JOURNAL OF ALGEBRA 115, 414–430 (1988)

A Generalization of Sylow's Third Theorem

KENNETH S. BROWN*

Department of Mathematics, White Hall, Cornell University, Ithaca, New York 14853

AND

JACQUES THÉVENAZ†

Département de Mathématiques, Ecole Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cedex 05, France

Communicated by Wilberd van der Kallen

Received February 9, 1987

Let G be a finite group and p a prime divisor of $|G|$. Sylow's third theorem says that the number of Sylow p-subgroups of G is congruent to 1 modulo p. This was generalized by Frobenius [6], who proved that if p^k divides $|G|$, then the number of subgroups of G of order p^k is congruent to 1 modulo p. Another result of Frobenius [7], which does not generalize Sylow's theorem but which is in some sense analogous to it, says that if n is any divisor of $|G|$, then the number of *elements* of G of order dividing n is divisible by n. More recently, there have been analogues of Sylow's theorem involving *chains* of subgroups. For example, it was shown by Brown [3] that the Euler characteristic of the poset of non-trivial p-subgroups of G is congruent to 1 modulo the highest power of p dividing $|G|$. A similar result concerning the poset of *all* proper, non-trivial subgroups of G was obtained by Kratzer and Thévenaz [10]; more generally, they studied for any subgroup H of G the Euler characteristic of the "interval" (H, G), consisting of all subgroups K such that $H < K < G$.

The purpose of this paper is to prove a very general theorem of this type, which has all the results stated above as corollaries. Let H be a proper subgroup of G and let D be a set of divisors of the index $|G: H|$. We study the Euler characteristic of the poset $S_D(H, G)$ consisting of all subgroups K such that $H \subseteq K \subseteq G$ and $|K: H| \in D$. For simplicity, we will confine ourselves in this introduction to the case $H = \{1\}$, in which case we write

* Partially supported by NSF Grant DMS-8502278.
† Partially supported by the Swiss NSF.

0021-8693/88 $3.00

$S_D(G)$ instead of $S_D(\{1\}, G)$. Our theorem in this case takes the form $\chi(S_D(G)) \equiv 1 \bmod m$, where m is a certain divisor of $|G|$ (depending on G and D). The precise definition of the modulus m is somewhat technical and will be deferred to Section 1. But here are some easily stated special cases:

(A) If all the elements of D are powers of a single prime p, then $\chi(S_D(G)) \equiv 1 \bmod p^{|D|}$. (If D consists of a single power of p, for instance, then we recover Frobenius's generalization of Sylow's theorem. At the other extreme, if D consists of all non-trivial powers of p that divide $|G|$, then we recover Brown's theorem.)

(B) Let n be a divisor of $|G|$ and let D be the set of non-trivial divisors of n. Then $\chi(S_D(G)) \equiv 1 \bmod n$. (This result is similar to the Frobenius theorem on the number of elements of G of order dividing n. In fact, the Frobenius theorem is an easy corollary of the analogue of (B) for $S_D(H, G)$ with H not necessarily trivial; see 1.6 below.)

(C) Let n be a divisor of $|G|$ and let D be the set of proper divisors of $|G|$ which do not divide n. Let G_{ab} be the abelianization of G. If $|G_{ab}|$ is relatively prime to $|G|/n$, then $\chi(S_D(G)) \equiv 1 \bmod |G|/n$.

The paper is organized as follows. In Section 1 we state the main theorem and discuss some special cases, including (A), (B), and (C) above. In Section 2 we prove the theorem when D consists of all proper, non-trivial divisors of $|G: H|$ which are powers of a fixed prime p. It turns out that the method used by Brown (when $H = \{1\}$) goes through with no essential change, except in one case; but in this case we can complete the proof by means of a slight variation on an argument of Thévenaz [17]. As a by-product of this proof, we obtain an improvement of the Kratzer–Thévenaz theorem.

We complete the proof in Section 3 by means of a simple inductive argument, using the result of Section 2 to start the induction. Finally, we give in Section 4 some results on $\chi(S_D(H, G))$ which take into account the *structure* of G and D rather than just their cardinalities. For example, we compute $\chi(S_D(G)) \bmod p^{|D|+1}$ in the situation of (A) and obtain, as a special case, P. Hall's generalization of Kulakoff's theorem. There are two short appendices related to Section 4: Appendix A discusses a result of Stanley [13], and Appendix B contains a lemma about p-groups.

This paper is a revised version of [15]. In particular, the results (A), (B), and (C) stated above are Theorems A, B, and C of [15]. After proving our main theorem, we learned of the work of Hawkes, Isaacs, and Özaydin, whose paper [9] contains two extensions of the results of [15], which are both special cases of our main theorem: (i) they prove the analogue of (B) for $S_D(H, G)$ with H not necessarily trivial and show that it implies the Frobenius theorem (see 1.4 and 1.6 below); and (ii) they prove a generalization of (C) which was conjectured in [15] (see 1.5 below).

1. Statement of the Main Theorem

We continue with the notation of the introduction. Thus G is a finite group, H is a proper subgroup, D is a set of divisors of $|G: H|$, and $S_D(H, G)$ is the poset of subgroups $K \supseteq H$ with $|K: H| \in D$. To avoid trivialities, we will assume that D consists of *proper*, *non-trivial* divisors of $|G: H|$. (Otherwise, $S_D(H, G)$ has a largest or smallest element and hence has Euler characteristic 1 for trivial reasons.) Throughout this paper the set of positive integers is viewed as a poset, ordered by divisibility, and intervals of integers are to be understood in the sense of this ordering. Thus the closed interval $[m, n]$ is the set of positive integers k such that $m \,|k|\, n$, and the open interval (m, n) and half-open interval $(m, n]$ are defined similarly. With this notation, D is a subposet of $(1, |G: H|)$. Similarly, $S_D(H, G)$ is a subposet of (H, G).

Let $\mu_D(H, G)$ be the *reduced Euler characteristic* $\chi(S_D(H, G)) - 1$ of this poset.[1] Equivalently, $\mu_D(H, G) = \sum_{q \geq 1} (-1)^q N_q$, where N_q is the number of chains $H = K_0 < \cdots < K_q = G$ with $K_i \in S_D(H, G)$ for $0 < i < q$. Note that $\mu_D(H, G)$ can also be described in terms of Möbius functions. Namely, it is the Möbius function associated to the poset $S_D(H, G) \cup \{H, G\}$, evaluated at the pair (H, G) (cf. [14, 3.8.5]).

Our goal, as stated in the introduction, is to prove a congruence of the form $\chi(S_D(H, G)) \equiv 1 \bmod m$. In terms of μ, this congruence simply says that $\mu_D(H, G)$ is divisible by m. It turns out that the statement of the result (i.e., the definition of m) is simpler if we weight $\mu_D(H, G)$ according to the number of G-conjugates of H. Thus, letting $N(H)$ be the normalizer of H in G, we set

$$\begin{aligned} \nu_D(H, D) &= |G: N(H)| \cdot \mu_D(H, G) \\ &= \sum_{J \sim H} \mu_D(J, G), \end{aligned}$$

where $\sim$ denotes G-conjugacy. In case $H = \{1\}$, we have $\nu_D(H, G) = \mu_D(H, G) = \chi(S_D(G)) - 1$.

An element $d \in D$ will be called a *cone point* of D if $\operatorname{lcm}(d, d') \in D$ for all $d' \in D$. (The terminology here comes from the fact that any such d induces a "conical contraction" of the poset D; cf. [12, 1.5].) By a *weak cone point* of D we will mean an element $d \in D \cup \{|G: H|\}$ such that $\operatorname{lcm}(d, d') \in D \cup \{|G: H|\}$ for all $d' \in D$. We will only be interested in cone points and weak cone points which are powers of a prime p. Note, in this case, that the only way d can be a weak cone point without being a cone point is if $d = |G: H|_p$, where the latter is the p-part of $|G: H|$.

[1] Recall that the Euler characteristic $\chi(S)$ of a finite poset S is defined to be the Euler characteristic of the simplicial complex $\Delta(S)$ whose simplices are the chains in S.

We are now ready to state our main theorem on the divisibility properties of $v_D(H, G)$. For this purpose we may obviously fix a prime p and consider divisibility by powers of p. Recall that G is said to be *p-perfect* if G admits no non-trivial p-group quotient, or, equivalently, if $|G_{ab}|$ is relatively prime to p. More generally, G is *p-perfect* mod H if the quotient of G by the normal closure of H is p-perfect.

(1.1) THEOREM. *Let c be the number of powers of p which are cone points of D. Then $v_D(H, G)$ is divisible by p^c. If G is p-perfect* mod H, *then $v_D(H, G)$ is divisible by p^w, where w is the number of powers of p which are weak cone points of D.*

(Note that, by a remark above, we have $w = c$ or $c + 1$. So the second statement improves the exponent c by at most 1 if G is p-perfect mod H. Note also that $|G: N(H)|$ divides $v_D(H, G)$ by definition, so the theorem is vacuous unless p^c or p^w is larger than $|G: N(H)|_p$. Thus the theorem is really about divisibility of $\mu_D(H, G)$ by a certain divisor of $|N(H): H|$; in particular, it is vacuous if $N(H) = H$.)

Suppose, for instance, that D consists of the entire interval $(1, |G: H|)$, and let $|G: H|_p = p^r$. Then $c = r - 1$ and $w = r$. The poset $S_D(H, G)$ is simply (H, G), and $\mu_D(H, G) = \mu(H, G)$, where μ is the Möbius function associated to the lattice of subgroups of G. So the theorem says, in this case, that $v(H, G) = |G: N(H)|\ \mu(H, G)$ is divisible by p^{r-1} in general and by p^r if G is p-perfect mod H. Applying this to all prime divisors p of $|G: H|$, we obtain the following result, which is a slight improvement of the Kratzer–Thévenaz theorem [10, Théorème 3.1] cited in the introduction:

(1.2) COROLLARY. *$v(H, G)$ is divisible by $|G: H|/|G: HG'|_0$, where G' is the commutator subgroup of G and, for any positive integer k, k_0 denotes the product of the distinct prime divisors of k.*

Remark. This result is stated here as a corollary only for the purpose of illustrating Theorem 1.1. In fact, we will give a direct proof of this case of the theorem in the next section (2.2, 2.3), and this direct proof yields a result which is stronger than (1.2). Corollary 1.2 was proved independently in [9, Theorem 4.5].

Suppose next that D consists entirely of powers of p. Then every element of D is a cone point, so $c = |D|$. And if $|G: H|$ is a power of p, then $w = |D| + 1$. So the theorem yields:

(1.3) COROLLARY. *If D consists of powers of p, then $v_D(H, G)$ is divisible by $p^{|D|}$.* If, in addition, *$|G: H|$ is a power of p and G is p-perfect* mod H, *then $v_D(H, G)$ is divisible by $p^{|D|+1}$.*

When $H = \{1\}$, this is the result stated as (A) in the introduction.

In our remaining applications of (1.1) there is no need, when counting weak cone points, to consider $|G: H|$ itself. For it can contribute to w in (1.1) only if it is a power of p, and this case is already covered by (1.3).

Suppose now that D is the half-open interval $(1, n]$, where n is a proper divisor of $|G: H|$. Then every element of D is a cone point, so $p^c = n_p$ ($=$ the p-part of n). The theorem therefore implies that $v_D(H, G)$ is divisible by n_p. Applying this to all prime divisors p of n, we obtain:

(1.4) COROLLARY. *If $D = (1, n]$ for some proper divisor n of $|G: H|$, then $v_D(H, G)$ is divisible by n.*

This result was proved independently in [9, Theorem 5.1]. When $H = \{1\}$, it reduces to (B) of the introduction. A different proof of (B), which is in fact the original proof given in [15], can be found in [16].

Next we look at the complementary situation.

(1.5.) COROLLARY. *Let n be a proper divisor of $|G: H|$ and let $D = (1, |G: H|) - (1, n]$. Let $k = \gcd(|G: H|/n, |G: HG'|)$. Then $v_D(H, G)$ is divisible by $|G: H|/nk_0$.*

We leave it to the reader to work out what this says one prime at a time and to check that it follows from the theorem. Note that (1.5) reduces to (1.2) if $n = 1$. Note also that it reduces to (C) of the introduction if $H = \{1\}$ and $k = 1$. The statement of (1.5) for $H = \{1\}$ but $k > 1$ appeared in [15] as Conjecture 4.2; it was proved independently in [9, Corollary 4.12].

We close this section by mentioning that the following famous theorem of Frobenius is a consequence of Corollary 1.4.

(1.6) COROLLARY (Frobenius [7]). *Let n be a divisor of $|G|$ and let α_n be the number of elements of G of order dividing n. Then α_n is divisible by n.*

For the proof, one shows that $\alpha_n = \sum \mu(H, K)\,|H|$, where H and K range over subgroups of order dividing n. (This is obtained by applying Möbius inversion to the equations $|H| = \sum_{K \leqslant H} f(K)$, where $f(K) = 0$ if K is not cyclic and $f(K) = \varphi(k)$ if K is cyclic of order k (φ denoting the Euler function), and then summing the resulting formulas.) The corollary follows by applying (1.4) after an easy rearrangement of the terms in the sum.

This approach to the Frobenius theorem was noticed independently by Hawkes, Isaacs, and Özaydin. The reader can refer to their paper [9, Theorem 6.3] for a more detailed proof.

2. Proof of the Theorem: A Special Case

Our main purpose in this section is to prove Theorem 1.1 when D consists of all the powers of p which are in $(1, |G: H|)$. In this case we will write $S_p(H, G)$, $\mu_p(H, G)$ and $v_p(H, G)$ for $S_D(H, G)$, $\mu_D(H, G)$, and $v_D(H, G)$. Our methods actually apply equally well to a number of other choices of D, but, for simplicity, the only other one we will explicitly mention is $D = (1, |G: H|)$. As in (1.2), we omit the subscript D in this case and write $\mu(H, G)$ and $v(H, G)$.

Recall that $\mu_p(G) = \mu_p(\{1\}, G)$ was shown in [3] to be divisible by $|G|_p$ if G is not a p-group. An equivalent formulation is that $\chi(S_p) \equiv 1 \bmod |G|_p$, where S_p is the poset $S_p(\{1\}, G)$ of non-trivial p-subgroups of G. We begin by reviewing Brown's proof, in order to see how far we can push it.

Let P be a Sylow p-subgroup of G, and consider the conjugation action of P on S_p. This induces an action of P on the simplicial complex $\Delta = \Delta(S_p)$. Let Δ' be the union of the fixed-point sets Δ^Q, where Q ranges over the non-trivial subgroups of P. Then P acts freely on the complement of Δ', so $\chi(\Delta) \equiv \chi(\Delta') \bmod |P|$. Note now that S_p^Q contains Q and has the property that $QN \in S_p^Q$ for any $N \in S_p^Q$. It follows that Δ^Q is contractible. Since the family $\{\Delta^Q\}$ is closed under intersection, the union Δ' is also contractible. In particular, $\chi(\Delta') = 1$, so $\chi(\Delta) \equiv 1 \bmod |P|$; i.e., $\chi(S_p) \equiv 1 \bmod |G|_p$. See [3] for more details.[2]

This proof can be generalized in many ways. Suppose, for example, that we are interested in subposets of an interval (H, G). Then it is natural to consider subposets S invariant under the action of $N(H)/H$ on (H, G) induced by conjugation. Given such an S, choose a Sylow p-subgroup P/H of $N(H)/H$ and consider its action on $\Delta = \Delta(S)$. Let p^k be a divisor of $|N(H): H|$ $(k \geqslant 0)$, and let Δ' be the union of the fixed-point sets $\Delta^{Q/H}$, where Q/H ranges over the subgroups of P/H with $|Q: H| > p^k$. Then for every simplex σ of Δ which is not in Δ', the stabilizer $(P/H)_\sigma$ has order dividing p^k, hence the cardinality of the orbit of σ is a multiple of $|P: H|/p^k$. Thus $\chi(\Delta) \equiv \chi(\Delta') \bmod |P: H|/p^k$. Suppose now that the following condition holds:

(C) *For every p-subgroup Q/H of $N(H)/H$ with $|Q: H| > p^k$, Q is a "cone point" for $S^{Q/H}$ in the following sense: Q is in S (and hence in $S^{Q/H}$), and for any $N \in S$ normalized by Q, $QN \in S$.*

It then follows exactly as in Brown's proof that Δ' is contractible. Consequently:

[2] Brown actually considered a more general situation, where the group G was allowed to be infinite. The proof in this generality involved some technicalities which do not arise for finite groups. The reader might therefore find it easier to refer to [12, Sect. 4], where Brown's proof as specialized to the case of finite groups is given.

(2.1) PROPOSITION. *Let S be an $N(H)$-invariant subposet of (H, G) such that* (C) *holds for some divisor p^k of $|N(H): H|$. Then $\chi(S) \equiv 1 \bmod (|N(H): H|_p)/p^k$.*

Remark. The reader who prefers a purely combinatorial proof can easily remove the topology from the discussion above. Indeed, our appeals to the contractibility of various complexes can be replaced by the following two combinatorial facts: (a) Suppose S is a poset with an element z such that z and x have a least upper bound in S for all $x \in S$; then $\chi(S) = 1$. (b) Suppose $X = \bigcup X_i$, where X is a finite simplicial complex and $\{X_i\}$ is a family of subcomplexes closed under intersection; if $\chi(X_i) = 1$ for all i, then $\chi(X) = 1$.

One can immediately deduce from Proposition 2.1 many cases of Theorem 1.1. We will confine ourselves to two:

(2.2) COROLLARY. (i) *If $|G: H|$ is not a power of p then $\mu_p(H, G)$ is divisible by $|N(H): H|_p$; hence $v_p(H, G)$ is divisible by $|G: H|_p$.*

(ii) *If G is p-perfect* mod H, *then $\mu(H, G)$ is divisible by $|N(H): H|_p$; hence $v(H, G)$ is divisible by $|G: H|_p$.*

Proof. Let S be either $S_p(H, G)$, under the assumption that $|G: H|$ is not a power of p, or (H, G), under the assumption that G is p-perfect mod H. We wish to verify condition (C) with $k = 0$. If Q/H is a non-trivial p-subgroup of $N(H)/H$, then certainly $Q \in S$ unless $Q = G$; but this is precluded by our hypotheses. Similarly, if $N \in S$ is normalized by Q, then $QN \in S$ unless $QN = G$; but QN/N is a p-group, so our hypotheses imply that $QN < G$. Thus (2.1) is applicable and yields (i) and (ii). ∎

To complete the proof of Theorem 1.1 for $v_p(H, G)$ and $v(H, G)$, it suffices to consider the latter; for if $|G: H|$ is a power of p, then the two sets D in question coincide. We wish to prove, then, that $v(H, G)$ is divisible by $(|G: H|_p)/p$ if G is not p-perfect mod H.

We begin by reviewing the Crapo complementation formula ([5, Theorem 3]; see also [2, 6.2]), which we will apply to the closed interval $[H, G]$. Recall that two elements J, K of the lattice $[H, G]$ are said to be *complements* of one another if J and K generate G and $J \cap K = H$. Fix $J \in [H, G]$ and let $J^\perp$ be the set of complements of J in $[H, G]$. Then the Crapo formula says

$$\mu(H, G) = \sum_{\substack{K, K' \in J^\perp \\ K \leqslant K'}} \mu(H, K)\, \mu(K', G).$$

This simplifies when $J = HN$ for some normal subgroup N of G. In this case any complement K has the property that the canonical map $K \to G/N$

is surjective and has kernel contained in H. Hence there are lattice isomorphisms $[H, K] \approx [J/N, G/N] \approx [J, G]$. It follows that there cannot be a proper inclusion $K < K'$ between complements of J, and the Crapo formula simplifies to

$$\mu(H, G) = \mu(J, G) \cdot \sum_{K \in J^\perp} \mu(K, G) = \mu(J/N, G/N) \cdot \sum_{K \in J^\perp} \mu(K, G). \qquad (*)$$

We can now prove that $v(H, G)$ is divisible by $(|G: H|_p)/p$. This is a consequence of the following more precise result, which was proved in [17, Corollary 4.15], for $H = \{1\}$. (A special case appears also in [9, Corollary 4.9].) Recall that $O^p(G)$ denotes the largest p-perfect subgroup of G, or, equivalently, the smallest normal subgroup of G of index a power of p.

(2.3) PROPOSITION. *Let* $J = O^p(G)\, H$. *Then* $\mu(H, G) = 0$ *unless the following three conditions are satisfied*: (a) J *is normal in* G; (b) G/J *is an elementary abelian* p*-group*; *and* (c) J *has a complement in the lattice* $[H, G]$. *Moreover, if* $|G: J| = p^n$ *and* $|J: H|_p = p^m$, *then* $v(H, G)$ *is divisible by* p^s, *where* $s = m + \binom{n}{2}$.

Proof. Suppose $\mu(H, G) \neq 0$. Then $(*)$ shows that (c) holds and that $\mu(J/O^p(G), G/O^p(G)) \neq 0$. Now if P is any p-group and Q is a subgroup, it is well known that $\mu(Q, P) \neq 0$ if and only if Q is normal in P with elementary abelian quotient, in which case $\mu(Q, P) = (-1)^r p^{\binom{r}{2}}$ where $|P: Q| = p^r$. (This follows from [14, 3.9.5, 3.10.2]; see also [10, 2.4].) This proves (a) and (b) and allows us to rewrite $(*)$ as

$$\mu(H, G) = (-1)^n p^{\binom{n}{2}} \cdot \sum_{K \in J^\perp} \mu(K, G).$$

To complete the proof, then, it suffices to show that $|G: N(H)| \cdot \sum_K \mu(K, G)$ is divisible by p^m. To this end we group the complements K into $N(H)$-conjugacy classes. Note that $N(K) \subseteq N(H)$ for any $K \in J^\perp$, since $N(K)$ normalizes K and J and hence also their intersection H. We therefore obtain

$$|G: N(H)| \cdot \sum_{K \in J^\perp} \mu(K, G) = |G: N(H)| \cdot \sum\nolimits^* |N(H): N(K)| \cdot \mu(K, G)$$

$$= \sum\nolimits^* v(K, G),$$

where the $*$'s indicate that K ranges over representatives for $J^\perp$ mod $N(H)$. Now G is p-perfect mod K for any $K \in J^\perp$; so (2.2) implies that $v(K, G)$ is divisible by $|G: K|_p = |J: H|_p = p^m$, hence the sum is also divisible by p^m. ∎

Remark. Note that, as a consequence of (2.3), we actually get divisibility of $v(H, G)$ by $|G: H|_p$, or even a higher power of p, in many cases. In fact, our divisor is $(|G: H|_p)/p$ only when $n = 1$ or 2 in (2.3).

3. Proof of the Main Theorem: The General Case

We begin by examining the effect on $v_D(H, G)$ of adjoining a new element to D. Let $D+ = D \cup \{e\}$, where e is an element of $(1, |G: H|)$ not in D. It is immediate from the definitions that

$$\mu_{D+}(H, G) = \mu_D(H, G) + \sum_{|E:H| = e} \mu_{D'}(H, E)\, \mu_{D''}(E, G), \tag{3.1}$$

where $D' = \{d \in D: d \mid e\}$ and $D'' = \{d/e: d \in D \text{ and } e \mid d\}$. Using standard poset notation, we can also write $D' = D_{<e}$ and $D'' = (1/e)\, D_{>e}$. Note that D' is a subset of $(1, e)$ and that D'' is a subset of $(1, f)$, where $f = |G: H|/e$. Applying (3.1) to all the conjugates of H and summing, we obtain

$$\begin{aligned} v_{D+}(H, G) &= v_D(H, G) + \sum_{J \sim H} \sum_{|E:J| = e} \mu_{D'}(J, E)\, \mu_{D''}(E, G) \\ &= v_D(H, G) + \sum_{|G:E| = f} \sum_{J \sim H} \mu_{D'}(J, E)\, \mu_{D''}(E, G). \end{aligned}$$

The inner sum here ranges over subgroups $J < E$ which are G-conjugate to H. Now group the E's into G-conjugacy classes and, for fixed E, group the subgroups J into E-conjugacy classes. Using $*$'s to denote summation over representatives for these conjugacy classes, the equation above becomes

$$v_{D+}(H, G) = v_D(H, G) + \sum_{|G:E| = f}^{*} \sum_{J \sim H}^{*} v_{D'}(J, E)\, v_{D''}(E, G). \tag{3.2}$$

We now prove the theorem, arguing by induction on $|G: H|$; thus we assume the theorem is known for all pairs of groups $K < L$ with $|L: K| < |G: H|$. Suppose first that D consists of powers of p, i.e., that we are in the situation of 1.3. Since the theorem is known by Section 2 when D consists of all powers of p in $(1, |G: H|)$, it suffices to show that if the theorem holds for $D+ = D \cup \{e\}$ (where e is a power of p not in D), then it holds for D. By the induction hypothesis, each term of the sum in (3.2) is divisible by $p^{|D'| + |D''|} = p^{|D|}$. Since $v_{D+}(H, G)$ is assumed to be divisible by $p^{|D|+1}$, it follows that $v_D(H, G)$ is divisible by $p^{|D|}$. And if $|G: H|$ is a power of p and G is p-perfect mod H, then $|G: E|$ is a power of p and G is p-perfect mod E; so each $v_{D''}(E, G)$ is divisible by $p^{|D''|+1}$ and we conclude that $v_D(H, G)$ is divisible by $p^{|D|+1}$. This completes the proof when D consists of powers of p.

Turning now to an arbitrary D, we can build D by starting with the powers of p in D and then successively adjoining the remaining elements, in descending order. So it suffices to show that if the theorem is true for D then it is true for $D+ = D \cup \{e\}$, where e now is *not* a power of p, and where the following condition holds: If $d \in D$ is not a power of p, then $d > e$ (for the ordinary ordering of the integers). This implies that any power of p which is a cone point (resp. weak cone point) for $D+$ is also a cone point (resp. weak cone point) for D. Thus the term $v_D(H, G)$ in (3.2) is divisible by p^c (or p^w if G is p-perfect mod H), where c (resp. w) is the number of powers of p which are cone points (resp. weak cone points) for $D+$.

We now count cone points (and weak cone points) of D' and D''. If q is a power of p which is a cone point (resp. weak cone point) for $D+$, then either $q \mid e$, in which case $q \in D'$, or else $\operatorname{lcm}(q, e)/e = q/e_p \in D''$ (resp. $D'' \cup \{f\}$). In the first case, q is a cone point for D'; in fact, our condition on e implies that D' consists entirely of powers of p, so every element is a cone point. And in the second case, it is easy to check that $\operatorname{lcm}(q, e)/e$ is a cone point (resp. weak cone point) for D''. (Just note that $\operatorname{lcm}(q, e)$ is a cone point or weak cone point for $D_{>e}$.) If we let c'' (resp. w'') be the number of cone points (resp. weak cone points) of D'', it follows that $|D'| + c'' \geqslant c$ and that $|D'| + w'' \geqslant w$. Our induction hypothesis now implies that each term of the sum in (3.2) is divisible by p^c (or p^w if G is p-perfect mod H), and the theorem is proved.

4. Further Results

Several of the steps in the proof of Theorem 1.1 involved the use of explicit formulas, such as the Crapo complementation formula in Section 2 and the formulas (3.1) and (3.2). One might therefore expect to be able to improve Theorem 1.1 in some cases by combining these formulas with information about the structure of G and/or D. We give in this section a few results of this type, which only deal with the case where D is a set of powers of a fixed prime p.

We can construct any such D by starting with all powers of p in $(1, |G: H|)$ and then removing one at a time, in descending order. If we apply (3.1) at each stage, it is easy to check by induction that we obtain the following result. Let $D^c = [1, |G: H|_p] - D$; i.e., D^c consists of the powers of p which divide $|G: H|$ and are not in D. Then

$$\mu_D(H, G) = \sum (-1)^q \mu(H_0, H_1) \cdots \mu(H_{q-1}, H_q) \mu_p(H_q, G), \qquad (4.1)$$

where the sum is taken over all chains $H = H_0 < \cdots < H_q < G$ $(q \geqslant 0)$ such that $|H_i: H| \in D^c$ for all i.

Chains $H = H_0 < \cdots < H_q < G$ as in (4.1) will be called D^c-*chains*.

Remark. Equation (4.1) is in fact a special case of a much more general formula [14, 3.14.4], which is proved in a similar way, and which describes the effect on the Möbius function when one passes from an arbitrary finite poset to an arbitrary subset.

Our main direct use of (4.1) will be for p-groups. For groups which are not p-groups, it is more convenient to rewrite (4.1) as

$$\mu_D(H, G) = -\sum_{\substack{H \leqslant P < G \\ |P:H| \in D^c}} \mu_{D(P)}(H, P)\, \mu_p(P, G), \tag{4.2}$$

where $D(P) = D_{<|P:H|}$.

(Note: For the purposes of this formula we make the convention that $\mu_\varnothing(H, H) = -1$.)

To prove (4.2), write the sum in (4.1) as

$$\sum_{|P:H| \in D^c} \sum (-1)^q\, \mu(H_0, H_1) \cdots \mu(H_{q-1}, P)\, \mu_p(P, G),$$

where the inner sum is taken over $D(P)^c$-chains. The result now follows from (4.1) applied to (H, P). Alternatively, one can get (4.2) directly by repeatedly applying (3.1), where D is constructed by starting with all powers of p in $(1, |G:H|)$ and removing one at a time, in ascending order.

In case $H = \{1\}$, we can group the terms in (4.2) according to G-conjugacy classes of the subgroups P. This yields

$$\mu_D(G) = -\sum_{|P| \in D^c}{}^{*}\, \mu_{D(P)}(P)\, \nu_p(P, G). \tag{4.3}$$

Note that each P in (4.3) is a p-group. More generally, each P in (4.2) is a p-group provided H is a p-group. Thus we need to understand the Möbius invariants $\mu_D(-, -)$ for p-groups in order to apply (4.2) and (4.3). This is a special case of what Stanley studied in [13]. But the results we need are in fact immediate consequences of (4.1), so we will give a self-contained treatment.

Suppose that G is a p-group. Recall from the proof of (2.3) that $\mu(P, Q) = 0$ for $P < Q$ in G unless P is normal in Q with elementary abelian quotient, in which case $\mu(P, Q) = (-1)^n p^{\binom{n}{2}}$, where $|Q:P| = p^n$. So we may confine our sum in (4.1) to D^c-chains which are *elementary*, in the sense that each group is normal in the next with elementary abelian quotient. If

C is an elementary D^c-chain $H = H_0 < \cdots < H_k = G$, we define an integer $b(C)$ by

$$b(C) = \sum_{i=1}^{k} \binom{n_i - n_{i-1}}{2},$$

where $|H_i| = p^{n_i}$. We can restate (4.1) as follows for p-groups:

(4.4) PROPOSITION. *If G is a p-group and $|G:H| = p^r$, then*

$$\mu_D(H, G) = \sum_C (-1)^{r-k+1} p^{b(C)},$$

where C ranges over all elementary D^c-chains $H = H_0 < \cdots < H_k = G$.

(Note: This result is somewhat different from Stanley's formula for $\mu_D(H, G)$; see Appendix A below for a discussion of Stanley's version.)

Proposition 4.4 gives, in particular, a divisibility result for $\mu_D(H, G)$ that is much sharper than (1.3) (when G is a p-group). Namely, it is easy to see that the minimum value b of $b(C)$ is attained when C is a *maximal* D^c-chain, i.e., when $k = r - |D|$ so that $\{|H_0: H|, \ldots, |H_k: H|\}$ is the entire complement of D in $[1, p^r]$. Thus $\mu_D(H, G)$ is divisible by p^b. This result, which was first proved by Stanley [13, Example 6.2, Corollary 6.5], is better than (1.3) unless H is normal in G and all of the differences $n_i - n_{i-1}$ in the definition of b are $\leqslant 2$. But even in the latter case, we can view (4.4) as an improvement of (1.3), in that it gives precise information about $\mu_D(H, G)$ in terms of the structure of G and D. For example, it implies:

(4.5) COROLLARY. *If G is a p-group and b is as above, then*

$$\mu_D(H, G) \equiv (-1)^{|D|-1} p^b e \bmod p^{b+1},$$

where $e = e_D(H, G)$ is the number of elementary maximal D^c-chains from H to G.

As a concrete illustration of this, take $H = \{1\}$ and let D be a singleton $\{p^s\}$, where $0 < s < r$ and $|G| = p^r$. Then $b = |D| = 1$. Assume, to avoid trivialities, that G is not cyclic. Then it is not hard to show that $e \equiv 1 \bmod p$ if p is odd (see Appendix B below); so (4.5), in this case, says that $\mu_D(G) \equiv p \bmod p^2$. Thus we have recovered Kulakoff's theorem:

(4.6) COROLLARY (Kulakoff [11]). *Let G be a non-cyclic group of order p^r, where p is an odd prime. If $0 < s < r$, then the number of subgroups of G of order p^s is congruent to $1 + p \bmod p^2$.*

We now turn to the case where G is *not* a p-group. For simplicity we take $H = \{1\}$, although the methods work equally well whenever H is a p-group. The results we give below arose from an attempt to answer a question of Stanley (private communication), who asked whether his divisibility results on $\mu_D(H, G)$ for p-groups G could be generalized to arbitrary finite groups.

Suppose first that G has a cyclic Sylow p-subgroup. Then for any subgroup P with $|P| \in D^c$, the poset $S_{D(P)}(P)$ is a chain with $|D(P)|$ elements. Hence $\mu_{D(P)}(P) = 0$ unless $D(P) = \varnothing$, in which case $\mu_{D(P)}(P) = -1$. Let p^s be the smallest element of D. Then the condition $D(P) = \varnothing$ simply means that $|P| < p^s$. We therefore obtain from (4.3):

$$\mu_D(G) = \sum_{|P| < p^s}^{*} v_p(P, G).$$

Now we know from (1.3) that each $v_p(P, G)$ is divisible by p^{r-s+1}, where $|G|_p = p^r$, so we obtain:

(4.7) PROPOSITION. *Suppose that G is not a p-group and that G has a cyclic Sylow p-subgroup of order p^r. If $D \subseteq (1, p^r]$ and p^s is the smallest element of D, then $\mu_D(G)$ is divisible by p^{r-s+1}.*

This is a better result than (1.3) would give, unless D consists of the entire interval $[p^s, p^r]$. In case D is simply the singleton $\{p^s\}$, for example, we recover the (elementary) result that the number of subgroups of order p^s is congruent to 1 modulo p^{r-s+1} (cf. [8, Lemma 4.61]).

If we no longer assume that G has a cyclic Sylow p-subgroup, then we cannot obtain quite as good a result, but we can still sharpen (1.3). It will be convenient, in what follows, to set $r(P) = r$ if P is a group of order p^r. (Note: This is just the rank of the lattice $[\{1\}, P]$, i.e., the length of any maximal chain, but it is not the rank of P in the sense of group theory unless P is elementary abelian.) We denote by $0 = n_0 < \cdots < n_t$ the exponents n such that $p^n \in D^c$. Then we can rewrite (4.3) as

$$\mu(G) = -\sum_{k=0}^{t} \sum_{r(P)=n_k}^{*} \mu_{D(P)}(P)\, v_p(P, G). \tag{4.8}$$

Consider a term in (4.8) with $r(P) = n_k$. Write $D = D_< \cup D_>$, where $D_< = \{p^i \in D: i < n_k\} = D(P)$ and $D_> = \{p^i \in D: i > n_k\}$. By the paragraph following (4.4), $\mu_{D(P)}(P)$ is divisible by p^{b_k}, where

$$b_k = \sum_{i=1}^{k} \binom{n_i - n_{i-1}}{2}.$$

On the other hand, $v_p(P, G)$ is divisible by $|G: P|_p = p^{|D_>| + (t-k)}$. Since

$b_k \geqslant |D_<| + |\{i: 1 \leqslant i \leqslant k,\ n_i - n_{i-1} \geqslant 3\}|$ and since obviously $t-k \geqslant |\{i: k+1 \leqslant i \leqslant t,\ n_i - n_{i-1} \geqslant 3\}|$, we obtain the following result:

(4.9) PROPOSITION. *Let D be a subset of $(1, |G|_p]$, let $0 = n_0 < \cdots < n_t$ be the exponents n such that $p^n \in D^c$, and let s be the cardinality of $\{i: 1 \leqslant i \leqslant t,\ n_i - n_{i-1} \geqslant 3\}$. Then $\mu_D(G)$ is divisible by $p^{|D|+s}$.*

We end this section by computing $\mu_D(G) \bmod p^{|D|+1}$. The analysis above shows that a term of (4.8) with $r(P) = n_k$ is divisible by $p^{|D|+(t-k)}$, hence is zero $\bmod p^{|D|+1}$ unless $k = t$. Now apply (4.4) to compute $\mu_{D(P)}(P)$ when $r(P) = n_t$. Every term of the sum in (4.4) is divisible by $p^{|D(P)|+1}$ except possibly the terms corresponding to maximal $D(P)^c$-chains. So (4.8) yields

$$\mu_D(G) \equiv (-1)^{|D(P)|}\, p^b \cdot \sum_{r(P)=n_t}^{*} e(P)\, v_p(P, G) \qquad \bmod p^{|D|+1}, \quad (4.10)$$

where $b = b_t$ and $e(P)$ is the number of elementary maximal $D(P)^c$-chains.

This takes its simplest form when $|G|_p = p^r \notin D$, so that $n_t = r$. In this case each P above is a Sylow p-subgroup, so the sum has only one term and $v_p(P, G) = |G: N(P)| \cdot (-1) \equiv -1 \bmod p$. Consequently:

(4.11) PROPOSITION. *Let $|G|_p = p^r$, let D be a subset of $(1, p^r)$, let $0 = n_0 < \cdots < n_t = r$ be the exponents n such that $p^n \in D^c$, and let $b = b_t$ be the sum of binomial coefficients defined above. Let $e = e(P)$ be the number of elementary chains $\{1\} = P_0 < \cdots < P_t = P$ with $r(P_i) = n_i$, where P is a Sylow p-subgroup of G. Then*

$$\mu_D(G) \equiv (-1)^{|D|-1}\, p^b e \bmod p^{|D|+1}.$$

(Note: We derived (4.9) and (4.11) assuming that G is not a p-group, but both results still hold when G is a p-group by (4.4).)

In case D is a singleton and p is odd, then $e \equiv 1 \bmod p$ by Appendix B if p is non-cyclic. Thus Proposition 4.11 reduces in this case to P. Hall's generalization of Kulakoff's theorem:

(4.12) COROLLARY ([8, Theorem 4.6]). *Let G have a non-cyclic Sylow p-subgroup of order p^r, where p is odd. If $0 < s < r$, then the number of subgroups of G of order p^s is congruent to $1 + p \bmod p^2$.*

APPENDIX A: STANLEY'S FORMULA

For the sake of completeness, we state here a result of Stanley [13] similar to Proposition 4.4, and we sketch briefly how it can be proved by our methods. Let G be a p-group and let $\{1\} = N_0 < \cdots < N_n = G$ be a

fixed chief series; thus each N_i is normal in G and $r(N_i) = i$. This induces a "canonical chain" $P = M_0 < \cdots < M_k = Q$ in the lattice $[P, Q]$ for any subgroups $P < Q$ of G, where $|Q: P| = p^k$; namely, form the series $\{P \cdot (Q \cap N_i)\}$ and remove all repetitions. Call a chain $P < R < Q$ *special* if R is complementary to M_i in $[P, Q]$, where $|Q: R| = p^i$. More generally, call an arbitrary chain $P_0 < \cdots < P_s$ in G *special* if the subchain $P_{j-1} < P_j < P_{j+1}$ is special for some j $(1 < j < s)$. Stanley's result, then, is

$$\mu_D(H, G) = (-1)^{|D|-1} p^b M, \tag{A1}$$

where M is the number of elementary maximal D^c-chains which are not special, and b is the number defined following (4.4) above.

(This is not stated explicitly in [13], but it is obtained by combining Theorem 1.2 and Lemma 6.4 of [13], applied to the dual of the lattice of subgroups of G. The "Loewy chains" counted by Stanley in his Lemma 6.4 correspond to our elementary maximal D^c-chains, and his condition on descent sets corresponds to our requirement that the chains not be special.)

We can also formulate (A1) as follows:

$$\mu_D(H, G) = (-1)^{m-1} \sum \mu(H_0, H_1) \cdots \mu(H_{m-1}, H_m), \tag{A2}$$

where the sum is taken over the elementary maximal D^c-chains $H = H_0 < \cdots < H_m = G$ which are not special. It is in this form that the result is easily proved by the method of Section 3.

Suppose, for instance, that D is the complement in $(1, |G: H|)$ of a single element e, and apply (3.1). This yields

$$\mu_D(H, G) = \mu(H, G) - \sum \mu(H, E)\, \mu(E, G),$$

where the sum is taken over all elementary D^c-chains $H < E < G$. On the other hand, we can use Crapo complementation to write $\mu(H, G) = \sum \mu(H, E)\, \mu(E, G)$, where the sum is now taken over all *special* elementary D^c-chains $H < E < G$; (A2) follows at once for this D. For arbitrary D, remove one element of $(1, |G: H|)$ at a time and repeatedly apply (3.1) and Crapo complementation.

Remarks. 1. The proof just sketched works for an arbitrary supersolvable lattice L and yields the results of Stanley [13] on the rank-selected Möbius invariants of such a lattice. Consider, for instance, his Theorem 1.2. One first proves, as above, an analogue (A2′) of (A2) for L (with "elementary" omitted). Since each interval $[H_{i-1}, H_i]$ is again supersolvable, this reduces Theorem 1.2 to the case of the ordinary Möbius number $\mu(0, 1)$, where 0 (resp. 1) is the smallest (resp. largest) element of L.

But the result in this case (i.e., Stanley's Corollary 1.4) is easily proved by Crapo complementation and induction on the rank of L.

Similarly, if L is a q-SS lattice as in Stanley's Lemma 6.4, then (A2′) immediately gives a formula analogous to (A1) for the rank-selected Möbius invariants of L.

2. Our methods also give an easy proof of Björner's theorem [1] that all the rank-selected posets in question, such as $S_D(H, G)$ for G a p-group, are spherical. One starts by proving this for $L - \{0, 1\}$ by "homotopy complementation" and induction on the rank of L (cf. [2, Theorem 5.1]). The rank-selected subposets are then treated as in Section 3 by removing one element at a time from the set of allowable ranks.

Appendix B: A Lemma about p-Groups

Let G be a p-group of order p^r, let D be the singleton $\{p^s\}$ for some s with $0 < s < r$, and let $e = e_D(G)$ be the number of elementary maximal D^c-chains from $\{1\}$ to G.

Lemma. *If p is odd and G is not cyclic, then $e \equiv 1 \bmod p$.*

(This was needed above in order to deduce (4.6) from (4.5), and again to get (4.12) from (4.11).)

Proof. We argue by induction on $|G|$. If $s = r - 1$, then $e = \sum_Q f(Q)$, where Q ranges over the normal subgroups of G such that G/Q is elementary abelian of rank 2, and $f(Q)$ is the number of maximal chains of subgroups of Q. Now $f(Q)$ is easily shown by induction to be congruent to 1 mod p, so e is congruent mod p to the number of Q's, i.e., to the number of subgroups of G/J of index p^2, where J is the Frattini subgroup of G. Since G is not cyclic, G/J has order at least p^2, so the number of subgroups of index p^2 is indeed congruent to 1 mod p.

If $s < r - 1$, on the other hand, then $e = \sum_Q e(Q)$, where Q now ranges over the subgroups of G of index p and $e(Q) = e_D(Q)$. If all such Q's are non-cyclic, then we are done by the induction hypothesis and the fact that the number of Q's is congruent to 1 mod p. If at least one Q is cyclic, on the other hand, then there are $p + 1$ subgroups Q of index p, and exactly p of these are cyclic; see, for instance, [4, IV.4.1, IV.4.4]. (It is here that one needs the assumption that p is odd; one also has to note that $r \geqslant 3$.) So $e(G) = e(Q)$ for the non-cyclic Q, and we are again done by the induction hypothesis. ∎

REFERENCES

1. A. BJÖRNER, Shellable and Cohen–Macaulay partially ordered sets, *Trans. Amer. Math. Soc.* **260** (1980), 159–183.
2. A. BJÖRNER AND J. W. WALKER, A homotopy complementation formula for partially ordered sets, *European J. Combin.* **4** (1983), 11–19.
3. K. S. BROWN, Euler characteristics of groups: The p-fractional part, *Invent. Math.* **29** (1975), 1–5.
4. K. S. BROWN, "Cohomology of groups," Springer-Verlag, New York/Heidelberg/Berlin, 1982.
5. H. H. CRAPO, The Möbius function of a lattice, *J. Combin. Theory* **1** (1966), 126–131.
6. G. FROBENIUS, Verallgemeinerung des Sylow'schen Satzes, *Berliner Sitzungsberichte* (1895), 981–993.
7. G. FROBENIUS, Über einen Fundamentalsatz der Gruppentheorie, *Berliner Sitzungsberichte* (1903), 987–991.
8. P. HALL, On a theorem of Frobenius, *Proc. London Math. Soc. (2)* **40** (1936), 468–501.
9. T. HAWKES, I. M. ISAACS, AND M. ÖZAYDIN, On the Möbius function of a finite group, preprint, 1986.
10. C. KRATZER AND J. THÉVENAZ, Fonction de Möbius d'un groupe fini et anneau de Burnside, *Comment. Math. Helv.* **59** (1984), 425–438.
11. A. KULAKOFF, Über die Anzahl der eigentlichen Untergruppen und der Elemente von gegebener Ordnung in p-Gruppen, *Math. Ann.* **104** (1931), 778–793.
12. D. QUILLEN, Homotopy properties of the poset of non-trivial p-subgroups of a group, *Adv. in Math.* **28** (1978), 101–128.
13. R. P. STANLEY, Supersolvable lattices, *Algebra Universalis* **2** (1972), 197–217.
14. R. P. STANLEY, "Enumerative Combinatorics," Vol. I, Wadsworth, Monterey, 1986.
15. J. THÉVENAZ, Generalizations of Sylow and Brown theorems, unpublished manuscript, 1986.
16. J. THÉVENAZ, Idempotents de l'anneau de Burnside et caractéristique d'Euler, Séminaire groupes finis III, *Publ. Math. Univ. Paris VII*, in press.
17. J. THÉVENAZ, Permutation representations arising from simplicial complexes, *J. Combin. Theory Ser. A* **46** (1987), 121–155.