, then $c(S, E_n) = 1$.

Note that this is simply the previous conjecture restricted to simple groups (using the fact that $r(S) = 2$ for finite simple groups).

There is not much progress towards proving these conjectures in general. We quote some results from the literature in this direction. In particular, we have the following collection of partial results (for alternating groups, projective special linear groups, and the Suzuki groups):

Theorem 3.14. *The n-th product replacement graph is connected for:*
(a) $G = \mathrm{PSL}(2, p)$ for $p \geq 5$ prime, and $n \geq 3$ (Gilman, [Gil77]).
(b) $G = \mathrm{PSL}(2, 2^m)$ for $m \geq 2$, and $n \geq 3$ (Evans, [Eva93]).
(c) $G = \mathrm{PSL}(2, p^m)$ for $p \geq 3$ prime and $m \geq 2$, and $n \geq 4$ (Garion, [Gar08]).
(d) $G = \mathrm{Sz}(2^{2m-1})$ for $m \geq 2$, and $n \geq 3$. (Evans, [Eva93]).
(e) $G = A_m$ for $6 \leq m \leq 11$, and $n \geq 3$ (David [Dav93], Cooperman and Pak [CP00]).

Note that none of these results or conjectures apply to the $n = r(G)$ case, and in fact that product replacement graph is in general not connected. It is easiest to see this for cyclic groups, for if $n = 1$ there are no basic elementary operations (they all need sequences of length 2 to define), so $L_1(G)$ and $E_1(G)$ are the trivial group, and $c(Z_m, E_1) = \varphi(m)$. The following theorem of Diaconis and Graham [DG99] extends this to compute $c(A, E_n)$ for $n = r(A)$ and A a finite abelian group. Recall that for any finite abelian group A has an "invariant factor decomposition", i.e. we have an isomorphism of A with a direct product of cyclic groups

$$A \cong Z_{m_1} \times \cdots \times Z_{m_n}$$

where $m_n > 1$ and m_i divides m_{i-1} for each $i > 1$.

Theorem 3.15 (Diaconis and Graham)**.** *Let A be a finite abelian group with invariant factor decomposition $Z_{m_1} \times \cdots \times Z_{m_n}$. Then $r(A) = n$, and $c(A, E_n) = c(A, L_n) = \varphi(m_n)$ (where φ is Euler's phi function).*

Proof. First note that $L_n(A) = E_n(A)$ since A is abelian, so we only need to worry about left operations. To see $r(A) = n$, note that as a product of n cyclic groups, $r(A) \leq n$ (as an application of Proposition 2.1). On the other hand, if p is a prime dividing m_n, then p divides each m_i by definition of the invariant factor decomposition. Then A has Z_p^n as a quotient. By Proposition 2.20, $r(Z_p^n) = n$, so $r(A) \geq n$. This forces equality.

So, we are interested in length n generating sequences $(a_1, \ldots, a_n)$ of $Z_{m_1} \times \cdots \times Z_{m_n}$. Each a_i is an n-tuple $(a_{i1}, \ldots, a_{in})$; thus we can express a generating sequence as a square matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix},$$

where the i-th row represents the element a_i of A, and the j-th column consists of elements of Z_{m_j}. Note that left operations consist of row operations on this matrix, in particular adding one row (or its negative) to another.

First, note that if we reduce each row modulo m_n, we get a $n \times n$ matrix with entries in $\mathbb{Z}/m_n\mathbb{Z}$. Since this is a commutative ring, there is a determinant function from $M_n(\mathbb{Z}/m_n\mathbb{Z})$ to $\mathbb{Z}/m_n\mathbb{Z}$ which is invariant under all row operations (so in particular the operations in $L_n(A)$). Now, if d generates $\mathbb{Z}/m_n\mathbb{Z}$, the matrix

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots \\ \vdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & d \end{bmatrix}$$

corresponds to a generating sequence of A, and moreover has determinant d. Since there are $\varphi(m_n)$ elements of $\mathbb{Z}/m_n\mathbb{Z}$ that generate, these matrices give us $\varphi(m_n)$ distinct orbits of $\Gamma_n(A)$ under $E_n(A)$. To prove that $c(A, E_n) = \varphi(m_n)$, we want to show that any generating sequence of A is equivalent to one of these matrices.

Let us start with a matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

From Lemma 3.9, we know any pair (a, b) of elements in some $\mathbb{Z}/m\mathbb{Z}$ is equivalent under elementary operations to $(d, 0)$, where $d = \gcd(a, b, m)$. Applying this to the first column of our matrix, we get

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \sim \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n-1,1} & a'_{n-1,2} & \cdots & a'_{n-1,n} \\ 0 & a'_{n,2} & \cdots & a'_{n,n} \end{bmatrix},$$

where $d_{n-1,1}$ is the GCD of $a_{n-1,1}$, $a_{n,1}$, and m_1. Iterating this process for each row of the first column, and noting that $\gcd(a_{1,1}, \ldots, a_{n,1}, m_1) = 1$ (since the projection of our generating sequence onto the first coordinate must generate) we get

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \sim \begin{bmatrix} 1 & a'_{1,2} & \cdots & a'_{1,n} \\ 0 & a'_{2,2} & \cdots & a'_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{n,2} & \cdots & a'_{n,n} \end{bmatrix},$$

for some set of entries $a'_{i,j}$. Now, we can apply the same process to the bottom $n-1$ columns to reduce the second column to $(a'_{1,2}, 1, 0, \ldots, 0)$, and this will not change the first column (since we're doing row operations on rows with their first entry zero). We get:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \sim \begin{bmatrix} 1 & a'_{1,2} & a'_{1,3} & \cdots & a'_{1,n} \\ 0 & d'_{2,2} & a'_{2,3} & \cdots & a'_{2,n} \\ 0 & 0 & a'_{3,3} & \cdots & a'_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a'_{n,3} & \cdots & a'_{n,n} \end{bmatrix},$$

where $d'_{2,2} = \gcd(a'_{2,2}, \ldots, a'_{n,2}, m_2)$. Then, note that since this is a generating sequence, $d'_{2,2}$ must equal 1 so that we can generate a sequence with 0 in the first column. We can continue to do this for each other column, and eventually get:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \sim \begin{bmatrix} 1 & b_{1,2} & \cdots & b_{1,n-1} & b_{1,n} \\ 0 & 1 & b_{2,3} & \cdots & b_{2,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & b_{n-1,n} \\ 0 & \cdots & 0 & 0 & d \end{bmatrix}.$$

Now, this matrix has determinant d after reducing modulo m_n; since it represents a generating sequence, d must be invertible in $\mathbb{Z}/m_n\mathbb{Z}$ and hence is coprime to m_n. This means every $b_{i,n}$ is congruent to a multiple of d modulo m_n. Using elementary operations we can "clean up" the last column by adding multiples of the bottom row to each other row, replacing all of the $b_{i,n}$ entries by 0's. We can similarly do this for each other entry above the diagonal, giving

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots \\ \vdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & d \end{bmatrix}.$$

□

3.3 General computations

We now turn to some more general techniques for understanding the orbits of the actions of $E_n(G)$ and $L_n(G)$. We start by defining:

Definition 3.16. If $(g_1, \dots, g_n)$ is a generating sequence, we say the i-th coordinate is *free* (or a *free spot*) if $(g_1, \dots, g_{i-1}, g_{i+1}, \dots, g_n)$ generates.

Note that we can always use left operations on $(g_1, \dots, g_n)$ to replace g_i with $w \cdot g_i$ where w is any word in the $g_1, \dots, g_{i-1}, g_{i+1}, \dots, g_n$. If the i-th spot is free, then w can be any element of G, so we can use left operations to replace g_i by any element $g \in G$. One consequence of this is that we can permute sequences that have free spots:

Lemma 3.17. *If $(g_1, \dots, g_n)$ is redundant, and $\sigma \in S_n$ is a permutation, then $(g_1, \dots, g_n)$ is equivalent to $(g_{\sigma(1)}, \dots, g_{\sigma(n)})$ under $L_n(G)$ (and hence also under $E_n(G)$).*

Proof. By redundancy, there is some coordinate k so that $(g_1, \dots, g_{k-1}, g_{k+1}, \dots, g_n)$ generates; we say that such a coordinate k is a "free spot". We can use left operations to replace g_k by any $g \in G$. In particular, we can build any g_i in the k-th spot. Then the i-th coordinate is a free spot, and we can build g_k there. Thus we can permute two coordinates g_i and g_k as long as one of them is in a free spot.

If i and j are two non-free coordinates, then pick a free coordinate k. We can build g_i in the k-th spot to make the i-th spot free. We can then build g_j in the i-th spot to make the j-th spot free, and then build g_i in the j-th spot to make the k-th spot free again. Finally, we can build g_k in the k-th spot again; the result is that we have swapped g_i and g_j.

We have thus shown that we can swap any g_i and g_j in our sequence. This proves the desired statement for transpositions σ, and by repeating the process (noting that a permutation of a redundant sequence is redundant) we can get any composition of transpositions and hence any permutation. □

Now, we show that if n is "large enough" then the action is transitive. It turns out that anything bigger than $\overline{r}(G)$ is "large enough". The proof of this fact is very similar to the proof of Tarski's Irredundant Basis Theorem 1.17:

Theorem 3.18. *If $n > \overline{r}(G)$, then $L_n(G)$ (and hence $E_n(G)$) acts transitively on $\Gamma_n(G)$.*

Proof. By Lemma 3.17, we can permute the entries of any sequence of length n. We will do this many times throughout the proof, (in particular we will often permute the coordinates so that the first entry is a "free spot").

Fix some irredundant generating sequence $s = (s_1, \dots, s_m)$ (so we have $m < n$); we want to show that an arbitrary generating sequence $(g_1, \dots, g_n)$ is equivalent to $(s_1, \dots, s_m, 1, \dots, 1)$. Following the proof of Theorem 1.17, we let $\ell(g)$ be the shortest length of a sequence of the elements s_i and their inverses that multiplies together to give g. For a generating sequence $t = (g_1, \dots, g_n)$, we again define $\ell(t) = \sum \ell(g_i)$, $m(t) = \max\{\ell(g_i)\}$, and $f(t) = |\{i : \ell(t_i) = m(t)\}|$. Recall that we view $\ell(t)$ as a measure of "how far" t is from s, and we use $m(t)$ and $f(t)$ as tiebreakers for sequences such that $\ell(t) = \ell(t')$.

Fix some $t = (g_1, \dots, g_n)$, and consider the set of all sequences t' which are equivalent to t under $L_n(G)$. Then, there is such a $t' = (g'_1, \dots, g'_n)$ that is minimal with respect to our measure, i.e. $(\ell(t), m(t), f(t))$ is minimal with respect to the lexicographical order.

Now, we claim that $m(t') = 1$. To prove this, first note that since we can permute the coordinates of t', we can assume without loss of generality that the first coordinate is free. Then, if $m(t') > 1$, some coordinate g'_i (which we can assume without loss of generality is g'_2) satisfies $\ell(g'_2) = m(t')$. Thus g'_2 can be written as $s^{\pm k} g''_2$ with $\ell(g''_2) < \ell(g'_2)$. We can then use left operations (and the fact that the first spot is free) to get

$$t' \sim (s_k, g'_2, \dots, g'_n) \sim (s_k, g''_2, \dots, g'_n)$$

However, since $\ell(g_2'') < \ell(g_2')$ and $\ell(s_k^{\pm 1}) = 1 < m(t')$, we have $\ell(t'') < \ell(t')$ or $\ell(t'') = \ell(t')$ and either $m(t'') < m(t')$ or $m(t') = m(t'')$ and $f(t'') < f(t')$. This contradicts minimality of t', so we must have $m(t') = 1$.

Now that we know $m(t') = 1$, this implies $\ell(g_i') = 1$ for each i, and therefore each g_i' is some $s_k^{\pm 1}$. Since t' generates, each s_k must appear in this way as some g_i'. We can then permute the coordinates to get

$$t' \sim (s_1^{\pm 1}, \ldots, s_k^{\pm 1}, g_{k+1}', \ldots, g_n').$$

Since the coordinates after the k-th are free, we can replace them all by 1's. Then if we have s_i^{-1} in the i-th coordinate, we can put s_i in the free n-th coordinate, put s_i in the i-th coordinate, and then put 1 back in the n-th coordinate. We thus get

$$t \sim t' \sim (s_1, \ldots, s_k, 1, \ldots, 1). \qquad \square$$

On the other hand, when $n = r(G)$ we generally don't expect that $E_n(G)$ or $L_n(G)$ acts transitively on $\Gamma_n(G)$. We can thus try to prove lower bounds on $c(G, E_n)$ and $c(G, L_n)$. In particular, consider the case where $n = r(G) = 2$. It turns out that commutators give an invariant of orbits:

Proposition 3.19 (Higman)**.** *Let G be a finite group with $r(G) = 2$. Then the function $f : \Gamma_n(G) \to G$ mapping (g_1, g_2) to the commutator $[g_1, g_2]$ is invariant under the action of $L_2(G)$. Moreover, the conjugacy class of $[g_1, g_2]$ is invariant under $E_2(G)$.*

Proof. To see $[g_1, g_2]$ is invariant under $L_2(G)$, we need to show that it is invariant under basic left operations, i.e.

$$[g_1, g_2] = [g_2^{\pm 1} g_1, g_2] = [g_1, g_1^{\pm 1} g_2].$$

This is a simple computation. For instance,

$$[g_2^{\pm 1} g_1, g_2] = (g_2^{\pm 1} g_1)^{-1} g_2^{-1} (g_2^{\pm 1} g_1) g_2 = g_1^{-1} g_2^{\mp 1} g_2^{-1} g_2^{\pm 1} g_1 g_2 = g_1^{-1} g_2^{-1} g_1 g_2 = [g_1, g_2].$$

The computation $[g_1, g_2] = [g_1, g_1^{\pm 1} g_2]$ is similar. So, $[g_1, g_2]$ is invariant under $L_2(G)$.

Showing that the conjugacy class of $[g_1, g_2]$ is invariant under $E_2(G)$ requires showing invariance under elementary right operations. For instance,

$$[g_1 g_2^{\pm 1}, g_2] = (g_1 g_2^{\pm 1})^{-1} g_2^{-1} (g_1 g_2^{\pm 1}) g_2 = g_2^{\mp 1} (g_1^{-1} g_2^{-1} g_1 g_2) g_2^{\pm 1} = g [g_1, g_2] g^{-1},$$

so these are indeed conjugate. A similar computation shows $[g_1, g_2]$ and $[g_1, g_2 g_1^{\pm 1}]$ are conjugate. $\square$

A result of Guralnick and Pak [GP03] proves that there is no corresponding invariant for $E_n(G)$ if $n \geq 3$. In particular, there is no word in the elements $g_1, \ldots, g_n$ and their inverses that is invariant under left operations on the sequence $(g_1, \ldots, g_n)$.

We can apply Higman's result to give a more explicit lower bound for $c(G, L_2)$:

Proposition 3.20. *Let G be a nonabelian finite group with $r(G) = 2$. Let (g, h) be a generating sequence of G, and let p be the smallest prime dividing $|g|$. Then $c_2(G, L_2) \geq p - 1$.*

Proof. Let $m = |g|$. Since p is the smallest prime dividing m, m and i are coprime for $1 \leq i \leq p-1$, and hence (g^i, h) generates G for $1 \leq i \leq p-1$ (as some power of g^i equals g). We claim that the sequences (g^i, h) are each in different orbits of $L_n(G)$. Since there are $p-1$ such sequences, this will imply there are at least $p-1$ orbits.

So, fix i, j with $1 \leq i < j \leq p-1$. To show (g^i, h) and (g^j, h) are in different orbits, it suffices to show their commutator invariants $[g^i, h]$ and $[g^j, h]$ are different. To see this, note that if we set $\ell = j - i$, we have

$$[g^j, h] = g^{-i} [g^\ell, h] g^i [g^i, h]$$

(this follows from a standard commutator identity that is easy to verify). Now, if $[g^j, h] = [g^i, h]$, we would have $g^{-i}[g^\ell, h]g^i = 1$ and hence $[g^\ell, h] = 1$. This would imply that g^ℓ and h commute; since ℓ is coprime to $|g|$, g is a power of g^ℓ, and we get that g and h commute. Then $G = \langle g, h \rangle$ is abelian, which is a contradiction. Thus we must have $[g^i, h] \neq [g^j, h]$, as desired. $\square$

We can apply this result to the case of finite simple groups. To do so, we need the following theorem, the "3/2 generation theorem." Its proof uses the classification theorem of finite simple groups; see [GK00].

Theorem 3.21. *Let S be a nonabelian finite simple group, and $g \in S$ a non-identity element. Then there is $h \in S$ with $\langle g, h \rangle = S$.*

Corollary 3.22. *Let S be a nonabelian finite simple group; then $c(S, L_2) \geq 4$ (in particular $L_2(S)$ does not act transitively on $\Gamma_2(S)$).*

Proof. By Burnside's $p^a q^b$ theorem (see, for instance, [DF04] chapter 19.2), there are at least three distinct primes dividing the order $|S|$. Let p be the largest such prime; then we must have $p \geq 5$. By Cauchy's Theorem, S has an element g of order p. By the 3/2 generation theorem, there is an element h such that (g, h) generates S. Then, by Proposition 3.20, $c(S, L_2) \geq p - 1 \geq 4$. □

While Proposition 3.22 is useful theoretically (in that it proves $c(S, L_2) > 1$), it is generally a poor lower bound on $c(S, L_2)$. For instance, $c(A_5, L_2) = 44$, and $c(S, L_2)$ is even higher for other small nonabelian finite simple groups that we tested. Higman's result Proposition 3.19 is much better - for A_5 it gives a sharp lower bound of $c(A_5, L_2) \geq 44$ (though it does not give a sharp lower bound in all cases). Of course, there is not an obvious way to apply Proposition 3.19 to get a good lower bound without using a computer to calculate commutator invariants.

4 Homogeneous Covers

Let G be a finite group, and $n \geq r(G)$. Recall that for a generating sequence $s = (g_1, \ldots, g_n)$ in $\Gamma_n(G)$, we define $\pi_s : F_n \to G$ as the surjective homomorphism taking the free basis $x_1, \ldots, x_n$ to $g_1, \ldots, g_n$, and let $K_s = \ker \pi_s$.

Definition 4.1. We define $K = \bigcap K_s$, the intersection of all of the kernels K_s for $s \in \Gamma_n(G)$. This is a normal subgroup of F_n. We define the *homogeneous cover of rank n* of G as F_n/K, and denote it by $H(n, G)$.

We let $\overline{x}_i$ denote the image of x_i in $H(n, G)$. Since $x_1, \ldots, x_n$ generate the free group F_n, the $\overline{x}_i$ generate $H(n, G)$. Also, for each generating sequence s, we have $K \subseteq K_s$ by construction, so by the universal mapping property for quotient groups π_s factors through $H(n, G)$. This gives a surjective homomorphism $\overline{\pi}_s : H(n, G) \to G$. If $s = (g_1, \ldots, g_n)$, then $\overline{\pi}_s(\overline{x}_i) = g_i$ by definition.

4.1 Homogeneous Groups

We start by defining two special types of subgroups, which were originally defined by Neumann and Neumann in [NN51]:

Definition 4.2. Let G be a group. We say a normal subgroup $H \trianglelefteq G$ is *hypercharacteristic* if, given any $N \trianglelefteq G$ with $G/H \cong G/N$, we have $H \leq N$. We say a normal subgroup $U \trianglelefteq G$ is *ultracharacteristic* if, given $N \trianglelefteq G$ with $G/U \cong G/N$, we have $U \geq N$.

Some basic properties of ultracharacteristic and hypercharacteristic subgroups are summarized in the next few propositions.

Proposition 4.3. *If U is ultracharacteristic in G, then $G/U \cong G/N$ in fact implies $U = N$ (i.e. U is the only normal subgroup of G with the quotient G/U).*

Proof. Let U satisfy this property. If $G/U \cong G/N$, then $N \leq U$ because U is ultracharacteristic. Assume that we in fact have $N < U$. Now, we know there is an isomorphism $h : G/N \to G/U$. By our assumption, U/N is a nontrivial subgroup of G/N, so $h[U/N]$ is a nontrivial subgroup of G/U.

By the correspondence theorem, there is $\widetilde{U} > U$ with $h[U/N] = \widetilde{U}/U$, and $\widetilde{U} \trianglelefteq G$. Then, by the third isomorphism theorem and our isomorphism h, we have

$$G/\widetilde{U} \cong \frac{G/U}{\widetilde{U}/U} = \frac{h[G/N]}{h[U/N]} \cong \frac{G/N}{U/N} \cong G/U.$$

Since $\widetilde{U} > U$, this contradicts our hypothesis. Therefore, our assumption must be false; it must be true that if $G/N \cong G/U$ then $U = N$. □

Proposition 4.4. *Ultracharacteristic implies hypercharacteristic, and hypercharacteristic implies characteristic.*

Proof. By Proposition 4.3, ultracharacteristic implies hypercharacteristic. If $H \leq G$ is hypercharacteristic, let α be an automorphism of G. The image $\alpha[H]$ is a normal subgroup of G, and α induces an isomorphism $G/H \cong G/\alpha[H]$. Since H is hypercharacteristic, $H \leq \alpha[H]$. Applying the same argument to α^{-1}, we get $H \leq \alpha^{-1}[H]$, and applying α gives $\alpha[H] \leq H$ and therefore equality. This means H is invariant under all automorphisms, so is a characteristic subgroup. □

Proposition 4.5. *If H is of finite index in G, H is hypercharacteristic if and only if it is ultracharacteristic.*

Proof. By Proposition 4.4, we just need to prove that if H is hypercharacteristic and of finite index in G, then H is ultracharacteristic. This follows from computing the index; if $G/N \cong G/H$, we know $H \leq N$ by hypothesis, and also $[G : N] = [G : H]$ because these are the orders of G/N and G/H. Then, we have $[G : H] = [G : N][N : H]$, which implies $[N : H] = 1$ and hence $H = N$. □

The most important example of an ultracharacteristic subgroup is the group K we used to define $H(n, G)$:

Proposition 4.6. *Let G be a finite group and n an integer, and as before define $K = \bigcap K_s$, the intersection of all kernels K_s for $s \in \Gamma_n(G)$. Then K is ultracharacteristic in the free group F_n.*

Proof. We need to show that if $L \trianglelefteq F_n$ satisfies $F_n/L \cong F_n/K$, then $L \subseteq K$. So, suppose we have such an L. Let $p : F_n \to F_n/L$ be the canonical projection, and let $h : F_n/L \to F_n/K$ be an isomorphism. For each $s \in \Gamma_n(G)$, consider the homomorphism

$$\widetilde{\pi}_s = \overline{\pi}_s \circ h \circ p,$$

where $\overline{\pi}_s : F_n/K \to G$ is the homomorphism taking $(\overline{x}_1, \ldots, \overline{x}_n)$ to s. Each $\widetilde{\pi}_s$ is a surjective homomorphism $F_n \to G$, so equals some $\pi_{s'} : F_n \to G$ for $s' \in \Gamma_n(G)$. We can thus define a map $\sigma : \Gamma_n(G) \to \Gamma_n(G)$ so that $\widetilde{\pi}_s = \pi_{\sigma(s)}$.

Now, suppose $\sigma(s) = \sigma(t)$. This means $\widetilde{\pi}_s = \widetilde{\pi}_t$, so

$$\overline{\pi}_s \circ h \circ p = \overline{\pi}_t \circ h \circ p.$$

Since p and h are surjective, this implies $\overline{\pi}_s = \overline{\pi}_t$, and hence $s = t$. So, σ is injective; since $\Gamma_n(G)$ is a finite set, it is also surjective, and hence is a permutation.

Further note that for each s, we have

$$L = \ker p \subseteq \ker \widetilde{\pi}_s = K_{\sigma(s)}.$$

Therefore, taking intersections and using the fact that σ is surjective,

$$L \subseteq \bigcap K_{\sigma(s)} = \bigcap K_s = K.$$

This proves K is ultracharacteristic, as desired. □

We can now give a number of equivalent definitions for a "'homogeneous group"', an idea first introduced by Gaschütz in [Gas55].

Definition 4.7. Let H be a finite group with $r(H) \leq n$. We say H is "homogeneous of rank n" if any of the following equivalent properties hold:
1) For any (and hence all) $s \in \Gamma_n(G)$, the kernel $K_s = \ker \pi_s$ is ultracharacteristic.
2) The homogeneous cover $H(n, H)$ is isomorphic to H.
3) The reduced Eulerian function $h_n(H)$ (see Definition 1.11) has value 1.
4) We have $|\operatorname{Aut}(H)| = |\Gamma_n(H)|$.

Proof of equivalence. We start by proving (1) and (2) are equivalent. By definition $H(n, H) = F_n/K$, where K is the intersection of all kernels K_s for $s \in \Gamma_n(G)$. If one K_s is ultracharacteristic, then every other K_t equals K_s, so $K = K_s$ and hence $H(n, H) = F_n/K = F_n/K_s \cong H$. Conversely, if $H(n, H) \cong H$, then K is ultracharacteristic. Since $F_n/K \cong H \cong F_n/K_s$ for each s, we have $K = K_s$, and hence K_s is ultracharacteristic.

Next, we prove (1) and (3) are equivalent. If any K_s is ultracharacteristic, it is equal to all of the other K_t. By Proposition 1.13, this means all of the elements of $\Gamma_n(H)$ are in the same orbit under the action of $\operatorname{Aut}(H)$. By Proposition 1.10, the automorphism group acts freely, and $h_n(H) = |\Gamma_n(H)|/|\operatorname{Aut}(H)|$ is the number of orbits, so must equal 1. Conversely, if $h_n(H) = 1$, we can run this argument backwards; there is only one orbit under $\operatorname{Aut}(H)$, which implies all of the kernels K_s are equal, and hence that their common value is ultracharacteristic.

Finally, it is trivial that (3) and (4) are equivalent, as $h_n(H) = |\Gamma_n(H)|/|\operatorname{Aut}(H)|$ by definition. □

An immediate consequence of this definition is:

Proposition 4.8. *If H is a homogeneous group of rank n, then any two elements $s, t \in \Gamma_n(H)$ are equivalent under the action of $\operatorname{Aut}(H)$.*

Proof. Since the action of the automorphism is free, we know $h_n(H) = 1$ is the number of orbits under $\operatorname{Aut}(H)$, so there must be an automorphism α with $\alpha \cdot s = t$. □

If we fix $s \in \Gamma_n(H)$ for a homogeneous group H, then the automorphisms of H are in one-to-one correspondence with $\Gamma_n(H)$ by associating $t \in \Gamma_n(H)$ with the automorphism taking s to t. This is analogous to the way automorphisms of a vector space are determined by taking one basis to another.

We can prove that the "rank n" in the definition of a homogeneous group in fact corresponds to the minimal length of a generating sequence $r(H)$:

Proposition 4.9. *If H is nontrivial and homogeneous of rank n, $r(H) = n$.*

Proof. Suppose $r(H) < n$. Then there is a generating sequence $h_1, \ldots, h_{n-1}$ of length $n-1$. Since H is nontrivial, there is a non-identity $h \in H$. Then, $(h_1, \ldots, h_{n-1}, 1)$ and $(h_1, \ldots, h_{n-1}, h)$ are elements of $\Gamma_n(H)$. By the previous proposition, these two generating sequences are equivalent under $\operatorname{Aut}(H)$, but looking at the last entry this means there is an automorphism α taking 1 to h. This is a contradiction, so we must have $r(H) = n$. □

We end this section with some propositions about homomorphisms out of homogeneous groups. The first is due to Gaschütz [Gas55].

Proposition 4.10. *Let H be homogeneous of rank n, and let $M, N \trianglelefteq H$. If $\alpha : H/M \to H/N$ is an isomorphism, there is $\beta \in \operatorname{Aut}(H)$ with $\alpha(hM) = \beta(h)M$.*

Proof. Let $s = (s_1, \ldots, s_n)$ be a length n generating sequence of H. This projects to a generating sequence $\overline{s}$ of H/M, which α carries to a generating sequence $\overline{t}$ of H/N. By Gaschütz's Lemma 1.21, $\overline{t}$ lifts to a generating sequence $t = (t_1, \ldots, t_n)$ of H. By Proposition 4.8 above, there is an automorphism $\beta \in \operatorname{Aut}(H)$ taking s to t. This means $\beta(s_i) = t_i$, and therefore $\alpha(s_i M) = t_i N = \beta(s_i)M$ by definition. Since the s_i generate H, we get $\alpha(hM) = \beta(h)M$ for each $h \in H$. □

Proposition 4.11. *Let H a homogeneous group of rank n, and G a finite group. If $\psi : H \to G$ is a surjective homomorphism, then for any $s \in \Gamma_n(G)$ there is a surjective homomorphism $\Psi_s : H \to H(n, G)$ with $\overline{\pi}_s \circ \Psi_s = \psi$.*

Proof. Let $\pi : F_n \to F_n/K = H(n,G)$ be the canonical projection. By Gaschutz's lemma, any $t \in \Gamma_n(G)$ lifts through ψ to a generating sequence $\hat{t}$ of H. Let $\gamma_t : F_n \to H$ be the surjective homomorphism taking the free basis $(x_1, \ldots, x_n)$ to $\hat{t}$. Then we have

$$\psi \circ \gamma_t = \pi_t = \overline{\pi}_t \circ \pi$$

because these maps each take x_i to t_i. Note that $H \cong F_n / \ker \gamma_t$, so by homogeneity $\ker \gamma_t = L$ is ultracharacteristic and thus independent of t. Note that

$$L = \ker \gamma_t \leq \ker \psi \circ \gamma_t = \ker \pi_t = K_t,$$

and hence $L \leq \bigcap K_t = K$. Therefore, $\pi : F_n \to H(n,G)$ factors through H by a homomorphism $\Psi_s : H \to H(n,G)$ so that $\Psi_s \circ \gamma_s = \pi$ (note Ψ_s is surjective because π is). Then, we have

$$\overline{\pi}_s \circ \Psi_s \circ \gamma_s = \overline{\pi}_s \circ \pi = \psi \circ \gamma_s.$$

Since γ_s is surjective, $\overline{\pi}_s \circ \Psi_s = \psi$, as desired. □

4.2 Homogeneous Covers and Subdirect Products

Combining Proposition 4.6 in the previous definition with the definition of homogeneous groups and homogeneous covers, we get:

Proposition 4.12. *The homogeneous cover $H(n,G)$ is a homogeneous group of rank n.*

The results of the previous section can therefore be used to study homogeneous covers. For instance, Proposition 4.11 implies that if H is a homogeneous group and G is a quotient of H, then $H(n,G)$ is a quotient of H.

Another way to study homogeneous covers is using the properties of "subdirect products":

Definition 4.13. Let $\{G_i : i \in I\}$ be a set of groups. A group G is called a *subdirect product* of the G_i if there is an embedding $h : G \to \prod G_i$, so that the projection $\pi_i[h[G]]$ onto the i-th coordinate equals all of G_i.

Proposition 4.14. *The homogeneous cover $H(n,G)$ is a subdirect product of $h_n(G)$ copies of G.*

Proof. By definition, there are $h_n(G)$ orbits of $\Gamma_n(G)$ under the action of $\mathrm{Aut}(G)$, and by Proposition 1.13, $K_s = K_t$ if and only if s and t are in the same orbit under $\mathrm{Aut}(G)$. Therefore, if we let $k = h_n(G)$, we can pick k sequences $s_1, \ldots, s_k$ so that the K_{s_i} are all distinct, and moreover any kernel K_s equals some K_{s_i}.

Then, define a homomorphism $h : F_n \to G^k$ as the product of the maps π_{s_i}, so $x \in F_n$ maps to $(\pi_{s_1}(x), \ldots, \pi_{s_k}(x))$. Note that $h(x) = 1$ if and only if $\pi_{s_i}(x) = 1$ for each x, if and only if $x \in \bigcap K_{s_i} = K$. So, h factors through to an injective homomorphism $h' : H(n,G) \to G^k$ (which is the product of the maps $\overline{\pi}_{s_i} : H(n,G) \to G$). Moreover, this map is surjective onto each coordinate, as projecting onto the i-th coordinate of G gives the surjective map $\overline{\pi}_{s_i}$. □

A subdirect product of a number of copies of the same group G inherits many of the properties of G. Applying our previous proposition, for instance, we get the following list of facts about homogeneous covers.

Theorem 4.15. *Let G be a finite group and $H(n,G)$ the n-th homogeneous cover. Then:*
1) $H(n,G)$ is a finite group, with order at most $|G|^{h_n(G)}$.
2) $H(n,G)$ is abelian if and only if G is.
3) $H(n,G)$ is nilpotent if and only if G is, and if they are their nilpotency classes are equal.
4) $H(n,G)$ is solvable if and only if G is, and if they are their derived lengths are equal.
5) A simple group S appears in the Jordan-Holder decomposition of $H(n,G)$ if and only if it appears in the decomposition of G.
6) The exponent $\exp(G)$ is equal to $\exp(H(n,G))$.

It is straightforward to prove that each of these holds, and in fact that a more general version of each statement holds for subdirect products in general. In fact, one can prove a general result in universal algebra describing how subdirect products inherit properties; see [Lyn59].

Another consequence of Proposition 4.14 is that we can compute the order of the elements $\overline{x}_i \in H(n, G)$ (recall that $(\overline{x}_1, \ldots, \overline{x}_n)$ is the projection of the free basis $(x_1, \ldots, x_n)$ of F_n):

Proposition 4.16. *If $n > r(G)$, $|\overline{x}_i| = \exp(G)$. If $n = r(G)$, then*

$$|\overline{x}_i| = \operatorname{lcm}\{|g| : g \text{ is in a generating sequence in } \Gamma_n(G)\}.$$

Proof. We start by proving that the equality

$$|\overline{x}_i| = \operatorname{lcm}\{|g| : g \text{ is in a generating sequence in } \Gamma_n(G)\}$$

actually holds for any n. We know G is isomorphic to a subgroup of G^k, and in particular that $\overline{x}_i$ corresponds to a k-tuple of elements in generating sequences of G. Therefore $|\overline{x}_i|$ divides the LCM in question. On the other hand, if g appears in any generating sequence s, then a permutation s' of this generating sequence has g in the i-th spot, and $\overline{\pi}_{s'}$ takes $\overline{x}_i$ to g. Thus $|g|$ divides $|\overline{x}_i|$; since this holds for every g in a sequence in $\Gamma_n(G)$, the LCM of the $|g|$ divides $|\overline{x}_i|$ and hence equality holds.

If $n > r(G)$, then every element $g \in G$ appears in some length n generating sequence, so we have actually proven

$$|\overline{x}_i| = \operatorname{lcm}\{|g| : g \in G\},$$

which equals $\exp(G)$ by definition. □

We call the value

$$|\overline{x}_i| = \operatorname{lcm}\{|g| : g \text{ is in a generating sequence in } \Gamma_n(G)\}$$

the "generating exponent" $\operatorname{gexp}(G)$, and note that $\operatorname{gexp}(G)$ always divides $\exp(G)$. It is often the case that $\operatorname{gexp}(G) = \exp(G)$, but there are examples where equality does not hold. The smallest group for which $\operatorname{gexp}(G) \neq \exp(G)$ is $G = (Z_3 \times Z_3) \rtimes Q_8$ (with the semidirect product corresponding to the embedding of Q_8 in $\mathrm{GL}_2(\mathbb{F}_3)$). This group has order 72, and its exponent is 12 while its generating exponent is 4.

We end this section with one more result, relating homogeneous covers of different ranks for the same group:

Proposition 4.17. *If G is a finite group and $m > n \geq r(G)$, there is a natural surjective group homomorphism $h : H(m, G) \to H(n, G)$.*

Proof. Let $f : F_n \to F_m$ be given by $f(x_i) = x_i$, and $g : F_m \to F_n$ be given by $g(x_i) = x_i$ for $i \leq n$ and $g(x_i) = 1$ for $i > n$ (both maps extend to be homomorphisms by the universal property for free groups). Note that $g \circ f$ is the identity on F_n. Also, given $s \in \Gamma_n(G)$, let $s' = f(s) = (s, 1, \ldots, 1)$ in $\Gamma_m(G)$, and note that $\pi_{s'} = \pi_s \circ g$. Let $K_s = \ker \pi_s$ for $s \in \Gamma_m(G)$, $L_s = \ker \pi_s$ for $s \in \Gamma_n(G)$, $K = \bigcap_{s \in \Gamma_m} K_s$, $K' = \bigcap_{s \in \Gamma_n} K_{s'}$, and $L = \bigcap_{s \in \Gamma_n} L_s$. Note $F_m/K = H(m, G)$, $F_n/L = H(n, G)$, and $K \leq K'$.

Now, we claim $f[F_n] \cdot K' = F_m$ and $f[F_n] \cap K' = f[L]$. To prove the first equality, note that for $i \leq n$, $x_i \in f[F_n]$, and for $i > n$, $\pi_{s'}(x_i) = 1$ because $s' = (s_1, \ldots, s_n, 1, \ldots, 1)$, so $x_i \in K'_s$ for each s and hence $x_i \in K'$. Thus every x_i is in either $f[F_n]$ or K', so $F_m = \langle x_1, \ldots, x_m \rangle \subseteq f[F_n] \cdot K'$. For the second equality $f[F_n] \cap K' = f[L]$, note that since $\pi_{s'} \circ f = \pi_s \circ g \circ f = \pi_s$,

$$f[F_n] \cap K_{s'} = \{f(x) : x \in F_n, \pi_{s'}(f(x)) = 1\} = \{f(x) : x \in F_n, \pi_s(x) = 1\} = f[L].$$

Finally, note that since $K \leq K'$, there is a canonical projection $F_m/K \to F_m/K'$, and then by the second isomorphism theorem there is an isomorphism

$$\frac{F_m}{K'} = \frac{f[F_n] \cdot K'}{K'} \cong \frac{f[F_n]}{f[F_n] \cap K'} = \frac{f[F_n]}{f[L]} \cong \frac{F_n}{L} = H(n, G)$$

These maps compose to give our surjective homomorphism $H(m, G) \to H(n, G)$, which is explicitly given by letting $h(xK)$ equal $x'L$ for some $x' \in F_n$ such that $f(x') \in xK'$. □

In general, there is not a natural homomorphism in the other direction, $H(n,G) \to H(m,G)$. However, such homomorphisms do exist in special cases, and understanding exactly when they do exist is an open question.

4.3 Computations of Homogeneous Covers

To understand a new concept, it is always useful to make computations. For our case, we can compute homogeneous covers of groups in some special cases. Since $H(n,G)$ is a quotient of F_n, we can always give a presentation for $H(n,G)$ with n variables. Our approach to doing this is usually to start by finding relations that must hold in $H(n,G)$ (for instance, if $\exp(G) = k$, then our knowledge that $\exp(H(n,G)) = k$ as well means the relations x_i^k must hold). Once we have some set of relations normally generating some $K_0 \trianglelefteq F_n$ with $K_0 \leq K$, we try to prove that $K_0 = K$ and hence $F_n/K_0 = H(n,G)$ by showing that if $x \notin K_0$ there is a surjective homomorphism $\pi : F_n \to G$ with $\pi(x) \neq 1$, which means $x \notin K$ by definition.

The easiest case is that of abelian groups. Recall that if A is a finite abelian group, then $A \cong Z_{m_1} \times \cdots \times Z_{m_n}$ where $m_n > 1$ and m_i divides m_{i-1} for each $i > 1$. Moreover, $r(A) = n$. Then:

Proposition 4.18. *If A is a finite abelian group as above, $H(k,A) \cong Z_{m_1}^k$ for each $k \geq n$.*

Proof. By definition, $H(k,A) \cong F_k/K$, where K is the kernel of all of the surjective maps $\pi_s : F_k \to A$. Since A is abelian, the kernels K_s each contain the commutator subgroup F_n', so K contains F_n'. Moreover, since A has exponent m_1, each kernel K_s contains $x_i^{m_1}$ for each element of the free basis x_i, and hence so does K. Therefore, K contains $K_0 = \langle\langle F_n', x_1^{m_1}, \ldots, x_k^{m_1}\rangle\rangle$ (the normal closure of these elements), so F_n/K is a quotient of $F_n/K_0 \cong Z_{m_1}^k$.

To see $K_0 = K$, first note that if $(a_1, \ldots, a_k) \in Z_{m_1}^k$ is nonzero then there is a surjective $\pi : Z_{m_1}^k \to A$ so that $\pi(a_1, \ldots, a_k)$ is nonzero (if a_i is nonzero, take the map $Z_{m_1}^k \to Z_{m_1} \times \cdots \times Z_{m_n}$ mapping the i-th coordinate to the first, and map the other coordinates of $Z_{m_1}^k$ onto to the coordinates of A). Thus if x is any element of F_n not in K_0, x corresponds to a nonzero $(a_1, \ldots, a_n) \in Z_{m_1}^k \cong F_n/K_0$, and π lifts to a map $\hat{\pi} : F_n \to A$ with $\pi(x) \neq 0$. Therefore $x \notin K$, which means $K = K_0$, and $H(m,A) \cong Z_{m_1}^k$. □

Another way to derive this result is by a lemma of F. Levi [Lev33], which proves that the only characteristic subgroups of F_n satisfying $F_n' < C < F_n'$ are $F_n' \cdot (F_n)^m$ (i.e. the subgroups $\langle\langle F_n', x_1^{m_1}, \ldots, x_k^{m_1}\rangle\rangle$ that came up in our proof).

Another situation where we can compute $H(m,G)$ explicitly is when G is a nonabelian finite simple group. This follows immediately from Theorem 2.22, which proved that $r(S^{h_m(S)}) = m$ (in fact, our proof consisted of showing $F_n/K \cong S^{h_m(S)}$).

Proposition 4.19. *Let S be a nonabelian finite simple group, and $n = h_m(G)$ (the reduced Eulerian function). Then $H(m,S) \cong S^n$.*

If we are given a presentation of a group, we can sometimes use this to compute a presentation for $H(m,G)$. An example where this works is the group T_p of 3×3 upper triangular matrices (with 1's on the diagonal) over the finite field $\mathbb{F}_p$, for a prime p. This is one of the two nonabelian groups of order p^3.

Proposition 4.20. *We have*

$$H(m,T_p) = \langle x_1, \ldots, x_n | x_i^p; [x_i,x_j]^p; [x_i,x_j] \text{ central}\rangle$$

(where saying "x central" means we include all commutators $[x,x_i]$, so x commutes with everything). In particular, this means $|H(m,T)| = p^{m(m+1)/2}$, and $H(2,T) \cong T_p$.

Proof. Properties of the group T_p are that every nonzero element has order p, the group is generated by two matrices

$$X_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad X_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

and that commutator $[X_1, X_2]$ generates $G' = Z(G)$. So, for each generating sequence $(g_1, \ldots, g_n)$, we have $g_i^p = 1$, that $[g_i, g_j]^p = 1$, and that $[g_i, g_j] \in G' = Z(G)$. Therefore, the group

$$G = \langle x_1, \ldots, x_n | x_i^p; [x_i, x_j]^p; [x_i, x_j] \text{ central} \rangle$$

is a quotient of F_n by a subgroup contained in K. To show that this quotient is equal to $H(m, T) = F_n/K$, we need to show that for every nontrivial $g \in G$ there is a surjective homomorphism $G \to T_p$ taking g to a nontrivial element of T_p.

We do this by noting that the relations defining G mean that any $g \in G$ can be written in the following "normal form":

$$g = x_1^{a_1} \cdots x_m^{a_m} \prod_{i<j} [x_i, x_j]^{\varepsilon_{ij}},$$

where $0 \leq a_i < p$ and $0 \leq \varepsilon_{ij} < p$. We can see this by noting that we can switch a pair $x_i x_j$ with $x_j x_i$ at the expense of adding a commutator $[x_i, x_j]$ that is in the center and can therefore be moved to the right side of the expression.

Therefore, suffices to show that if we have an element

$$g = x_1^{a_1} \cdots x_m^{a_m} \prod_{i<j} [x_i, x_j]^{\varepsilon_{ij}},$$

that maps to 1 in T_p under any map $G \to T_p$ mapping $(x_1, \ldots, x_n)$ to a generating sequence, then a_i or ε_{ij} nontrivial. Now, for any $i < j$, we can consider the generating sequence with X_1 in the i-th spot and X_2 in the j-th spot, and 1's everywhere else. The corresponding map into T_p takes g to

$$X_1^{a_i} X_2^{a_j} [X_1, X_2]^{\varepsilon_{ij}}.$$

By assumption, this element is the identity in T_p, so we have

$$\begin{bmatrix} 1 & a_1 & a_i a_j + \varepsilon_{ij} \\ 0 & 1 & a_j \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for each i, j. Therefore, $a_i = a_j = \varepsilon_{ij} = 0$, proving the desired statement.

So, $G = H(m, T_p)$, and distinct normal forms represent distinct elements. We can check that there are $p^m \cdot p^{m(m-1)/2} = p^{m(m+1)/2}$ distinct normal forms, so this is the order $|H(m, T_p)|$. For $m = 2$, this implies $|H(2, T_p)| = p^3 = |T_p|$; since T_p is a quotient of $H(2, T_p)$ by definition, this means $H(2, T_p) \cong T_p$. □

A similar but slightly more complicated computation works for the quaternion group Q_8:

Proposition 4.21. *We have $H(2, Q_8) \cong Q_8$ and*

$$H(m, Q_8) \cong \langle x_1, \ldots, x_m | x_i^4; [x_i, x_j]^2; x_i^2 \text{ and } [x_i, x_j] \text{ central} \rangle$$

for $m > 2$. In particular, $|H(m, Q_8)| = 2^{m(m+3)/2}$ for $m > 2$.

Proof. First consider the case $m = 2$. For any generating sequence (g_1, g_2), we have $g_i^4 = 1$ and $g_1^2 = g_2^2 = [g_1, g_2] \in Z(Q_8)$. Therefore, $H(2, Q_8)$ is a quotient of

$$\langle x_1, x_2 | x_1^4; x_2^4; x_1^2 = x_2^2 = [x_1, x_2] \text{ central } \rangle.$$

Any element of this group can be put in the normal form

$$[x_1, x_2]^{\varepsilon} x_1^{k_1} x_2^{k_2}$$

for $\varepsilon, k_1, k_2 \in \{0, 1\}$. Hence this group has order at most 8. Since Q_8 is a quotient group of it that has order 8, this means $H(m, Q_8)$ must be isomorphic to Q_8.

Then consider the $m > 2$ case. If $m > 2$ and $(g_1, \ldots, g_m)$ is any generating sequence of Q_8, we know $g_i^4 = 1$, $[g_i, g_j]^2 = 1$, and that $[g_i, g_j]$ and g_i^2 are central (as $Q_8' = Z(G) = \langle -1 \rangle$). So, $H(m, Q_8)$ is a quotient of

$$\langle x_1, \ldots, x_n | x_s^4; [x_s, x_t]^2; x_s^2 \text{ and } [x_s, x_t] \text{ central} \rangle.$$

We can put elements of this group in a normal form

$$\prod_{s<t} [x_s, x_t]^{\varepsilon_{st}} \prod_s x_s^{k_s}$$

with $0 \leq k_s < 4$ and $0 \leq \varepsilon_{st} < 2$. We claim distinct normal forms represent distinct elements of $H(m, Q_8)$, which will prove that $H(m, Q_8)$ has this presentation and that it has order $4^m 2^{\binom{m}{2}} = 2^{m(m+3)/2}$. Note that it suffices to show that the only normal form representing the identity has $k_s = \varepsilon_{st} = 0$. For convenience, if $s > t$ we will let ε_{st} denote ε_{ts} (as $[x_s, x_t]$ has order 2, $[x_s, x_t] = [x_t, x_s]$, so any commutators in the "wrong order" can be switched to the correct order).

So, suppose a normal form

$$\prod_{s<t} [x_s, x_t]^{\varepsilon_{st}} x_s^{k_t}$$

does represent the identity element. By using the generating sequence with $g_s = i$ and $g_t = j$, and the other g_r equal to 1, projecting $(x_1, \ldots, x_n)$ to this sequence in Q_8 gives

$$1 = (-1)^{\varepsilon_{st}} i^{k_s} j^{k_t} \tag{4}$$

Similarly, if we take the generating sequence with $g_s = i$, $g_t = j$, $g_u = -1$ for distinct s, t, u (which we can do because $m \geq 3$), and set the other $g_r = 1$, we get

$$1 = (-1)^{\varepsilon_{st}} i^{k_s} j^{k_t} (-1)^{k_u}. \tag{5}$$

So, for any coordinate u, we can set equations 4 and 5 equal to each other, and get $1 = (-1)^{k_u}$. This means each k_u is either 0 or 2. Equation 4 then reduces to

$$1 = (-1)^{\varepsilon_{st}} (-1)^{k_s/2} (-1)^{k_t/2} = (-1)^{\varepsilon_{st} + k_s/2 + k_t/2} \tag{6}$$

Finally, consider the generating sequence with $g_s = i$, $g_t = j$, and $g_u = i$ (and other coordinates 1). This gives

$$1 = (-1)^{\varepsilon_{st}} (-1)^{\varepsilon_{tu}} i^{k_s} j^{k_t} i^{k_u} = (-1)^{\varepsilon_{st} + \varepsilon_{tu} + k_s/2 + k_t/2 + k_u/2} \tag{7}$$

Combining equations 6 and 7, we get that $\varepsilon_{tu} = k_u/2$ for any coordinates t, u. Switching the coordinates t and u gives the equation $\varepsilon_{ut} = k_t/2$. Thus $k_t = k_u$ for all t, u, and therefore there is a constant $c \in \{0, 1\}$ so that every $k_s/2$ and ε_{st} is equal to c. If c was 1, then equation 6 would give $1 = (-1)^3 = -1$, which is a contradiction. So we must have $c = 0$, meaning each k_s and each ε_{st} is zero. Thus our normal form was the trivial one, as desired. □

5 Computation of $c(G, E_n)$ and $c(G, L_n)$ for $n = r(G)$

This section describes work done by Keith Dennis and me, to try to understand the number of orbits of G under the actions of $E_n(G)$ and $L_n(G)$ in the case $n = r(G)$. Theorem 3.15 of Diaconis and Graham, which answers this question when G is abelian, is the foundation of this study. We have tried various theoretical and computational techniques to consider various other cases.

5.1 Determinant Functions

Recall the argument of Diaconis and Graham to prove their theorem. The essence of the argument is that, given a finite abelian group $A = Z_{m_1} \times \cdots \times Z_{m_n}$, there is a "determinant function" on $\Gamma_n(A)$ that takes values in $(\mathbb{Z}/m_n\mathbb{Z})^\times$ (this determinant function is defined by first projecting s to a generating sequence $\overline{s}$ of $Z_{m_n}^n$, viewing $\overline{s}$ as a full-rank $n \times n$ matrix S over $\mathbb{Z}/m_n\mathbb{Z}$, and then

taking the determinant of S). The rest of the proof shows that this determinant function is exactly the invariant that is needed - two generating sequences s and t are equivalent if and only if their determinants are equal.

Moreover, we note that this construction of a determinant function can be extended to work for any sequence in A^n (though the resulting value is no longer guaranteed to be in the group of units $(\mathbb{Z}/m_n\mathbb{Z})^\times$). Note that most of the characterizing properties of the determinant function carry over to this $d : A^n \to \mathbb{Z}/m_n\mathbb{Z}$:

- d is multilinear: if $a = (a_1, \ldots, a_n) \in A^n$ and $a' = (a_1, \ldots, a_i', \ldots, a_n)$ differs from a in one entry, then $d(a) + d(a') = d(a_1, \ldots, a_i + a_i', \ldots, a_n)$.
- d is skew-symmetric: if $\sigma \in S_n$ is a permutation, we have $d(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = \operatorname{sgn}\sigma d(a_1, \ldots, a_n)$.

For convenience, we also list the properties of d that come from the theorem of Diaconis and Graham:

- If $(a_1, \ldots, a_n)$ is a generating sequence of A, then $d(a_1, \ldots, a_n)$ is a unit in $\mathbb{Z}/m_n\mathbb{Z}$.
- If $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are equivalent under left operations, then $d(a_1, \ldots, a_n) = d(b_1, \ldots, b_n)$.
- If $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are generating sequences, then $d(a_1, \ldots, a_n) = d(b_1, \ldots, b_n)$ means they are equivalent under left operations.

Now, we want to generalize this by constructing functions $d : G^n \to \mathbb{Z}/m\mathbb{Z}$ that satisfy similar properties for nonabelian groups G. The first example is for the dihedral group of order $2n$, D_{2n}. Define a determinant function $d : (D_{2n})^2 \to \mathbb{Z}/n\mathbb{Z}$ by:

$$\begin{aligned} d(R^i, R^j) &= 0 \\ d(R^iF, R^j) &= -j \\ d(R^i, R^jF) &= i \\ d(R^iF, R^jF) &= j - i \end{aligned}$$

We can check (through easy but somewhat tedious calculations) that d is skew-symmetric and that it is invariant under left operations. Moreover, recall that in Example 1.9 we showed that the length 2 generating sequences of D_{2n} are given by:

1. Sequences (R^iF, R^j) with $1 \leq i, j \leq n$ and $(j, n) = 1$.
2. Sequences (R^i, R^jF) with $1 \leq i, j \leq n$ and $(i, n) = 1$.
3. Sequences (R^iF, R^jF) with $1 \leq i, j \leq n$ and $(j - i, n) = 1$.

By construction, we can see that $d(a, b)$ is a unit in $\mathbb{Z}/n\mathbb{Z}$ if and only if (a, b) is a generating sequence of D_{2n}. Moreover, we can check that if (a, b) is a generating sequence and $d(a, b) = d(a', b')$, then (a, b) are equivalent to (a', b') under left operations. To see this, start with the case (R^i, R^jF), which has determinant i (which is a unit in $\mathbb{Z}/n\mathbb{Z}$). By left multiplication, we get

$$(R^i, R^jF) \sim (R^i, R^{ik+j}F) = (R^i, R^JF);$$

since i is a unit, we can get any element R^J, and hence (R^i, R^jF) is equivalent to any sequence of type (1) with determinant i. Moreover, if (R^IF, R^JF) is a sequence of type (2) with determinant i, then $J - I = i$, and we get

$$(R^i, R^jF) \sim (R^i, R^JF) \sim (R^JFR^i, R^JF) = (R^{J-i}F, R^JF) = (R^IF, R^JF).$$

Finally, if (R^IF, R^{-i}) is a sequence of type (3) with determinant i, we get

$$(R^i, R^jF) \sim (R^IF, R^{I+i}F) \sim (R^IF, R^IFR^{I+i}F) = (R^IF, R^{-i})$$

Summarizing, d satisfies a similar list of properties to the determinant function from the argument of Diaconis-Graham:

- d is alternating: $d(a,b) = -d(b,a)$
- A sequence (a,b) generates D_{2n} if and only if $d(a,b)$ generates $\mathbb{Z}/n\mathbb{Z}$ (i.e. is a unit).
- If (a,b) and (a',b') are equivalent under left operations, then $d(a,b) = d(a',b')$.
- If (a,b) and (a',b') are generating sequences with $d(a,b) = d(a',b')$, then (a,b) and (a',b') are equivalent under left operations.

We note that the second condition is in fact stronger than the corresponding condition for Diaconis-Graham's determinant, as for D_{2n} we have (a,b) generates if and only if $d(a,b)$ is a unit, while for abelian groups only one implication holds.

Finally, we can show that this d satisfies an analogue to multilinearity, namely

$$d(a\alpha, b) = {}^{\alpha}d(a,b) + d(\alpha, b),$$

where we let ${}^{\alpha}d(a,b)$ denote $d(\alpha^{-1}a\alpha, \alpha^{-1}b\alpha)$. This is straightforward but quite tedious to check.

Another situation in which we can construct a determinant function is for nonabelian groups of order pq, where p, q are distinct primes with p dividing $q-1$. In this situation, we know there is a unique (up to isomorphism) nonabelian group of order pq, with presentation

$$G = \langle X, Y | X^p = Y^q = 1, XY = Y^r X \rangle,$$

where r is an integer which has order p in $(\mathbb{Z}/q\mathbb{Z})^{\times}$. From this presentation, and the fact that we know $|G| = pq$, we can see that every element of G can be written as X^iY^j, and two elements X^iY^j and X^IY^J are equal if and only if $i \equiv I \pmod{p}$ and $j \equiv J \pmod{q}$.

We can use this presentation to compute in G. If we let s be the inverse of r modulo q, then we can compute $YX = XY^s$, and then get $Y^jX = XY^{js}$ and $Y^jX^i = X^iY^{js^i}$. This allows us to compute in G in terms of this presentation. In particular, we can compute powers of an arbitrary $X^iY^j \in G$:

$$(X^iY^j)^k = X^{ki}Y^{j\omega(i,k)}$$

where $\omega(i,k)$ is an integer given by:

$$\omega(i,k) = \sum_{m=0}^{k-1} s^{mi} = \begin{cases} 0 & k = 0 \\ k & i = 0 \\ \frac{s^{ki}-1}{s^i-1} & k, i \neq 0 \end{cases}.$$

Note that $r(G) = 2$ in this case, as G is not cyclic but is generated by the two elements X and Y. So, we define a determinant function $d : G^2 \to \mathbb{Z}/q\mathbb{Z}$ by:

$$d(X^iY^j, X^kY^\ell) = (s^i - 1)\ell - (s^k - 1)j.$$

Note that since $|Y| = q$, $|X| = p$, and s has order p in $(\mathbb{Z}/q\mathbb{Z})^{\times}$, this is well-defined (as if $X^iY^j = X^IY^J$ we know $J \equiv j \pmod{q}$, and also that $I \equiv i \pmod{p}$ implies $s^I \equiv s^i \pmod{q}$).

We can then check that this determinant function d satisfies similar properties to the others we have studied. For instance, it is clear d is antisymmetric (i.e. $d(a,b) = -d(b,a)$). To see that it is invariant under left operations, we can compute

$$(X^kY^l) \cdot (X^iY^j) = X^kX^iY^{\ell s^i}Y^j = X^{k+i}Y^{j+\ell s^i}$$

and thus

$$d(X^kY^\ell X^iY^j, X^kY^\ell) = d(X^{k+i}Y^{j+\ell s^i}, X^kY^\ell) = (s^{k+i} - 1)\ell - (s^k - 1)(\ell s^i + j)$$
$$= (s^{k+i} - 1)\ell - (s^k s^i - s^i)\ell - (s^k - 1)j = (s^i - 1)\ell - (s^k - 1)j$$

That d is invariant under the other possible left operations follows from a similar computation.

Next, we claim that a sequence (a,b) generates if and only if $d(a,b)$ is a unit in $\mathbb{Z}/q\mathbb{Z}$, i.e. if and only if $d(a,b) \neq 0$. To see this, we consider an arbitrary sequence $(a,b) = (X^iY^j, X^kY^\ell)$ and split into four cases based on the value of i and k:

- $i = k = 0$: In this case $(a, b) = (Y^j, Y^\ell)$, which doesn't generate, and $d(a, b) = 0$.
- $i = 0$, $k \neq 0$: In this case $(a, b) = (Y^j, X^k Y^\ell)$. This generates as long as $j \neq 0$, and $d(a, b) = -(s^k - 1)j$ is nonzero as long as $j \neq 0$.
- $k = 0$, $i \neq 0$: Apply antisymmetry and the previous case.
- $i, k \neq 0$: Note that $(X^i Y^j, X^k Y^\ell)$ does not generate G if and only if $X^k Y^\ell$ is a power of $X^i Y^j$ (if this holds, the pair generates a cyclic subgroup; if not, then it must generate G by order considerations). Then, the computation $(X^i Y^j)^n = X^{in} Y^{j\omega(i,n)}$ means that (a, b) fails to generate if and only if $\ell = j\omega(i, n)$, where n satisfies $in = k$. Above, we showed $\omega(i, n) = \frac{s^{in}-1}{s^i-1} = \frac{s^k-1}{s^n-1}$. So, the pair fails to generate if and only if $\ell = j\frac{s^k-1}{s^i-1}$ if and only if $d(a, b) = (s^i - 1)\ell - (s^k - 1)j = 0$.

Moreover, we can check that this determinant function d determines orbits of $L_2(G)$ exactly, in that if $d(a, b) = d(a', b') \neq 0$ then (a, b) and (a', b') are equivalent generating sequences under $L_2(G)$. To show this, we show that any sequence (a, b) with determinant $d = d(a, b)$ is equivalent to $(X, Y^{d/(s-1)})$. This is a straightforward computation; we can use left operations to get (a, b) to a sequence of the form (XY^j, Y^ℓ), check that $d(XY^j, Y^\ell) = d$ implies $\ell = d/(s-1)$, and then multiply the first entry on the left by an appropriate power of Y^ℓ to get X.

This determinant function also satisfies the same multilinearity as the one for the dihedral groups, namely

$$d(a\alpha, b) = {}^\alpha d(a, b) + d(\alpha, b)$$

(where ${}^\alpha d(a, b)$ denotes $d(\alpha^{-1}a\alpha, \alpha^{-1}b\alpha)$). This is again a straightforward computation. Letting $a = X^i Y^j$, $\alpha = X^I Y^J$, and $b = X^k Y^\ell$, note first we can compute

$$a\alpha = X^i Y^j X^I Y^J = X^{i+I} Y^{js^I + J}$$

$$\alpha^{-1} a \alpha = Y^{-J} X^{-I} X^i Y^j X^I Y^J = Y^{-J} X^{-I} X^i X^I Y^{J + js^I} = X^i Y^{J + js^I - Js^i}$$

and similarly

$$\alpha^{-1} b \alpha^{-1} = X^k Y^{J + \ell s^I - Js^k}.$$

We then can compute

$$\begin{aligned}
{}^\alpha d(a, b) + d(\alpha, b) &= d(X^i Y^{J+js^I - Js^i}, X^k Y^{J+\ell s^I - Js^k}) + d(X^I Y^J, X^k Y^\ell) \\
&= (s^i - 1)(\ell s^I - J(s^i - 1)) - (s^k - 1)(js^I - J(s^i - 1)) + (s^I - 1)\ell - (s^k - 1)J \\
&= (s^i - 1)s^I \ell - (s^k - 1)s^I j + (s^I - 1)\ell - (s^k - 1)J \\
&= (s^{i+I} - 1)\ell - (s^k - 1)(js^I + J) = d(a\alpha, b)
\end{aligned}$$

We have shown that we can construct a determinant function for two classes of nonabelian groups, the dihedral groups and the nonabelian groups of order pq. We note that these are both semidirect products, and in fact semidirect products of a cyclic group by another cyclic group. We also note that these determinant functions are closely related to Higman's commutator invariant (proposition 3.19). For instance, for the dihedral group we have

$$[a, b] = R^{2d(a,b)},$$

where R is the generator of the normal subgroup $\langle R \rangle \cong \mathbb{Z}/n\mathbb{Z}$.

One can ask whether it is possible to construct determinant functions in more general situations. A natural place to start might be for an arbitrary semidirect product of two cyclic groups. Ultimately, we might hope to construct a determinant function for other types of semidirect products, or even possibly for arbitrary solvable groups.

5.2 p-Groups

The determinant functions of the previous section give a good understanding of the action of $L_n(G)$ for some specific classes of groups G. However, it is not at all apparent how to understand this action for more complicated groups. Moreover, the sets $\Gamma_n(G)$ quickly get large, making it difficult to work out details of the action of $L_n(G)$ and get intuition for this action. To better understand this action, I have used computer programs to calculate values of $c(G, E_n)$. These programs are written using the GAP system [GAP08].

Since the value of $c(G, E_n)$ is known for abelian groups, the obvious "next step" is to work with nilpotent groups, and particularly p-groups. Also, we restrict our attention to the $n = r(G) = 2$ case, as computations are the easiest and fastest in this situation. We give some computations in Table 5.2, for groups G satisfying $r(G) = 2$ and $|G| = p^k$ for $2 \leq k \leq 5$ and $p = 2, 3$. Though this is a relatively small number of groups, we can already see interesting patterns.

The first two columns of each table give the group structure. The second column gives GAP's name of that group in its small groups library (where, for instance, $(16, 9)$ is the 9th group in the library of groups of order 16). The first column gives a simple description of the group if one exists (letting QD_m denote quasidihedral groups and Q_m denote generalized quaternion groups). Then, the third column lists the value of $c(G, E_2)$ and last column gives the nilpotence class of G.

The most striking fact about this data is that for every 2-group listed, $c(G, E_2)$ is either 1, 2, or 4, i.e. is a power of 2, and that for every 3-group listed, $c(G, E_2)$ is either 2, 6, or 18. Indeed, all of the computations we have run for 2-groups have found that $c(G, E_n)$ is a power of 2, and all of the computations for 3-groups have found $c(G, E_n)$ is 2 times a power of 3. This leads us to the general conjecture that if P is a p-group for some prime p, then $c(P, E_n) = (p-1)p^k$ for some k. This holds in every case that we have checked, but we have not yet been able to prove this. The following proposition seems to be a first step towards this:

Proposition 5.1. *Suppose G has nilpotency class 2, i.e. G is nonabelian and the commutator subgroup G' is contained in the center $Z(G)$. Let $n = r(G)$. Then $r(G/G') = n$, and $c(G, E_n) = c(G/G', E_n)$; this value is given by Diaconis and Graham's Theorem 3.15.*

Proof. The fact that $r(G/G') = n$ follows from the fact that G' is contained in the Frattini subgroup $\Phi(G)$ for a nilpotent group G and from Proposition 1.30. Now, we want to show that if $s, s' \in \Gamma_n(G)$ then $s \sim s'$ if and only if $\pi(s) \sim \pi(s')$, where $\pi : G \to G/G'$ is the canonical projection; this means the components of $E_n(G)$ correspond exactly to the components of $E_n(G/G')$. That $s \sim s'$ implies $\pi(s) \sim \pi(s')$ holds in general, so we need to prove the converse. If $\pi(s) \sim \pi(s')$, then the sequence of basic elementary operations in $E_n(G/G')$ giving this equivalence lifts to a sequence of basic elementary operations in $E_n(G)$ giving that s is equivalent to some s'' so that $\pi(s'') = \pi(s')$. Thus, it remains to show that if $\pi(s) = \pi(s')$ then $s \sim s'$.

Now, G' is the group generated by all commutators $[g, h] = g^{-1}h^{-1}gh$ for $g, h \in G$. It is easy to check that the following commutator identities hold in general:

$$[ab, c] = b^{-1}[a, c]b[b, c] \qquad [a, bc] = [a, c]c^{-1}[a, b]c$$

Since $G' \leq Z(G)$ in our case, these reduce to

$$[ab, c] = [a, c][b, c] \qquad [a, bc] = [a, b][a, c].$$

This identity allows us to show

$$[b, a] = [a, b]^{-1} = [a^{-1}, b] = [a, b^{-1}].$$

If $(g_1, \ldots, g_n)$ generates G, we can use these identities to write any commutator $[g, h]$ as a product of commutators $[g_i, g_j]$, and therefore get that G' is generated by $[g_i, g_j]$.

Then, we can show that if we have another generating sequence $(g_1z_1, \ldots, g_nz_n)$ (where each z_i is in $G' \leq Z(G)$), then for any indices i, j, k we can use elementary operations to get

$$(g_1z_1, \ldots, g_nz_n) \sim (g_1z_1, \ldots, g_iz_i[g_j, g_k], \ldots, g_nz_n).$$

Structure	Group	$c(G, E_2)$	Class
$Z_2 \times Z_2$	(4,2)	1	1
$Z_4 \times Z_2$	(8,2)	1	1
D_8	(8,3)	1	2
Q_8	(8,4)	1	2
$Z_4 \times Z_4$	(16,2)	2	1
$(Z_4 \times Z_2) \rtimes Z_2$	(16,3)	1	2
$Z_4 \rtimes Z_4$	(16,4)	1	2
$Z_8 \times Z_2$	(16,5)	1	1
$Z_8 \rtimes Z_2$	(16,6)	1	2
D_{16}	(16,7)	2	3
QD_{16}	(16,8)	1	3
Q_{16}	(16,9)	1	3
$(Z_4 \times Z_2) \rtimes Z_4$	(32,2)	2	2
$Z_8 \times Z_4$	(32,3)	2	1
$Z_8 \rtimes Z_4$	(32,4)	2	2
$(Z_8 \times Z_2) \rtimes Z_2$	(32,5)	1	2
	(32,6)	1	3
$(Z_8 \rtimes Z_2) \rtimes Z_2$	(32,7)	1	3
	(32,8)	1	3
$(Z_8 \times Z_2) \rtimes Z_2$	(32,9)	2	3
$Q_8 \rtimes Z_4$	(32,10)	1	3
$(Z_4 \times Z_4) \rtimes Z_2$	(32,11)	1	3
$Z_4 \rtimes Z_8$	(32,12)	1	2
$Z_8 \rtimes Z_4$	(32,13)	2	3
$Z_8 \rtimes Z_4$	(32,14)	2	3
	(32,15)	1	3
$Z_{16} \times Z_2$	(32,16)	1	1
$Z_{16} \rtimes Z_2$	(32,17)	1	2
D_{32}	(32,18)	4	4
QD_{32}	(32,19)	2	4
Q_{32}	(32,20)	2	4

Structure	Group	$c(G, E_2)$	Class
$Z_3 \times Z_3$	(9,2)	2	1
$Z_9 \times Z_3$	(27,2)	2	1
$(Z_3 \times Z_3) \rtimes Z_3$	(27,3)	2	2
$Z_9 \rtimes Z_3$	(27,4)	2	2
$Z_9 \times Z_9$	(81,2)	6	1
$(Z_9 \times Z_3) \rtimes Z_3$	(81,3)	2	2
$Z_9 \rtimes Z_9$	(81,4)	2	2
$Z_{27} \times Z_3$	(81,5)	2	1
$Z_{27} \rtimes Z_3$	(81,6)	2	2
$(Z_3)^3 \rtimes Z_3$	(81,7)	2	3
$(Z_9 \times Z_3) \rtimes Z_3$	(81,8)	2	3
$(Z_9 \times Z_3) \rtimes Z_3$	(81,9)	6	3
	(81,10)	2	3
$(Z_9 \times Z_3) \rtimes Z_9$	(243,2)	6	2
	(243,3)	18	3
	(243,4)	6	3
	(243,5)	2	3
	(243,6)	6	3
	(243,7)	2	3
	(243,8)	6	3
	(243,9)	6	3
$Z_{27} \times Z_9$	(243,10)	6	1
$Z_{27} \rtimes Z_9$	(243,11)	6	2
$(Z_{27} \times Z_3) \rtimes Z_3$	(243,12)	2	2
	(243,13)	6	3
$(Z_9 \times Z_3) : Z_9$	(243,14)	6	3
$(Z_9 \times Z_3) : Z_9$	(243,15)	6	3
$(Z_{27} \rtimes Z_3) \rtimes Z_3$	(243,16)	2	3
$(Z_9 \times Z_3^2) \rtimes Z_3$	(243,17)	6	3
$(Z_9 \rtimes Z_3) \rtimes Z_3$	(243,18)	2	3
$(Z_{27} \times Z_3) \rtimes Z_3$	(243,19)	2	3
$(Z_{27} \times Z_3) \rtimes Z_3$	(243,20)	2	3
$Z_9 \rtimes Z_{27}$	(243,21)	2	2
$Z_{27} \rtimes Z_9$	(243,22)	2	3
$Z_{81} \times Z_3$	(243,23)	2	1
$Z_{81} \rtimes Z_3$	(243,24)	2	2
$(Z_9 \times Z_9) \rtimes Z_9$	(243,25)	6	4
$(Z_9 \times Z_9) \rtimes Z_9$	(243,26)	18	4
	(243,27)	6	4
$(Z_9 \rtimes Z_9) \rtimes Z_3$	(243,28)	6	4
	(243,29)	6	4
	(243,30)	6	4

Table 1: Data for small 2-groups and 3-groups

If $i \neq j, k$ then we can use right operations to build $[g_j z_j, g_k z_k]$ to the right of $g_i z_i$, and note that $[g_j z_j, g_k z_k] = [g_j, g_k]$ because $z_j, z_k \in Z(G)$. If $i = k$ and $i \neq j$, then we can use left and right operations to replace $g_i z_i$ by

$$(g_j z_j)^{-1}(g_i z_i)(g_j z_j) = (g_j^{-1} g_i g_j) z_i = (g_i [g_i, g_j]) z_i = g_i z_i [g_i, g_j].$$

So, we can build an arbitrary commutator $[g_i, g_j]$ in an arbitrary position of a generating sequence $(g_1 z_1, \ldots, g_n z_n)$. Since each z_i is in G' and hence a product of these commutators $[g_i, g_j]$, we can use this process repeatedly to build z_i^{-1} in the i-th spot for each i. Thus, we get $(g_1, \ldots, g_n) \sim (g_1 z_1, \ldots, g_n z_n)$ under $E_n(G)$. This proves that if $\pi(s) = \pi(s')$ then $s \sim s'$, which finishes the proof. □

Note that no such theorem can hold for nilpotency class 3; we have examples where it fails (as the dihedral group D_{16} has nilpotency class 3 and has more components than its abelianization, as does the group denoted (81,9) by GAP). Understanding the behavior of the groups D_{16}, QD_{16}, Q_{16}, (81,7), (81,8), (81,9), and (81,10) seems to be the next step in gaining a better theoretical understanding of the components of $E_n(G)$ for p-groups. These seven groups all have nilpotency class 3, and therefore necessarily have similar structures regarding centers and commutators. One would hope that understanding what makes D_{16} and (81,9) behave differently than the others would lead to a theoretical understanding of the values of $c(G, E_n)$ for p-groups G.

5.3 Other Computations

These computational techniques also shed light on various other questions regarding the actions of $E_n(G)$ and $L_n(G)$ on $\Gamma_n(G)$. One such question is whether the orbits of these actions are all of the same size. We note that the determinant functions of Section 5.1 allow us to prove that the orbits of the action of $L_n(G)$ are all the same size if G is abelian, dihedral, or nonabelian of order pq. In each of these cases, the determinant function allows us to explicitly describe the orbits and check that they have the same cardinality.

We can computationally check the size of the orbits of $E_n(G)$ or $L_n(G)$ for any given finite group G, as an extension of the method for computing the number of orbits. Indeed, for each solvable group G that I have tested, the orbits of $E_n(G)$ are all the same size, as are the orbits of $L_n(G)$. This leads us to conjecture that this is true for all solvable groups G, and in particular for the special case of p-groups. One would expect that techniques for investigating $c(G, E_n)$ for p-groups would also yield information about the uniformity of orbit sizes. We could also hope to prove this uniformity in some cases (or perhaps in general) by finding a more general construction for determinant functions.

However, we note that that the orbits of $E_n(G)$ or $L_n(G)$ are not uniformly sized for every group. In fact, computations show that there are differently sized components for even the simplest nonsolvable groups. Table 5.3 contains the list of component sizes for $E_n(G)$ and $L_n(G)$ for a few of the smallest nonabelian simple groups, a few small groups with A_5 in their Jordan-Holder series, and $A_5 \times A_5$. An expression such as "50 ($\times$24)" in the list of component sizes denotes that there are 24 components of size 50.

Another question we can ask is, given a group G and a quotient group $\overline{G}$, if $c(\overline{G}, E_n)$ necessarily divides $c(G, E_n)$ (or if $c(\overline{G}, L_n)$ divides $c(G, L_n)$). By Proposition 3.6, we know $c(H, E_n) \leq c(G, E_n)$ and $c(H, L_n) \leq c(G, L_n)$.

Note that if our conjecture about the number of orbits of p-groups (from the previous section) is satisfied, then if P is a p-group and $\overline{P}$ a quotient (which is also a p-group), $c(P, E_n) = p^n(p-1)$ and $c(\overline{P}, E_n) = p^m(p-1)$ for $m \leq n$, which means $c(\overline{P}, E_n)$ does indeed divide $c(P, E_n)$ (and the same holds for L_n).

The data for nonsolvable groups in Table 5.3 is striking. Note $\mathrm{SL}(2,5)$, $Z_2 \times A_5$, and $\mathrm{GL}(2,4)$ all have A_5 as quotients (while S_5 does not). If G is one of these three groups, then not only do we have $c(G, E_2) = c(A_5, E_2)$ and $c(G, L_2) = c(A_5, L_2)$, but for each G the orbit sizes are exactly integer multiples of the orbit sizes of A_5 (the multiple is $[\![G : A_5]\!]_2$ in each case). This suggests that the projection onto quotient groups behaves uniformly with respect to the orbits.

Group	$c(G, E_2)$	Component Sizes	$c(G, L_2)$	Component Sizes
A_5	3	$600\ (\times 2), 1080$	44	$50\ (\times 24), 54\ (\times 20)$
$\mathrm{PSL}(2,7)$	5	$1176\ (\times 2), 5376\ (\times 2), 6048$	188	$49\ (\times 48), 108\ (\times 56), 128\ (\times 84)$
A_6	10	$5400\ (\times 4), 7200\ (\times 2),$ $8640\ (\times 2), 11520\ (\times 2)$	792	$75\ (\times 288), 96\ (\times 180),$ $100\ (\times 144), 128\ (\times 180)$
$\mathrm{PSL}(2,8)$	7	$27216\ (\times 4), 35280\ (\times 3)$	440	$486\ (\times 224), 490\ (\times 216)$
S_5	3	$1800, 2160, 2880$	74	$75\ (\times 24), 96\ (\times 30), 108\ (\times 20)$
$\mathrm{SL}(2,5)$	3	$2400\ (\times 2), 4320$	44	$200\ (\times 24), 216\ (\times 20)$
$Z_2 \times A_5$	3	$1800\ (\times 2), 3240$	44	$150\ (\times 24), 162\ (\times 20)$
$\mathrm{GL}(2,4)$	3	$4800\ (\times 2), 8640$	44	$400\ (\times 24), 432\ (\times 20)$
$A_5 \times A_5$	9	$324000\ (\times 4), 648000\ (\times 4),$ 1036800	1936	$2250\ (\times 576), 2592\ (\times 400),$ $2700\ (\times 960)$

Table 2: Component sizes for small nonsolvable groups

Moreover, the data for $A_5 \times A_5$ shows that $c(A_5^2, E_n) = c(A_5, E_n)^2$ and $c(A_5^2, L_n) = c(A_5, L_n)^2$. The orbits for A_5^2 seem to match up with pairs of orbits for A_5. For instance, for the orbits under elementary operations, there are four orbits of size $324000 = 540 * 600$ (which correspond to pairs of orbits of size 600) there are four orbits of size $648000 = 600 * 1080$ (which correspond to pairs including the orbit of size 1080 and one of the orbits of size 600), and one orbit of size $1036800 = 960 * 1080$, corresponding to the single pair of two copies of the orbit of 1080. From this data, it seems clear that there is something interesting going on, though so far we can't prove much about it. (Note however we do not believe that $c(G^k, E_n)$ is always the k-th power of $c(G, E_n)$, and in fact we suspect that $c(G/N, E_n)$ does not always divide $c(G, E_n)$. A good place to look for a counterexample would be taking $G = A_5^{19}$, the largest power of A_5 that is generated by two elements.)

References

[AC65] J. J. Andrews and M. L. Curtis, *Free groups and handlebodies*, Proc. Amer. Math. Soc. **16** (1965), 192–195.

[Ber86] Ya. G. Berkovich, *Finite groups with a minimal set of generators of maximum length (Russian)*, Publ. Math. Debrecen **33** (1986), no. 3-4, 329–332.

[BLM05] Alexandre Borovik, Alexander Lubotzky, and Alexei Myasnikov, *The finitary andrews-curtis conjecture*, Infinite Groups: Geometric, Combinatorial and Dynamical Aspects, Progress in Mathematics, vol. 248, Birkhäuser Verlag Basel, 2005, pp. 15–30.

[BS] Stanley Burris and H Sankappanavar, *A course in universal algebra*, Available at `http://www.math.uwaterloo.ca/~snburris/htdocs/ualg.html`.

[CLGM+95] Frank Celler, Charles R. Leedham-Green, Scott H. Murray, Alice C. Niemeyer, and E. A. O'Brien, *Generating random elements of a finite group.*, Comm. Algebra **23** (1995), no. 13, 4931–4948.

[CP00] Gene Cooperman and Igor Pak, *The product replacement graph on generating triples of permutations*, Available at `http://www.math.ucla.edu/~pak/papers/research.htm`, 2000.

[Dav93] C. David, *T_3-systems of finite simple groups*, Rend. Sem. Mat. Univ. Padova **89** (1993), 19–27.

[Den09] R. Keith Dennis, *Generating sequences of finite groups*, Preprint, March 2009.

[DF04] David S. Dummit and Richard M. Foote, *Abstract algebra*, 3rd ed., John Wiley and Sons, Inc, 2004.

[DG99] Persi Diaconis and Ronald Graham, *The graph of generating sets of an abelian group*, Colloquium Mathematicum **80** (1999), no. 1, 31–38.

[DSC98] P. Diaconis and L. Saloff-Coste, *Walks on generating sets of groups.*, Invent. Math. **134** (1998), no. 2, 251–299.

[Dun70] M. J. Dunwoody, *Nielsen transformations*, Computational Problems in Abstract Algebra (Proc. Conf., Oxford, 1967), Pergamon, Oxford, 1970, pp. 45–46.

[Eva93] Martin J. Evans, *T-systems of certain finite simple groups*, Math. Proc. Cambridge Philos. Soc. **113** (1993), no. 1, 9–22.

[FJ08] Michael D. Fried and Moshe Jarden, *Field arithmetic*, third ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 11, Springer-Verlag, Berlin, 2008.

[GAP08] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.4.12*, 2008, (`http://www.gap-system.org`).

[Gar08] Shelly Garion, *Connectivity of the product replacement algorithm in $PSL(2,q)$*, J. Group Theory **11** (2008), 765–777.

[Gas55] Wolfgang Gaschütz, *Zu einem von B. H. und H. Neumann gestellten Problem*, Math. Nachr. **14** (1955), 249–252.

[Gil77] Robert Gilman, *Finite quotients of the automorphism group of a free group*, Canad. J. Math. **29** (1977), no. 3, 541–551.

[GK00] Robert M. Guralnick and William M. Kantor, *Probabilistic generation of finite simple groups*, J. Algebra **234** (2000), no. 2, 743–792.

[GP03] Robert Guralnick and Igor Pak, *On a question of b. h. neumann.*, Proc. Amer. Math. Soc. **131** (2003), no. 7, 2021–2025.

[Hal36] Philip Hall, *The Eulerian functions of a group*, Quarterly Journal of Mathematics **7** (1936), 134–151.

[Hal37] ———, *Complemented groups*, Journal of the London Mathematical Society **12** (1937), 205–208.

[Lev33] Friedrich Levi, *Über die Untergruppen der freien Gruppen*, Math. Z. **37** (1933), no. 1, 90–97.

[Lyn59] Roger C. Lyndon, *Properties preserved in subdirect products*, Pacific J. Math. **9** (1959), 155–164.

[NN51] Bernhard H. Neumann and Hanna Neumann, *Zwei Klassen charakteristischer Untergruppen und ihre Faktorgruppen*, Math. Nachr. **4** (1951), 106–125.

[Pak99] Igor Pak, *What do we know about the product replacement algorithm?*, Groups and computation, III, Ohio State Univ. Math. Res. Inst. Publ., vol. 8, 1999, pp. 301–347.

[Rot95] Joseph Rotman, *An introduction to the theory of groups*, Springer, 1995.

[Tar75] Alfred Tarski, *An interpolation theorem for irredundant bases of closure structures*, Discrete Math. **12** (1975), 185–192.

[vLW03] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, Cambridge University Press, 2003.

[Whi00] Julius Whiston, *Maximal independent generating sets of the symmetric group.*, J. Algebra **232** (2000), no. 1, 255–268.