ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 30, Number 1, Spring 2000

ALPERIN'S CONJECTURE AND THE SUBGROUP STRUCTURE OF A FINITE GROUP

I.M. ISAACS

1. Introduction. In this mostly expository paper, we show how information about the subgroup structure of a finite group G can be used to determine certain numerical invariants of the group. Among the invariants we will consider are the number $k(G)$ of conjugacy classes of G and the closely related quantity $k_\pi(G)$, which is the number of classes of π-elements of G, where π is a set of prime numbers. For each prime p we will also consider the more subtle invariant $z_p(G)$, defined to be the number of ordinary irreducible characters of G that have p-defect zero. (Recall that these are the characters $\chi \in \mathrm{Irr}\,(G)$ such that the integer $|G|/\chi(1)$ is not divisible by p.) If p does not divide $|G|$, we see that every irreducible character has p-defect zero, and thus $z_p(G) = k(G)$ in this case. In general, however, it is much more difficult to determine $z_p(G)$ from the structure of G. The problem of doing this was proposed by Brauer [**1**] and solved by Robinson [**6**]. Unfortunately, Robinson's solution is fairly technical, and it does not seem to yield a direct algorithm for computing $z_p(G)$ from the subgroup structure of G.

The validity of our computation of class numbers is very easily established, but the application of our method to the determination of $z_p(G)$ relies on the unproved Alperin weight conjecture (AWC), which we shall explain in Section 3. (But the AWC is known to hold for p-solvable groups [**4**], and so our algorithm for determining $z_p(G)$ is definitely valid in that case.)

In fact, the AWC is one of the most significant outstanding problems in finite group representation theory, and it has been a focus of intense interest and activity by many researchers. It seems appropriate, therefore, that an account of this conjecture and of at least one of its potential applications should be made accessible to as wide an audience as possible. That is one of the goals of this paper, which we have

Received by the editors on April 24, 1998.
Research partially supported by a grant from the National Security Agency.

attempted to make self-contained. Although a number of variations of the AWC have been formulated, we will limit our attention to the most basic version, and we use it to give an elementary argument that proves the validity of our algorithm for computing $z_p(G)$ from the subgroup structure of G. We stress, however, that there is little in this paper that is really new, except possibly for the point of view. (In particular, others have observed that the AWC yields methods for determining the number of p-defect zero characters of a group from its subgroup structure.)

Let us be precise about what we mean by the "subgroup structure" of a finite group G. If $H \subseteq G$ is a subgroup, we write $[H]$ to denote the orbit of H in the conjugation action of G on its subgroups, and we let $\mathcal{H}$ denote the collection of all such classes of subgroups of G. We define the nonnegative-integer valued function B, of two variables on $\mathcal{H}$, by setting $B([K],[H])$ to be the number of G-conjugates of K contained in each conjugate of H. This function can, of course, be presented as a square matrix with rows and columns indexed by $\mathcal{H}$, and in which the $([K],[H])$-entry is $B([K],[H])$. (This matrix is closely related to the so-called "Burnside matrix" of G.) In addition to the containment information encoded in B, we shall also need to know the orders of the various subgroups of G. For the purposes of this paper, therefore, we define the *subgroup-structure* of G to be the matrix corresponding to B together with the column vector, indexed by $\mathcal{H}$, in which the $[H]$-entry is the order $|H|$ of the subgroup H. We stress that, by our definition, the subgroup structure of G is a certain array of integers; it does not explicitly include information about the isomorphism types of the various subgroups.

The following easy result is essentially due to Hirsch [**3**]. We shall see that it enables us to compute the class number $k(G)$ from a knowledge of the subgroup structure of G together with exactly one bit of additional information for each member of $\mathcal{H}$: if $[H] \in \mathcal{H}$, we need to know whether or not H is abelian.

Theorem A (Hirsch). *Given a finite group, G, there exists a unique integer-valued function f defined on $\mathcal{H}$ such that $f([K]) = 0$ if K is* *nonabelian, and such that for every subgroup $H \subseteq G$, we have*

$$|H|k(H) = \sum_{[K]\in\mathcal{H}} f([K])B([K],[H]).$$

In fact, $f([K]) = 0$ unless K can be generated by two elements.

Theorem A yields a recursive algorithm for computing $k(H)$, and also $f([H])$, for every subgroup $H \subseteq G$. We can suppose that we already know both $k(K)$ and $f([K])$ for all subgroups $K \subseteq G$ such that $|K| < |H|$. In the equation of Theorem A, we need to consider only those subgroup classes $[K] \in \mathcal{H}$ for which $B([K],[H]) \neq 0$. But $B([K],[H]) \neq 0$ only when $[K] = [H]$ or $|K| < |H|$. By the inductive hypothesis, therefore, we know all of the relevant quantities that appear in the equation except $k(H)$ and $f([H])$. But we also know one of these, since $k(H) = |H|$ if H is abelian and, by Theorem A, $f([H]) = 0$ if H is nonabelian. (Recall that we are assuming that we know whether or not H is abelian.) Since $|H| \neq 0$ and $B([H],[H]) = 1 \neq 0$, we can solve for the one remaining unknown quantity. This argument clearly also establishes the uniqueness of the function f, as asserted in the statement of the theorem.

To see how this works in a specific example, consider the case where G is nonabelian of order 6, so that G has exactly four classes of subgroups, of orders 1, 2, 3 and 6, and the subgroup structure of G is the following matrix and column:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 6 \end{bmatrix}.$$

Since in this example the four subgroup orders happen to be distinct, we can use these orders to name the subgroup classes. For example, the entry 3 in the matrix tells us that $B(2,6) = 3$, and this corresponds to the fact that there are three conjugates of the subgroup of order 2 in the subgroup of order 6.

When $[H] = 1$, we know that H is abelian, and thus $k(H) = |H| = 1$. Theorem A yields $1 = |H|k(H) = f(1)$, and so $f(1) = 1$. Also when $[H] = 2$, we know that H is abelian. Thus $k(H) = |H| = 2$, and we have $4 = |H|k(H) = f(1) + f(2)$, and we deduce that $f(2) = 3$. Similarly, when $[H] = 3$, we get $9 = |H|k(H) = f(1) + f(3)$, and thus $f(3) = 8$. Finally, when $[H] = 6$, Theorem A yields $6k(6) = f(1) + 3f(2) + f(3) + f(6)$. But we know that H is not abelian in this case, and so $f(6) = 0$. We deduce that $6k(6) = 1 + 9 + 8 = 18$, and thus $k(G) = 3$. This, of course, is the correct answer.

A very similar algorithm can be used to compute the number of π-classes of an arbitrary finite group G, where π is a set of prime numbers. In fact, Theorem A goes through essentially unchanged.

Theorem B. *Given a finite group G and a prime set π, there exists a unique integer-valued function f defined on $\mathcal{H}$ such that $f([K]) = 0$ if K is nonabelian, and such that for every subgroup $H \subseteq G$, we have*

$$|H|k_\pi(H) = \sum_{[K]\in\mathcal{H}} f([K])B([K],[H]).$$

In fact, $f([K]) = 0$ unless K can be generated by two elements.

Observe that, if H is abelian, then $k_\pi(H) = |H|_\pi$, and so in this situation, too, we see that for every member $[H] \in \mathcal{H}$, one of the quantities $k_\pi(H)$ or $f([H])$ is known, depending on whether H is abelian or nonabelian. Exactly as in the earlier situation, if we know which members of $\mathcal{H}$ correspond to abelian subgroups, then the numbers $k_\pi(H)$ and $f([H])$ can be computed recursively. In particular, this establishes the uniqueness of f.

Continuing with the example where G is nonabelian of order 6, and taking $\pi = \{3\}$, we get $1 = f(1)$, $2 = f(1) + f(2)$ and $9 = f(1) + f(3)$, so that $f(2) = 1$ and $f(3) = 8$. Since $f(6) = 0$, we obtain $6k_\pi(G) = 1 + 3 + 8 = 12$, and thus $k_\pi(G) = 2$, which is correct.

As we shall see in Section 2, the proofs of Theorems A and B are nearly trivial. This is not the case for the following result, however, which, together with the AWC, underlies our algorithm for computing $z_p(G)$. As with our algorithms for $k(G)$ and $k_\pi(G)$, we assume that we know the subgroup structure of G and a bit more. The additional information we need to compute $z_p(G)$ is the knowledge for each member $[H] \in \mathcal{H}$ of whether or not the derived subgroup H' is a p-group. (Recall that for the class-number computations, the required additional information was very similar: we needed to know whether or not H' was trivial.) This result is essentially due to Thévenaz [**7**].

Theorem C (Thévenaz). *Assume that all groups involved in G satisfy the Alperin weight conjecture for the prime p. Then there exists a unique integer-valued function f on $\mathcal{H}$ such that $f([K]) = 0$ unless the commutator subgroup K' is a p-group, and such that for every subgroup $H \subseteq G$ we have*

$$|H| z_p(H) = \sum_{[K] \in \mathcal{H}} f([K]) B([K], [H]).$$

In fact, $f([K]) = 0$ unless K has an elementary abelian normal Sylow p-subgroup E and K/E can be generated by two commuting elements.

We mention that if a group G has a nontrivial normal p-subgroup P, then $z_p(G) = 0$. Perhaps the easiest way to see this is to recall that if $\chi \in \mathrm{Irr}\,(G)$ has p-defect zero, then χ vanishes on all elements of order divisible by p in G. It follows that the character inner product $[\chi_P, 1_P] = \chi(1)/|P| \neq 0$, and hence χ_P must have a principal constituent. As $P \triangleleft G$, we deduce that $P \subseteq \ker(\chi)$. But this is impossible if $P > 1$ since χ vanishes on the nonidentity elements of P.

If $K \subseteq G$ and K' is not a p-group, we know in the situation of Theorem C that we have $f([K]) = 0$. But when K' is a p-group, we can easily determine $z_p(K)$. In this case, K has a normal Sylow p-subgroup P, and so $z_p(K) = 0$ if P is nontrivial. If $P = 1$, on the other hand, then p does not divide $|K|$, and in this case K is abelian and $z_p(K) = k(K) = |K|$. It follows from this, and assuming the AWC, that if we know the subgroup structure of G and we also know for every member $[K] \in \mathcal{H}$ whether or not K' is a p-group, we can determine the numbers $z(H)$ and $f([H])$ for every subgroup $H \subseteq G$. In particular, we can compute $z_p(G)$, and we know that f is unique, as asserted.

To see how this works in practice, let us return to the nonabelian group of order 6, taking $p = 2$. If $[H] = 1$, then $|H| = 1$ is not divisible by p and H' is a p-group. We thus have $1 = |H| z_p(H) = f(1)$. If $[H] = 2$, then $|H| = 2$ is divisible by p and H' is a p-group, and thus $0 = |H| z_p(H) = f(1) + f(2)$, and thus $f(2) = -1$. When $[H] = 3$, we see that $|H|$ is not divisible by p and H' is a p-group, and thus $9 = |H|z_p(H) = f(1) + f(3)$, and so $f(3) = 8$. For $[H] = 6$, we know that H' is not a p-group, and thus $f(6) = 0$. We thus have $6z_p(H) = f(1) + 3f(2) + f(3) = 1 - 3 + 8 = 6$, and we deduce that $z_2(G) = 1$, which is correct.

Before proceeding with the proofs of our theorems, we pause to acknowledge with thanks the help of Geoff Robinson and Jacques Thévenaz who read a preliminary version of this paper and made a number of useful comments.

2. Counting classes.

Since Theorem A is a special case of Theorem B, it suffices to prove the latter result.

Proof of Theorem B. Let X denote the set of π-elements of G, and note that if $H \subseteq G$, then $k_\pi(H)$ is the number of orbits in the action of H on $X \cap H$. By the standard orbit-counting formula, often wrongly attributed to Burnside, this number is equal to $(1/|H|)\sum_{h\in H} c(h)$, where we have written $c(h)$ to denote the number of elements of $X \cap H$ that are centralized by $h \in H$. It follows that $|H|k_\pi(H)$ is equal to the cardinality of the set $\mathcal{S}(H) = \{(h,x) \mid h \in H, x \in X \cap H \text{ and } xh = hx\}$.

For subgroups $K \subseteq G$, write $f(K)$ to denote the number of pairs $(k,x) \in \mathcal{S}(K)$ such that $K = \langle k, x\rangle$, and note that if $f(K) \neq 0$, then K is generated by a pair of commuting elements, and hence it is abelian. If $H \subseteq G$ and (h,x) is any member of $\mathcal{S}(H)$, then the subgroup $K = \langle h, x\rangle$ is clearly contained in H. Also, if K is any subgroup of H, then there are exactly $f(K)$ pairs (h,x) in $\mathcal{S}(H)$ such that $\langle h, x\rangle = K$, and as K runs over all subgroups of H, we see that these account for all of $\mathcal{S}(H)$. It follows that

$$|H|k_\pi(H) = |\mathcal{S}(H)| = \sum_{K\subseteq H} f(K).$$

Observe that the number $f(K)$ is determined by the isomorphism type of K and the prime set π and, in particular, f is constant on the subgroup class $[K]$ in G. We can thus view f as being defined on the set $\mathcal{H}$ of subgroup classes, where we write $f([K]) = f(K)$. With this notation, we see that

$$\sum_{K \subseteq H} f(K) = \sum_{[K] \in \mathcal{H}} f([K])B([K],[H]),$$

and the result follows. □

3. Alperin's weight conjecture. Let G be a finite group and fix a prime p. Following Alperin, we define a *weight* of G to be an ordered pair (P, α) where P is a p-subgroup of G and α is an ordinary irreducible character of $\mathbf{N}_G(P)/P$ that has p-defect zero. (Of course, we can also view α as an irreducible character of $\mathbf{N}_G(P)$ with P contained in its kernel, and such that the p-part of $\alpha(1)$ is equal to the p-part of the index $|\mathbf{N}_G(P) : P|$.) The number of weights of G with first component P is just the number of p-defect zero irreducible characters of $\mathbf{N}_G(P)/P$, and we can write this as $z(\mathbf{N}_G(P)/P)$. (Since the prime p is being held fixed in what follows, we will write $z(\ )$ in place of $z_p(\ )$ for the function that counts p-defect zero irreducible characters of a group.)

The group G acts by conjugation on the set of weights of G in a natural way. To see this, consider a weight (P, α), and view $\alpha \in \mathrm{Irr}\,(N)$, where $N = \mathbf{N}_G(P)$. If $g \in G$, then P^g is a p-subgroup of G having normalizer N^g, and we can define $\alpha^g \in \mathrm{Irr}\,(N^g)$ by setting $\alpha^g(x^g) = \alpha(x)$ for all $x \in N$. It is then clear that $P^g \subseteq \ker(\alpha^g)$ and that the pair (P^g, α^g) is a weight. It is easy to check that the formula $(P, \alpha)^g = (P^g, \alpha^g)$ defines an action of G on the set of weights of G, and we observe that the stabilizer of (P, α) in G is exactly $N = \mathbf{N}_G(P)$. (This is because $\alpha^n = \alpha$ if $n \in N$.)

In the simplest and most accessible of its several variations, the Alperin weight conjecture (AWC) asserts that the number of G-orbits of weights of G is equal to the number $l(G)$ of conjugacy classes of p-regular elements of G. (Note that $l(G)$ is just $k_{p'}(G)$, in our earlier notation.) Since the size of the orbit containing the weight (P, α) of G is exactly $|G : \mathbf{N}_G(P)|$, we see that we can count the number of orbits of weights of G by summing the quantity $1/|G : \mathbf{N}_G(P)|$ over all weights (P, α). Since the number of weights having first component P is $z(\mathbf{N}_G(P)/P)$, it follows that the number of orbits of weights of G is exactly $\sum_P (z(\mathbf{N}_G(P)/P)/|G : \mathbf{N}_G(P)|)$, where the sum is taken over all p-subgroups $P \subseteq G$. We can thus restate the AWC for the group G in the form

$$(*) \qquad |G|l(G) = \sum_P z(\mathbf{N}_G(P)/P)|\mathbf{N}_G(P)|.$$

The term in the sum on the right side of equation $(*)$ corresponding to $P = 1$ is $|G|z(G)$, and thus the AWC implies that $l(G) \geq z(G)$. (This inequality is easy to prove directly using standard elementary facts from representation theory.) From the AWC, we see that the difference $l(G) - z(G)$ is equal to a quantity that can be computed in terms of the p-local subgroups of G, which are the subgroups $\mathbf{N}_G(P)$, where P is a nontrivial p-subgroup of G. The fact that the AWC implies that the quantity $l(G) - z(G)$ is "determined p-locally" is explored more fully and more precisely in [**7**].

Suppose that p does not divide $|G|$. The only p-subgroup of G is 1, and so we see that assertion of the AWC in this case is simply that $z(G) = l(G)$. But every irreducible character of G has p-defect zero and every element is p-regular, and thus $z(G) = |\mathrm{Irr}\,(G)|$ and $l(G) = k(G)$. Since it is always true that $|\mathrm{Irr}\,(G)| = k(G)$, we see that the AWC holds, for trivial reasons, in the case where p does not divide $|G|$. At the opposite extreme, suppose now that G is a p-group, so that $l(G) = 1$, and the left side of equation $(*)$ is equal to $|G|$. If $P < G$, then $\mathbf{N}_G(P)/P$ is a nontrivial p-group, and so it has no irreducible characters with p-defect zero. The only nonzero term on the right of $(*)$, therefore, is $z(G/G)|G| = |G|$, which occurs for $P = G$. Thus equation $(*)$ is valid and the AWC holds in this case too. More generally, it is known that the AWC holds for all p-solvable groups, and it has been verified for many other groups, including all symmetric groups, for all primes.

4. Chains of p-subgroups. The AWC is intimately associated with the theory of chains of p-subgroups of a finite group. (By a *chain* of subgroups, we simply mean a collection that is linearly ordered by inclusion.) The connection between the AWC and p-group chains was first observed by Knörr and Robinson, who presented in [**5**] a reformulation of the AWC in terms of such chains. In fact, it is not very difficult to construct such a reformulation, and it appears to be quite useful to do so. (Our proof of Theorem C, for example, relies on a chain version of the AWC.) In this section, we introduce some of the relevant ideas and we prove a form of the Knörr-Robinson result.

Fix a prime number p, and let G be a finite group. We write $\mathcal{N} = \mathcal{N}(G)$ to denote the collection of all chains σ of *nontrivial* p-subgroups of G such that each member of σ is normal in the largest. (We shall refer to such chains as *normal* chains, and we observe that the empty chain is normal by default.) Note that G permutes $\mathcal{N}$ by conjugation, and if $\sigma \in \mathcal{N}$, we write G_σ to denote the stabilizer in G of the chain σ. Observe that if σ is nonempty, then G_σ is just the intersection of the normalizers in G of all the members of the chain σ, and if σ is empty, then, of course, $G_\sigma = G$.

Suppose P is a nontrivial p-subgroup of G, and let $N = \mathbf{N}_G(P)$. There is a natural injective map $\tau \mapsto \tau^*$ from $\mathcal{N}(N/P)$ into $\mathcal{N}(G)$, defined as follows. If $\tau \in \mathcal{N}(N/P)$, we take τ^* to be the chain of p-subgroups of G consisting of P and all of the preimages of the members of τ. In particular, all of the members of τ^* are contained in N, and so each contains P as a normal subgroup, and, in fact, it is easy to see that τ^* is a normal chain. Observe that $|\tau^*| = 1 + |\tau|$, and, in particular, τ^* is nonempty. Conversely, if σ is any normal chain in G having P as its smallest member, then clearly $\sigma = \tau^*$ for some unique (possibly empty) chain $\tau \in \mathcal{N}(N/P)$.

Before we state our result relating $\mathcal{N}(G)$ to the AWC, we recall that a *section* of a group G is a group of the form H/K, where $K \lhd H \subseteq G$, and this section is *proper* if $|H/K| < |G|$.

Theorem 4.1 (Knörr-Robinson). *Assume that the* AWC *holds for all proper sections of G. Then the* AWC *holds for G if and only if*

$$(**) \qquad |G|z(G) = \sum_{\sigma \in \mathcal{N}(G)} (-1)^{|\sigma|} |G_\sigma| l(G_\sigma).$$

For future reference, we will write $R(G)$ to denote the quantity on the right side of equation $(**)$.

Proof of Theorem 4.1. In the sum $R(G)$, the term corresponding to the empty chain is $|G|l(G)$, and we collect the remaining terms according to the minimum member of the corresponding chain. For each nontrivial p-subgroup P of G, we write $r(P)$ to denote the sum of the terms in $R(G)$ corresponding to those normal chains σ having minimal member P. We can thus write $R(G) = |G|l(G) + \sum_P r(P)$, where P runs over the nontrivial p-subgroups of G.

To evaluate $r(P)$, we use the fact that the chains $\sigma \in \mathcal{N}(G)$ with minimum member P are exactly the chains τ^* for $\tau \in \mathcal{N}(N/P)$, where $N = \mathbf{N}_G(P)$. Since $P \in \tau^*$, we see that $G_{\tau^*} \subseteq N$, and, in fact, $G_{\tau^*}/P = (N/P)_\tau$. Next, we observe that the canonical homomorphism from G_{τ^*} onto $(N/P)_\tau$ maps the p-regular classes of G_{τ^*} to p-regular classes of $(N/P)_\tau$. This map on classes is easily seen to be surjective, and because P is a p-group, it is also injective. (The injectivity follows from the fact that all complements to a normal Sylow subgroup in a finite group are necessarily conjugate.) It follows that $l(G_{\tau^*}) = l((N/P)_\tau)$. We thus have

$$\begin{aligned} r(P) &= \sum_{\tau \in \mathcal{N}(N/P)} (-1)^{|\tau^*|} |G_{\tau^*}| l(G_{\tau^*}) \\ &= \sum_{\tau \in \mathcal{N}(N/P)} (-1)^{|\tau|+1} |P| |(N/P)_\tau| l((N/P)_\tau) \\ &= -|P| R(N/P), \end{aligned}$$

where $R(N/P)$ represents the right side of equation $(**)$ as applied to the group N/P.

Since we are considering nontrivial p-subgroups P, we see that N/P and all of its proper sections are proper sections of G, and hence these groups all satisfy the AWC by hypothesis. Working by induction on $|G|$, we can thus assume that equation $(**)$ holds for N/P, and thus $|N/P|z(N/P) = R(N/P)$. Substitution of this into the previous equation yields $r(P) = -|N|z(N/P)$, and it follows that

$$|G|l(G) - R(G) = -\sum_P r(P) = \sum_P |\mathbf{N}_G(P)| z(\mathbf{N}_G(P)/P),$$

where the sum is taken over all nontrivial p-subgroups P of G. Adding $|G|z(G)$ to both sides, we obtain

$$|G|l(G) - R(G) + |G|z(G) = \sum_P |\mathbf{N}_G(P)| z(\mathbf{N}_G(P)/P),$$

where now the sum is taken over all p-subgroups of G, including the identity. The sum on the right side of this equation is exactly the right side of equation $(*)$, while $|G|l(G)$ is the left side of $(*)$. We conclude that equation $(*)$ holds if and only if $|G|z(G) = R(G)$, and we recall that this is precisely equation $(**)$. Since we know that $(*)$ is equivalent to the AWC, the proof is complete. □

A surprising variation on Theorem 4.1, which was first observed by Knörr and Robinson in [**5**], is that if the quantity $l(G_\sigma)$ in the sum on the right side of equation $(**)$ is replaced by $k(G_\sigma)$, then the value of the sum remains unchanged. We get the same final total, in other words, if we count all the classes in the chain stabilizers instead of only the p-regular classes. (This result can be derived as a consequence of our argument establishing Theorem C, and we present a proof in Section 7.)

5. Simplification of the right hand side. Our goal in this section is to obtain a simpler formula for the quantity $R(G) = \sum_\sigma (-1)^{|\sigma|}|G_\sigma|l(G_\sigma)$ occurring on the right side of equation $(**)$ in Theorem 4.1. (Recall that, in this sum, σ runs over $\mathcal{N} = \mathcal{N}(G)$, the set of all normal chains of nontrivial p-subgroups of G.)

Let X be the set of all p-regular elements of G. We observe that $l(G_\sigma)$ is just the number of orbits in the conjugation action of G_σ on the set $X \cap G_\sigma$, and so we can use the "Burnside" orbit counting formula to determine that $|G_\sigma|l(G_\sigma) = \sum_{g \in G_\sigma} c(g)$, where we have written $c(g)$ to denote the number of elements of $X \cap G_\sigma$ that are centralized by g. In other words, $|G_\sigma|l(G_\sigma)$ is precisely the number of ordered pairs (g, x), where $g \in G_\sigma$, $x \in X \cap G_\sigma$ and $gx = xg$. If we write $\mathcal{T}$ to denote the set of all ordered triples (g, x, σ) such that $g \in G$, $x \in X$, $gx = xg$ and x and g both stabilize $\sigma \in \mathcal{N}$, it follows that

$$R(G) = \sum_{(g,x,\sigma) \in \mathcal{T}} (-1)^{|\sigma|}.$$

If σ is any nonempty chain of nonidentity p-subgroups of G, then, of course, σ has a unique maximum member, and we write $\max(\sigma)$ to denote this nontrivial p-group. It is convenient to extend the definition to the empty chain, and to define $\max(\sigma) = 1$, the trivial subgroup, if

σ is empty. For each triple $(g, x, \sigma) \in \mathcal{T}$, we write $\langle g, x, \sigma\rangle$ to denote the subgroup $\langle g, x, \max(\sigma)\rangle$.

If $K \subseteq G$ is any subgroup, we write $f(K)$ to denote the (often zero) integer given by $\sum(-1)^{|\sigma|}$, where the sum is taken over all triples $(g, x, \sigma) \in \mathcal{T}$ such that $\langle g, x, \sigma\rangle = K$. We thus have $R(G) = \sum f(K)$, where the sum runs over all subgroups $K \subseteq G$. Since the function f is clearly constant on the G-orbit $[K] \in \mathcal{H}$ of a subgroup K, we can write $f([K[)$ in place of $f(K)$, and we see that

$$R(G) = \sum_{[K]\in\mathcal{H}} f([K])B([K],[G]),$$

and more generally, if $H \subseteq G$, we have

$$R(H) = \sum_{[K]\in\mathcal{H}} f([K])B([K],[H]),$$

where $R(H)$ is, of course the right side of the equation of Theorem 4.1 as applied to the subgroup H. If all sections of G satisfy the AWC, then for each subgroup $H \subseteq G$, Theorem 4.1 guarantees that $|H|z(H) = R(H)$, and thus

$$|H|z(H) = \sum_{[K]\in\mathcal{H}} f([K])B([K],[H]).$$

To complete the proof of Theorem C, therefore, it suffices to prove the following.

Theorem 5.1. *If $K \subseteq G$, then $f(K) = 0$ unless K has an elementary abelian normal Sylow p-subgroup E and K/E can be generated by two commuting elements.*

We shall give the complete proof of this result in the next section, but for now, we observe that, if $f(K) \neq 0$, then there must be at least one triple $(g, x, \sigma) \in \mathcal{T}$ such that $K = \langle g, x, \sigma\rangle$, and thus K is generated by the two commuting elements g and x and some p-subgroup $P \triangleleft K$, where $P = \max(\sigma)$. It follows that $f(K) = 0$ unless the derived subgroup K' is a p-group, in which case we see that K has a normal Sylow p-subgroup E. In other words, all of Theorem 5.1 and Theorem C have been proved except for the assertion that E is elementary abelian when

$f(K) \neq 0$. (Note that the missing part of Theorem C is not relevant to the algorithm for counting characters of defect zero that we discussed in Section 1.)

6. Möbius functions. To establish that the normal Sylow p-subgroup E of K must be elementary abelian if $f(K) \neq 0$, we shall need some of the elementary theory of Möbius functions, and we proceed to review this now. (See [**2**], for example, for more detail on this material.)

Recall that if $\mathcal{P}$ is a finite poset, then the Möbius function $\mu = \mu_{\mathcal{P}}$ is the unique function of two variables on $\mathcal{P}$ satisfying the conditions that $\mu(x,x) = 1$ for all $x \in \mathcal{P}$, that $\mu(x,y) = 0$ if $x \not\leq y$ and that if $x < y$, then $\sum_z \mu(x,z) = 0$, where z runs over all elements in $\mathcal{P}$ such that $x \leq z \leq y$.

The connection between Möbius functions and chains is given by the following:

Lemma 6.1. *Let $x < y$ in a finite poset $\mathcal{P}$. Then*

$$\mu(x,y) = \sum_{\sigma} (-1)^{|\sigma|},$$

where σ runs over chains of elements $z > x$ in $\mathcal{P}$ such that the largest member of σ is y.

Proof. Use induction on the length of the longest chain above x with maximum element y. See Lemma 2.2 of [**2**] for the details. □

We shall be especially concerned with posets formed by collections of subgroups of a group, where the partial order is given by containment.

Lemma 6.2. *Let $\mathcal{P}$ be an intersection-closed collection of subgroups of a finite group $\mathcal{P}$. Assume that $1 \in \mathcal{P}$, but that the intersection of all of the maximal members of $\mathcal{P}$ is nontrivial. Then $\sum_{Q \in \mathcal{P}} \mu(1,Q) = 0$.*

Proof. For each maximal member M of $\mathcal{P}$, write $\mathcal{P}(M) = \{Q \in \mathcal{P} \mid Q \subseteq M\}$. Then $\mathcal{P} = \cup_M \mathcal{P}(M)$, and the sum S of the lemma can be computed using the inclusion-exclusion principle. Specifically, we have

$S = \sum \pm S_D$, where D runs over all possible intersections of maximal members of $\mathcal{P}$ and $S_D = \sum \mu(1, Q)$, where Q runs over all members of $\mathcal{P}$ contained in D. By the defining property of the Möbius function, each of the sums $S_D = 0$ because, for each of the intersections D, we have $D \in \mathcal{P}$ and $D > 1$ by hypothesis. See Lemma 2.4 of [**2**] for further detail. □

Now fix a subgroup $A \subseteq G$ where G is any finite group. For each A-invariant p-subgroup P of G, we define integers $a(P)$ and $n(P)$ as follows. We set $a(P) = \sum (-1)^{|\sigma|}$, where σ runs over all chains of A-invariant nonidentity subgroups of P such that $P = \max(\sigma)$. The definition of $n(P)$ is similar, except that we sum only over *normal* chains σ. (Recall that the unique chain σ for which $\max(\sigma) = 1$ is the empty chain, and thus $a(1) = 1 = n(1)$.)

Corollary 6.3. *For every A-invariant p-subgroup $P \subseteq G$, we have $a(P) = n(P)$. Also, this quantity vanishes unless P is elementary abelian.*

Proof. If P is abelian, we clearly have $a(P) = n(P)$, and so it suffices to show that each of these quantities vanishes unless P is elementary abelian. We can compute $a(P)$ and $n(P)$ simultaneously by letting $\mathcal{P}$ be either the collection of all A-invariant subgroups of P or the collection of all A-invariant normal subgroups of P. If we apply Corollary 6.1 with $x = 1$ and $y = P$, we see that either $a(P) = \mu(1, P)$ or $n(P) = \mu(1, P)$, depending on our choice of the poset $\mathcal{P}$.

By the defining properties of Möbius functions, and since we can assume that $1 < P$, we see that $\mu(1, P) = -\sum \mu(1, Q)$, where Q runs over $\mathcal{P} - \{P\}$. We can now apply Corollary 6.2 to the subgroup collection $\mathcal{P} - \{P\}$ to deduce that $\mu(1, P) = 0$ unless the intersection of the maximal members of $\mathcal{P} - \{P\}$ is trivial.

Suppose now that M is any maximal member of $\mathcal{P} - \{P\}$. Then M is either maximal among all A-invariant proper subgroups of P, or else it is maximal among all normal A-invariant proper subgroups of P. In the former case, we see that since $\mathbf{N}_P(M) > M$ is A-invariant, we have $\mathbf{N}_P(M) = P$. In either case, therefore, M is normal, and it is maximal among proper A-invariant normal subgroups of P. It follows that P/M is elementary abelian. We know that $\mu(1, P) = 0$ unless the intersection of all possible choices for M is trivial, and thus with either definition of $\mathcal{P}$, we see that $\mu(1, P) = 0$ unless P is elementary abelian. The result now follows. □

The following fairly standard result is essentially Lemma 2.7 of [**2**].

Lemma 6.4. *Let $\mathcal{P}$ be a poset of subgroups of a group E, and assume that both E and the trivial subgroup 1 are members of $\mathcal{P}$. Suppose that $U \subseteq E$ is a subgroup such that $\langle U, P\rangle \in \mathcal{P}$ for every member $P \in \mathcal{P}$. If $U > 1$, then $\sum_P \mu(1, P) = 0$, where P runs over all members of $\mathcal{P}$ such that $\langle U, P\rangle = E$.*

Proof. More generally, we show that if $U \subseteq H \in \mathcal{P}$, then $\sum_P \mu(1, P) = 0$, where this sum runs over all members $P \in \mathcal{P}$ such that $\langle U, P\rangle = H$. Suppose this is false, and let H be minimal among counterexample members of $\mathcal{P}$. (Note that since the sum is nonzero and $U > 1$, we must have $H > 1$.)

By the defining properties of the Möbius function, we know that

$$0 = \sum_{\substack{P\in\mathcal{P}\\ 1\subseteq P\subseteq H}} \mu(1, P) = \sum_{\substack{P\in\mathcal{P}\\ \langle U,P\rangle = H}} \mu(1, P) + \sum_{\substack{P\in\mathcal{P}\\ \langle U,P\rangle < H}} \mu(1, P).$$

We can decompose the second sum on the right as a sum of sums, one for each possibility for the group $\langle U, P\rangle < H$. Since each of these subsums vanishes by the minimality of H, the result follows. □

In order to complete the proof of Theorem 5.1, we introduce some additional notation. Recall that for subgroups $K \subseteq G$, we defined $f(K)$ to be the sum of $(-1)^{|\sigma|}$ over all triples $(g, x, \sigma) \in \mathcal{T}$ for which $K = \langle g, x, \sigma\rangle$. For each subgroup $A \subseteq G$, we now define $f_A(K)$ to be the subsum corresponding to those triples $(g, x, \sigma) \in \mathcal{T}$ for which $\langle g, x\rangle = A$. (Note that $f_A(K) = 0$ unless A is an abelian two-generator subgroup of K and $K = PA$ for some p-subgroup $P \lhd K$.) We can thus write $f(K) = \sum_A f_A(K)$, where this sum runs over all subgroups $A \subseteq K$. By the defining properties of the collection $\mathcal{T}$ of triples, we see that we can factor $f_A(K) = u(A)v_A(K)$, as follows. We set $u(A)$

to be the number of pairs (g, x) of elements of G such that $gx = xg$, x is p-regular and $\langle g, x\rangle = A$, and we define $v_A(K) = \sum_\sigma (-1)^{|\sigma|}$, where σ runs over normal chains of nonidentity A-invariant p-subgroups of K such that $\max(\sigma) \cdot A = K$.

If $f(K) \neq 0$, then certainly, $f_A(K) \neq 0$ for some subgroup A, and thus $u(A) \neq 0$ and $v_A(K) \neq 0$. Since $u(A) \neq 0$, we see that A is an abelian two-generator group, and because $v_A(K) \neq 0$, we know that $K = PA$ for some p-subgroup $P \lhd K$. The following result therefore includes Theorem 5.1.

Theorem 6.5. *Assume that $v_A(K) \neq 0$ where A is abelian. Then A is a p'-group, and $K = EA$, where E is a normal elementary abelian Sylow p-subgroup of K.*

Proof. As we have seen, we must have $K = PA$ for some p-subgroup $P \lhd K$. In particular, K has a normal Sylow p-subgroup E, and our task is to show that E must be elementary abelian and that p does not divide $|A|$.

We have $0 \neq v_A(K) = \sum_\sigma (-1)^{|\sigma|}$, where σ runs over normal chains of nonidentity A-invariant p-subgroups of K such that $\max(\sigma) \cdot A = K$. If P is any A-invariant p-subgroup of K, then, as before, we write $n(P)$ to denote the subsum of $v_A(K)$ corresponding to those normal chains σ for which $\max(\sigma) = P$, and we write $a(P)$ to denote the corresponding sum, but without the normality requirement on the chains. By Corollary 6.3 and Lemma 6.1, we have

$$0 \neq v_A(K) = \sum_P n(P) = \sum_P a(P) = \sum_P \mu(1, P),$$

where p runs over the A-invariant p-subgroups of K such that $PA = K$, and where the Möbius function μ is computed with respect to the poset $\mathcal{P}$ of all A-invariant p-subgroups of K. Since E is a normal Sylow p-subgroup of K, we see that $\mathcal{P}$ can also be described as the poset of all A-invariant subgroups of E.

Write $U = A \cap E$, and note that the condition that $PA = K$ is equivalent to $PU = E$. Since we know that the sum of $\mu(1, P)$ over all members $P \in \mathcal{P}$ for which $PU = E$ is nonzero, it follows by Lemma 6.4 that we must have $U = 1$, and thus A is a p'-group. We conclude that the only possibility for P is that $P = E$, and hence $0 \neq v_A(K) = a(E)$. It follows by Corollary 6.3 that E is elementary abelian, as required. $\square$

7. A variation on Theorem 4.1. Let us reexamine the quantity $R(G)$ that appeared on the right side of equation $(**)$ in Theorem 4.1. Recall that $R(G)$ was the sum of the quantity $(-1)^{|\sigma|}|G_\sigma|l(G_\sigma)$ over all normal chains $\sigma \in \mathcal{N}(G)$. What happens if we replace the quantity $l(G_\sigma)$ by $k(G_\sigma)$ in this sum? In other words, suppose that, instead of computing the "alternating sum" of the number of p-regular classes in the chain stabilizers, we were to compute the corresponding alternating sum of the total number of classes in the chain stabilizers. Perhaps surprisingly, the effect of this change is nil; the corresponding sum has exactly its original value. To see why this is so, let us write $S(G)$ to denote the analog of $R(G)$ obtained by replacing $l(G_\sigma)$ by $k(G_\sigma)$.

In Section 5, we obtained the simpler formula $R(G) = \sum f(K)$, where K runs over all subgroups of G, and $f(K)$ is the sum of $(-1)^{|\sigma|}$ over all triples $(g, x, \sigma) \in \mathcal{T}$ such that $\langle g, x, \sigma \rangle = K$. (Recall that $\mathcal{T}$ was defined to be the set of triples (g, x, σ), where g and x are commuting elements that stabilize the normal chain σ, and where x is required to be p-regular. Exactly the same reasoning leads us to consider the set of $\mathcal{S}$ of triples (g, h, σ), defined analogously to $\mathcal{T}$, but without the requirement that the second component must be p-regular. We get $S(G) = \sum e(K)$, where again we sum over all subgroups K, and we define $e(K)$ to be the sum of $(-1)^{|\sigma|}$ over all triples $(g, h, \sigma) \in \mathcal{S}$ such that $K = \langle g, h, \sigma \rangle$.

Recall that we proved Theorem 5.1 by decomposing the sum defining $f(K)$ and writing $f(K) = \sum_A f_A(K)$, where A runs over all subgroups of K and $f_A(K)$ is the sum of $(-1)^{|\sigma|}$ over all of those triples $(g, x, \sigma) \in \mathcal{T}$ such that $\langle g, x \rangle = A$. Analogously, we can write $e(K) = \sum_A e_A(K)$, where $e_A(K)$ is the sum of $(-1)^{|\sigma|}$ over all triples $(g, h, \sigma) \in \mathcal{S}$ for which $\langle g, h \rangle = A$.

We factored $f_A(K) = u(A)v_A(K)$, where $u(A)$ counts the number of pairs (g, x) of commuting elements of G such that x is p-regular and $\langle g, x \rangle = A$, and where $v_A(K)$ is the sum of $(-1)^{|\sigma|}$ over all normal A-invariant chains σ of nonidentity p-subgroups of K such that $K = \max(\sigma) \cdot A$. We see that the second factor of the analogous factorization of $e_A(K)$ is exactly the same number $v_A(K)$. We can thus write $e_A(K) = w(A)v_A(K)$, where we define $w(A)$ to be the total number of commuting pairs (g, h) of elements of G such that $\langle g, h\rangle = A$. Since there is no requirement that h should be p-regular in the definition of $w(A)$, we see that, in general, we have $w(A) \geq u(A)$, but if $|A|$ is not divisible by p, then clearly $w(A) = u(A)$.

We proved in Theorem 6.5 that if $v_A(K)$ is nonzero, then p cannot divide $|A|$. We know in this case that $w(A) = u(A)$, and it follows that for all subgroups $A \subseteq G$ and $K \subseteq G$ we have

$$f_A(K) = u(A)v_A(K) = w(A)v_A(K) = e_A(K),$$

and thus $f(K) = e(K)$ for all subgroups K. We conclude from this that $R(G) = S(G)$ for all groups G. Since $R(G)$ was exactly the right side of equation $(**)$ in Theorem 4.1, we can now restate that result in terms of the full class numbers of the appropriate subgroups instead of the numbers of p-regular classes.

Theorem 7.1 (Knörr-Robinson). *Assume that the* AWC *holds for all proper sections of G. Then the* AWC *holds for G if and only if*

$$|G|z(G) = \sum_{\sigma\in\mathcal{N}(G)} (-1)^{|\sigma|}|G_\sigma|k(G_\sigma).$$

We stress that our proof of the equality $R(G) = S(G)$ is independent of the AWC; it is the fact that this quantity is equal to $|G|z(G)$ that depends on Alperin's conjecture.

REFERENCES

1. R. Brauer, *Representations of finite groups*, Wiley, New York, 1963.

2. T. Hawkes, I.M. Isaacs and M. Ozaydm, *On the Möbius function of a finite group*, Rocky Mountain J. Math. **19** (1989), 1003–1034.

3. K.A. Hirsch, *On a theorem of Burnside*, Quart. J. Math. Oxford **1** (1950), 97–99.

4. I.M. Isaacs and G. Navarro, *Weights and vertices for characters of π-separable groups*, J. Algebra **177** (1995), 339–366.

5. R. Knörr and G.R. Robinson, *Some remarks on a conjecture of Alperin*, J. London Math. Soc. **38** (1989), 48–60.

6. G.R. Robinson, *The number of blocks with a given defect group*, J. Algebra **84** (1983), 493–502.

7. J. Thévenaz, *Locally determined functions and Alperin's conjecture*, J. London Math. Soc. **45** (1992), 446–468.

Mathematics Department, University of Wisconsin, 480 Lincoln Drive,
Madison, WI 53706
E-mail address: isaacs@math.wisc.edu