GROUPS OF ORDER p^n

ALAN LEE, ANUJ THAKUR

1. Summary

It is often useful to quantify the number of groups of a certain order for many reasons such as finding an isomorphism between two groups. In this paper, we estimate the number of groups of order p^n, where p is prime and n is a positive integer. We come up with an expression that, as n goes to infinity, behaves like a function in terms of p and n. Before we proceed with the key result and proof, let us first define a few terms.

Definition 1.1. Let be p a prime. A group G is a *p-group* if all of its elements have order p^n for some integer n.

Theorem 1.2. *All groups of order p^n are p-groups.*

Proof. All groups of order p^n must only have elements with orders that divide p^n. These groups are p-groups because the order of each element is a divisor of the order of the group, and the only numbers that divide p^n are powers of p that are at most n. Hence, throughout this paper, we use "p-group" and "group of order p^n" interchangeably.

□

Definition 1.3. The *index* of a subgroup H in a group G, denoted $[G : H]$, is the number of cosets of the subgroup in G.

Definition 1.4. *Big O notation* is used to represent the limiting behavior of a function by determining the order of the function. We write $f(x) = O(g(x))$ for some function $g(x)$ if there exists real numbers M and x_0 such that $|f(x)| \leq Mg(x)$ for all $x \geq x_0$.

For asymptotics, we often look at a function's big O notation in order to find out which term "dominates". We often omit the constant term in big O notation because the M above can just be multiplied by the constant instead.

Example. We can find the big O notation for the function $f(x) = 3x^4 - 6x^3 + 13x - \log x - 15$. To do so, we first may omit the terms that do not contribute to the highest growth rate, leaving only $3x^4$. By omitting the coefficient, we have $f(x) = O(x^4)$.

Throughout the course of the paper, we will be proving the following statements:

First, we will show how Graham Higman proved that the number of p-groups with order p^n, denoted $f(n, p)$, must be greater than or equal to $p^{\frac{2}{27}\left(n^3-6n^2\right)}$. Higman also showed an upper bound that approached $p^{\frac{2}{15}n^3}$, but since the next result offers a stronger bound, we

Date: June 4, 2020.

will not cover this case.

Second, we will show how Charles Sims proved the upper bound of $f(n,p)$ to be $p^{\frac{2}{27}n^3+O\left(n^{\frac{8}{3}}\right)}$. Together, this will yield the following formula, the centerpiece of this paper:

Higman-Sims asymptotic formula. *The number of p-groups of order p^n is*

$$f(n,p) = p^{\frac{2}{27}n^3+O(n^{\frac{8}{3}})}$$

[Kan90]. Since the second part of the exponent is in big-O notation, as n (not p) becomes very large, $f(n,p)$ will begin to behave like $p^{\frac{2}{27}n^3+cn^{8/3}}$ for some constant c.

After proving these results, we will assess this formula for smaller primes p and smaller numbers n for which the number of groups with order p^n is already known.

2. Lower Bound

The method Higman used to prove the lower bound was by taking a chain of inequalities, with the largest value being equal to the number of p-groups of order p^n and the smallest being a value we can enumerate. He proved the following theorem first, which enumerates the number of p-groups with a central elementary abelian subgroup of index p^r and order p^s. We shall denote this value as $g(r,s:p)$(note the variables we use in this paper(besides p) have no meaningful significance).

Theorem 2.1. *If $s > \frac{1}{2}r\,(r+1)$, then $g(r,s:p) = 0$. Otherwise, we have*

$$p^{\frac{1}{2}sr(r+1)-s^2-r^2} \leq g(r,s:p) \leq p^{\frac{1}{2}sr(r+1)-s^2+s}.$$

The proof of this theorem is outlined in [Hig60].

□

We now use this theorem to prove the remainder of the lower bound. First, define $f_2(n,p)$ to be the number of p-groups of order p^n that have central elementary abelian subgroups containing an elementary abelian quotient group. When $r+s=n$, $f_2(n,p) \geq g(r,s:p)$ because all groups counted by the latter are also in the former. Now, we substitute $r = \frac{2n-\delta}{3}$ and $s = \frac{n+\delta}{3}$ such that $\delta \in \{0,1,2\}$ and $r,s \in \mathbb{Z}$.

Comparing s and $\frac{1}{2}r\,(r+1)$ will show us how we can apply Theorem 2.1 to $g(r,s:p)$. We set

$$s = \frac{n+d}{3}$$

and

$$\frac{1}{2}r\,(r+1) = \frac{1}{2}\frac{(2n-\delta)(2n+2\delta)}{9} = \frac{(2n-\delta)(n+\delta)}{9} = \frac{2n^2+n\delta-\delta^2}{9}.$$

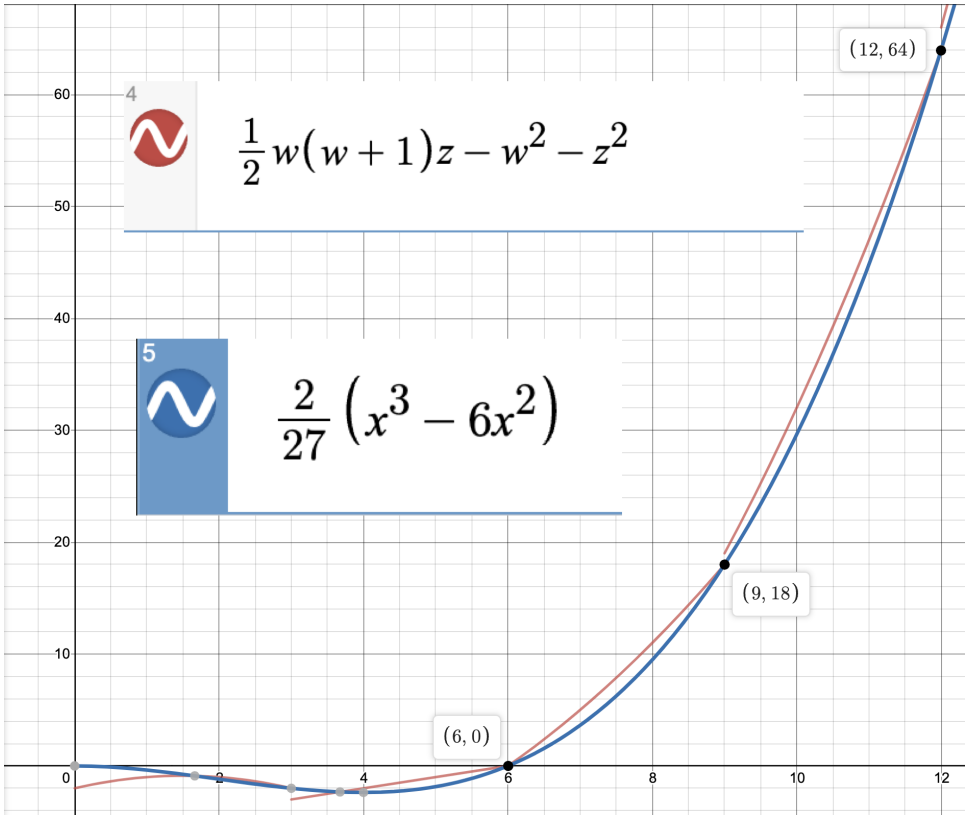


Figure 1. Comparing the asymptotic lower bound with the given inequality's exponent

Since there is a term of degree 2 on the $\frac{1}{2}r\,(r+1)$ side, these values of r and s do not satisfy $s > \frac{1}{2}r\,(r+1)$ for large values of n, namely $n \geq 3$. Thus, we conclude that

$$p^{\frac{1}{2}sr(r+1)-s^2-r^2} \leq g(r,s:p) \leq f_2(n,p).$$

Now substituting our values for r and s, our inequality becomes

$$p^M \leq f_2(n,p)$$

where M varies depending on δ.

If $\delta = 0$:

$$M = \frac{n}{3} \times \frac{n}{3} \times \frac{2n+3}{3} - \frac{5n^2}{9} = \frac{2n^3}{27} - \frac{12n^2}{27} = \frac{2}{27}\left(n^3 - 6n^2\right).$$

If $\delta = 1$:

$$M = \frac{1}{2}(2n-1)(n+1)\frac{2n+2}{27} - \frac{4n^2-4n+1}{9} - \frac{n^2+2n+1}{9}$$

$$= 2n^3 - 12n^2 + \frac{6n-7}{27}$$

$$= \frac{2}{27}\left(n^3 - 6n^2\right) + \frac{2}{9}\left(n - \frac{7}{6}\right).$$

If $\delta = 2$:

$$M = (n-1)(n+2)(2n+1)/27 - \left(4n^2 - 8n + 4 + n^2 + 4n + 4\right)/9$$
$$= \frac{(2n^3 - 12n^2 + 9n - 26)}{27}$$
$$= \frac{2}{27}\left(n^3 - 6n^2\right) + \frac{1}{3}\left(n - \frac{26}{9}\right).$$

All three of these values are greater or equal to $\frac{2}{27}\left(n^3 - 6n^2\right)$ for $n \geq 6$. Graphing the values as in **Figure 1** further supports this- whenever the blue line approaches the red line, the red line "kinks" upwards, and hence is always greater than or equal to the blue line (note that in the diagram, $w = r$, $z = s$, and $x = n$).

It now follows that $f_2(n,p) \geq p^{\frac{2}{27}\left(n^3-6n^2\right)}$. Since all groups in $f_2(n,p)$ are also in $f(n,p)$, the number of p-groups with order p^n, as our final lower bound, we have the following chain of inequalities, and hence our lower bound

$$f(n,p) \geq f_2(n,p) \geq g(r,s:p) \geq p^{\frac{2}{27}n^3 + O\left(n^2\right)}.$$

□

3. Upper Bound

Charles Sims proved part of the upper bound by using results from anti-symmetric bilinear maps and generators in [Sim65]. Because we are trying to find an upper bound, it makes sense to take derivatives to find critical points and maximums. This is what we do in the following section. Note that is in the lower bound the choice of our variables x, y, z doesn't have any significance. The results mentioned previously yielded an upper bound for p^M for

$$M = \frac{x^2(1-x-z)}{2} + \left(yz - \frac{z^2}{2}\right)(1-x-y) + \frac{z(1-x-y)^2}{2} + O(n^{\frac{1}{3}}),$$

where $x + y \leq 1$ and $0 \leq z \leq \min(x,y)$. Taking derivatives helps us find critical points so doing that with respect to y we get the derivative of

$$\frac{x^2(1-x-z)}{2} + \left(yz - \frac{z^2}{2}\right)(1-x-y) + \frac{z(1-x-y)^2}{2} + O(n^{\frac{1}{3}}),$$

which is

$$\frac{d}{dy}(-y-x+1) * \left(zy - \frac{z^2}{2}\right) + \frac{z}{2}\frac{d}{dy}(-y-x+1)^2 + \frac{d}{dy}\left(\frac{x^2(-z-x+1)}{2}\right)$$

$$= \left(-\frac{d}{dy}(y) + \frac{d}{dy}(-x) + \frac{d}{dy}(1)\right)\left(zy - \frac{z^2}{2}\right) + (-y - x + 1)\left(z\frac{d}{dy}(y) + \frac{d}{dy} - \frac{z^2}{2}\right)$$

$$+(-y - x + 1)(z)\left(-\frac{d}{dy}(y) + \frac{d}{dy}(-x) + \frac{d}{dy}(1)\right)$$

$$= (-1)\left(zy - \frac{z^2}{2}\right) + (-y - x + 1)(z(1) + 0) + (-y - x + 1)(z)(-1)$$

$$= \frac{z^2}{2} - yz$$

$$= -\left(yz - \frac{z^2}{2}\right) \leq 0.$$

This is the completes part of our proof.

Now that we have this we assume that $y = z$ and thus $y \leq x$.

If we let M be represented as $A(x, y, z)$ because it is a function of three variables we can take $B(x, y)$ such that $B(x, y) = A(x, y, y)$. This gives $B(x, y) = x^2(1 - x - y)/2 + y^2(1 - x - y)/2 + y(1 - x - y)^2/2$.

If we now take the derivative of this expression with respect to y we get $\frac{1-2x}{2} - y(1 - x)$. This is equal to 0 when $x \neq 1$ and $y = \frac{1-2x}{2(1-x)}$.

Now we want to simplify this. In our expression with two variables $B(x, y)$ we saw the expression $1 - x - y$. We make a substitution and call it u. We now have some function $C(x, u) = B(x, y) = (u(x^2 + (1 - x - u)^2 + u - ux - u^2)/2$. We now take the derivative of this with respect to u which is $(u(2x - 2(1 - x - u) - u))/2$. We note that if $u = 0$ then C is 0. We assume that $u > 0$ so the derivative is 0 when $4x + u - 2 = 0$(or $y = 3x - 1$).

This means that there is only a critical point of $B(x, y)$ when $y = 3x-1$. This is equivalent to $x = \frac{5-\sqrt{7}}{6}$ and $y = \frac{3-\sqrt{7}}{6}$. Now we plug this into $B(x, y)$ and we actually get $\frac{7-\sqrt{17}}{27}$ which is indeed less than $\frac{2}{27}$.

We aren't completely done yet but we are definitely close. We have to check that the maximum of $B(x, y)$ is less than or equal to $\frac{2}{27}$ with the bounds that $x + y = 1$, $y = 0$, or $x = y$. For the first case we have $B(x, y) = 0 + 0 + 0 = 0$. The second case we have $B(x, 0) = \frac{x^2(1-x)}{2}$ with $0 \leq x \leq 1$. This achieves a maximum when $x = \frac{2}{3}$ so the maximum is $\frac{2}{27}$. Finally we have $B(x, x) = \frac{x(1-2x)}{2}$ which has a max of $\frac{1}{16}$ when $x = \frac{1}{4}$.

Hence, we have that $B(x, y) \leq \frac{2}{27}$ and the upper bound is complete as well. Thus, we have the general result of the Higman-Sims formula.

□

4. Assessing the Formula

Since it is hard to enumerate the exact number of p-groups by hand, not many exact results for these numbers exist as of today. However, we can see the Higman-Sims asymptotic formula in action for a small prime such as $p = 2$. We are able to obtain the number of 2-groups for $n \leq 10$, and we can then compare these numbers to the formula $2^{\frac{2}{27}x^3}$. Before doing so, however, we shall take the logarithm base 10 of both sides in order to work with smaller numbers.

Next, we examine the rates of increase between consecutive integers between the two equations. In effect, for $h(x) = \log(2^{\frac{2}{27}x^3})$, we compare the following for each x we have data for:

$$\frac{h(x)}{h(x-1)} - \frac{\log(\text{number of groups with order } 2^x)}{\log(\text{number of with groups order } 2^{x-1})}.$$

An example would be when $x = 5$. Given that there are 51 groups of order $2^5 = 32$ and 14 groups of order $2^4 = 16$, we examine the above value:

$$\frac{h(5)}{h(4)} - \frac{\log(51)}{\log(14)} \approx 0.4633.$$

All of this is summarized in graphical format in **Figure 2**. The black line represents the logarithm of the asymptotic formula for $p = 2$ and the black dots represent the actual number of groups for each 2^x. While it appears that the black line is increasing much faster than the line drawn by the black points, this soon reverses for larger values of x. Indeed, the blue dots (which represent the difference between the rates of change between the two black figures- dots and lines) are already becoming negative for $x \geq 9$.

This negative trend indicates that the differences in the rates of change between the two black figures (line and points) are approaching each other, showing that the number of p-groups of order p^n begins to behave like the Higman-Sims asymptotic formula for larger values of n.

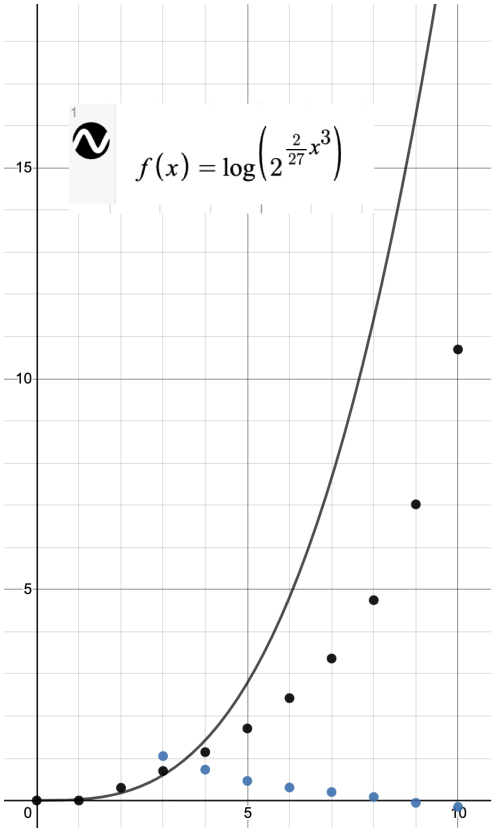
This test shows that the Higman-Sims asymptotic formula is indeed a good approximation for $p = 2$. Other small primes such as $p = 5$ and $p = 7$ also follow the same pattern: a decreasing sequence of blue dots. Larger primes are harder to test out as not many numbers k have known values of how many groups of order p^k there are, especially for larger p.

5. More Advances in Enumerating Groups

After this significant formula was discovered regarding p-groups in 1965, 4 years later, Peter Neumann devised a formula that applied to enumerating *all* groups, whose proof can be found in [Neu69].

Theorem 5.1. *The number of groups with order n is less than*

$$n^{\frac{1}{2}(\log_2 n)^2}.$$



x_2	y_2	y
0	$\log 1$	------
1	$\log 1$	------
2	$\log 2$	------
3	$\log 5$	$\frac{f(3)}{f(2)} - \frac{\log 5}{\log 2}$
4	$\log 14$	$\frac{f(4)}{f(3)} - \frac{\log 14}{\log 5}$
5	$\log 51$	$\frac{f(5)}{f(4)} - \frac{\log 51}{\log 14}$
6	$\log 267$	$\frac{f(6)}{f(5)} - \frac{\log 267}{\log 51}$
7	$\log 2328$	$\frac{f(7)}{f(6)} - \frac{\log 232}{\log 26}$
8	$\log 56092$	$\frac{f(8)}{f(7)} - \frac{\log 560}{\log 23}$
9	$\log 10494213$	$\frac{f(9)}{f(8)} - \frac{\log 104}{\log 5}$
10	$\log 49487365422$	$\frac{f(10)}{f(9)} - \frac{\log 49}{\log}$

Figure 2. Comparing actual number of p-groups to the estimated numbers

However, as can easily be seen by testing $n = 1024$, when the above formula yields a maximum of 1024^{50}, compared to the actual value of $49,487,365,422 \approx 1024^{3.553}$, and due to the broadness of this theorem, it is not always a strong upper bound. Rather, in 1993, a stronger analogous statement for non-prime powers was shown by László Pyber in [Pyb93].

In the future, if we are able to accurately measure the number of p-groups for higher powers of p, these formulae will be available as measures of accuracy.

References

[Hig60] Graham Higman. Enumerating p-groups. i: Inequalities. *Proceedings of the London Mathematical Society*, 3(1):24–30, 1960.

[Kan90] William M Kantor. Some topics in asymptotic group theory. *Groups, Combinatorics and Geometry (Durham*, pages 403–421, 1990.

[Neu69] Peter M Neumann. An enumeration theorem for finite groups. *The Quarterly Journal of Mathematics*, 20(1):395–401, 1969.

[Pyb93] László Pyber. Enumerating finite groups of given order. *Annals of Mathematics*, pages 203–220, 1993.

[Sim65] Charles C Sims. Enumerating p-groups. *Proceedings of the London Mathematical Society*, 3(1):151–166, 1965.