Generalizations of Fermat's Little Theorem via Group Theory

I. M. Isaacs, M. R. Pournaki*

Abstract

Let p be a prime number and a be an integer. Fermat's little theorem states that $a^p \equiv a \pmod{p}$. This result is generally established by an appeal to the theorem of elementary group theory that asserts that $x^{|G|} = 1$ for every element x of a finite group G. In this note we describe another way that group theory can be used to establish Fermat's little theorem and related results.

Keywords: Finite group, Congruence, Arithmetic function.

2000 Mathematics Subject Classification: Primary 20A05; Secondary 11A07, 11A25.

1 Introduction

As is well known, Fermat's so-called little theorem asserts that if p is prime, then $a^p \equiv a \pmod{p}$ for all integers a. This is usually proved by considering two cases: the result is trivial if $a \equiv 0 \pmod{p}$, and it follows from the fact that $a^{p-1} \equiv 1 \pmod{p}$ if a is not divisible by p. This latter fact is generally established by an appeal to the theorem of elementary group theory that asserts that $a^{|G|} = 1$ for every element a of a finite group G.

In this note we describe another way that group theory can be used to establish Fermat's little theorem and related results. In particular, this approach yields the following striking generalization, in which the modulus is not restricted to being prime. We feel that this result and this technique of proof both deserve to be more widely known than they seem to be.

Theorem A *Let a be an arbitrary integer. Then for every positive integer n,*

$$\sum_{d|n} \mu(n/d)a^d \equiv 0 \pmod{n}, \tag{1}$$

where μ is the Möbius function.

Recall that the Möbius function μ is defined so that $\mu(n) = 0$ unless n is square-free, in which case $\mu(n) = (-1)^t$, where t is the number of (distinct) prime divisors of n. In particular, $\mu(1) = 1$ and $\mu(p) = -1$ if p is prime. If n is prime, therefore,

*The research of the author was in part supported by a grant from IPM.

the left side of equation (1) is $a^n - a$, and hence Theorem A is exactly Fermat's little theorem in this case.

A history of equation (1) is given in Dickson's *History of the Theory of Numbers* [1, pp. 82-86]. According to Dickson, the case of Theorem A where a is prime was established by Gauss, and his proof was published posthumously in 1863. But it was not until around 1880 that a proof of the full result appeared. In fact, in the years 1880-83 four independent proofs were published by Kantor, Weyr, Lucas, and Pellet. Other proofs continue to be found; for example, there is a fairly recent proof of a somewhat more general theorem by C. J. Smyth in this MONTHLY [3]. (Smyth's "coloring proof", which generalizes ideas of Petersen and Thue, is related to the more general group-theoretic technique that we present here.)

2 Orbit Counting

Let G be a finite group that acts on a finite set Ω. Recall that this means that each element g of G effects a mapping $\alpha \longmapsto \alpha \cdot g$ on Ω such that $\alpha \cdot 1 = \alpha$ and $(\alpha \cdot g) \cdot h = \alpha \cdot (gh)$ for all points α of Ω and all elements g and h of G. As is well known, the action of G induces a partition of Ω into subsets called *orbits*, where the orbit containing the point α is the set $\{\alpha \cdot g : g \in G\}$.

The number N of orbits is easily computed if we know the number of points of Ω fixed by each element of G. The relevant formula is

$$N = \frac{1}{|G|} \sum_{g \in G} \pi(g)\,, \tag{2}$$

where we have written $\pi(g) = |\{\alpha \in \Omega : \alpha \cdot g = \alpha\}|$. The fixed-point counting function π is called the *permutation character* associated with the action of G on Ω, and we see that equation (2) tells us that the number N of orbits is exactly the average value of the permutation character over the group G. (We will prove a slight generalization of (2) later.) This orbit-counting formula is often credited to W. Burnside, but as was pointed out in [2], a more accurate attribution would be to Cauchy and Frobenius.

There are a number of combinatorial problems that can be reduced to counting orbits in an appropriate action, and of course, formula (2) provides the key to the solution of such problems. Since this counting technique was popularized by G. Pólya, his name, too, is occasionally attached to versions of this formula. (An example of a problem that is easily solved by "Pólya counting" is this: find the number of essentially different ways to assign colors to the faces of a cube if a palette of exactly n colors is available.)

For many of its applications, the Cauchy-Frobenius orbit-counting formula is not applied directly to the given action of G on Ω, but instead is used to count orbits of the induced action of G on the set S of all functions from Ω into some arbitrary finite set A. (Note that S can be viewed as the set of all possible colorings of the points of Ω with colors chosen from the set A.) To define the action of G on S, suppose that f belongs to S and let g be an element of G. The function $f \cdot g : \Omega \to A$ is constructed by setting $(f \cdot g)(\alpha) = f(\alpha \cdot g^{-1})$. Then $f \cdot g$ is a member of S, and it is routine to check that this defines an action of G on S.

Consider the permutation character χ associated with the action of G on S. (Thus $\chi(g)$ is the number of functions f in S such that $f \cdot g = f$.) It is easy to see that f is fixed by the group element g precisely when $f(\alpha) = f(\alpha \cdot g)$ for all points α of Ω, and thus f is g-fixed if and only if f is constant on each of the orbits of the cyclic group $\langle g \rangle$ acting on Ω.

Now for each element g of G, write $c(g)$ to denote the number of orbits of $\langle g \rangle$ on Ω, and note that $c(g)$ is the total number of cycles, including trivial "1-cycles", when the permutation of Ω induced by g is written in cycle notation. Thus, for example, if G is the symmetric group on six points and $\Omega = \{1, 2, 3, 4, 5, 6\}$ is the set on which G acts naturally, then for $g = (1\ 3)(2\ 4\ 6)$ we have $c(g) = 3$.

If $|A| = a$, it follows that $\chi(g) = a^{c(g)}$, and thus if we apply equation (2) to the permutation character χ, we see that

$$\frac{1}{|G|} \sum_{g \in G} \chi(g) = \frac{1}{|G|} \sum_{g \in G} a^{c(g)}$$

is the number of orbits of G on S. In particular, this quantity is an integer, and it follows that

$$\sum_{g \in G} a^{c(g)} \equiv 0 \pmod{|G|} \tag{3}$$

for every nonnegative integer a. (Note that we are given the action of G on Ω, and this determines the function $c(g)$, but the nonnegative integer $a = |A|$ is arbitrary because we get to choose the finite set A.)

For example, suppose that G is an arbitrary group of finite order n, and let G act on itself by right multiplication. If g in G has order $o(g) = m$, it is easy to see that the permutation induced by g on the set G consists of exactly n/m cycles of length m, and thus $c(g) = n/m = n/o(g)$. In this situation, congruence (3) yields

$$\sum_{g \in G} a^{n/o(g)} \equiv 0 \pmod{n} \tag{4}$$

for all nonnegative integers a.

Let us apply this general result in the case where G is cyclic of prime order p. Since G contains one element of order 1 and $p-1$ elements of order p, congruence (4) yields $0 \equiv a^p + (p-1)a \equiv a^p - a \pmod{p}$ for all integers $a \geq 0$. This is almost Fermat's little theorem, but it is missing the cases where $a < 0$. This is not a serious gap, however, since if $p > 2$, then replacing a with $-a$ in the expression $a^p - a$ yields the negative of the original expression.

Similarly, with every choice of the permutation group G we get a polynomial expression in a that is guaranteed to be divisible by $|G|$ for every nonnegative integer a. But actually, the restriction that $a \geq 0$ is never really necessary because of the following general lemma, which we will prove later.

Lemma B *Let $f(X)$ be a polynomial with rational coefficients and assume that $f(a)$ is in $\mathbb{Z}$ for all integers a such that $0 \leq a \leq d$, where $d = \deg(f)$. Then $f(a)$ is in $\mathbb{Z}$ for all integers a.*

We can use our technique to prove the following generalization of Fermat's little theorem. (Compare this with Theorem A.)

Theorem C *Let a be an arbitrary integer. Then for every positive integer n,*

$$\sum_{d|n} \varphi(n/d) a^d \equiv 0 \pmod{n}, \tag{5}$$

where φ is the Euler totient function.

Proof. Let G be a cyclic group of order n. Then for each divisor d of n there are exactly $\varphi(n/d)$ elements of order n/d in G, and so the result follows by congruence (4), where we appeal to Lemma B to handle negative values of a. $\square$

3 Homomorphisms

To prove Theorem A, we generalize the Cauchy-Frobenius orbit-counting formula. Let G act on Ω, as before, and fix a homomorphism λ from G into $\mathbb{C}^\times$, the multiplicative group of the complex numbers. If $\mathcal{O}$ is a G-orbit on Ω, we say that $\mathcal{O}$ is a λ-*good* if the stabilizer in G of every point in $\mathcal{O}$ is contained in $\ker(\lambda)$. Note that the stabilizers in G of the various points in $\mathcal{O}$ are conjugate in G, and so if any one of them is contained in the normal subgroup $\ker(\lambda)$, they all are, and $\mathcal{O}$ is λ-good. Note also that if λ is the trivial homomorphism, which maps every group element to the complex number 1, then every G-orbit is λ-good. The next theorem, therefore, includes the orbit-counting formula (2).

Theorem D *Let λ be a homomorphism from a finite group G into $\mathbb{C}^\times$. Suppose that G acts on some finite set Ω and let M be the number of λ-good orbits for this action. Then*

$$M = \frac{1}{|G|} \sum_{g \in G} \lambda(g)\pi(g) \,,$$

where π is the permutation character associated with the action.

We need the following easy fact, which is a special case of character orthogonality.

Lemma E *Let $\lambda : G \to \mathbb{C}^\times$ be a nontrivial homomorphism, where G is a finite group. Then*

$$\sum_{g \in G} \lambda(g) = 0 \,.$$

Proof. Let S be the sum in the statement of the lemma and choose h in G such that $\lambda(h) \neq 1$. Then

$$\lambda(h)S = \lambda(h) \sum_{g \in G} \lambda(g) = \sum_{g \in G} \lambda(hg) = S \,,$$

whence $S = 0$. $\square$

To prove Theorem D, we recall that if G acts on a set Ω and α is a point in the G-orbit $\mathcal{O}$, then $|\mathcal{O}| = |G|/|G_\alpha|$, where G_α is the stabilizer of α in G. This formula holds because there is a natural bijection τ from the set of right cosets in G of the subgroup G_α onto the G-orbit $\mathcal{O}$. The relevant map here is $\tau : (G_\alpha)g \longmapsto \alpha \cdot g$.

Proof of Theorem D. The permutation character π of G on Ω is the sum of the permutation characters of G acting on the various orbits. Now fix one orbit $\mathcal{O}$, and let σ be the permutation character of G acting on $\mathcal{O}$. Let

$$S = \sum_{g \in G} \lambda(g)\sigma(g),$$

and observe that it suffices to show that

$$S = \begin{cases} |G| & \text{if } \mathcal{O} \text{ is } \lambda\text{−good,} \\ 0 & \text{otherwise.} \end{cases}$$

Now in the sum S the complex number $\lambda(g)$ is counted $\sigma(g)$ times. It is counted once, therefore, for each ordered pair (x, g), where x is in $\mathcal{O}$, g is in G, and $x \cdot g = x$. It follows that

$$S = \sum_{x \in \mathcal{O}} \sum_{g \in G_x} \lambda(g) \,.$$

By Lemma E applied to the group G_x, the inner sum is zero if $G_x \not\subseteq \ker(\lambda)$ and it equals $|G_x|$ otherwise. Also, we note that the quantity $|G_x| = |G|/|\mathcal{O}|$ is constant as x runs over the orbit $\mathcal{O}$. If $\mathcal{O}$ is not λ-good, we get $S = 0$, as desired, and otherwise $S = |\mathcal{O}||G_x| = |G|$. $\square$

Now let G act on a finite set Ω. As before, let S be the set of all mappings from Ω into some finite set A of cardinality a and consider the induced action of G on S. We saw that the associated permutation character χ is given by the formula $\chi(g) = a^{c(g)}$. If we apply Theorem D in this situation and appeal to Lemma B to handle negative a, we obtain the following:

Corollary F *Let G be a finite group acting on a finite set Ω, and let λ be an arbitrary homomorphism from G into $\mathbb{C}^\times$. Then for each positive integer a,*

$$\sum_{g \in G} \lambda(g) a^{c(g)} \equiv 0 \pmod{|G|}, \tag{6}$$

where $c(g)$ is the number of orbits of $\langle g \rangle$ on Ω.

Next, we need a well known fact.

Lemma G *For each positive integer n, the sum of the primitive nth roots of unity in $\mathbb{C}$ is $\mu(n)$, where μ is the Möbius function.*

Proof. Write $F(n)$ to denote the sum of the primitive nth roots of unity in $\mathbb{C}$, and let $G(n)$ be the sum of *all* complex nth roots of unity. Then

$$G(n) = \sum_{m|n} F(m)\,,$$

and Möbius inversion yields

$$F(n) = \sum_{m|n} \mu(m) G(n/m)\,.$$

But $G(1) = 1$ and $G(m) = 0$ if $m > 1$. (Note that this fairly obvious fact is really a special case of Lemma E.) It follows that $F(n) = \mu(n)$, as asserted. $\square$

We can now prove our main result.

Proof of Theorem A. Let G be the (cyclic) group of order n consisting of the nth roots of unity in $\mathbb{C}$, and let $\lambda : G \to \mathbb{C}^\times$ be the identity map. Take $\Omega = G$ and let G act on Ω by right multiplication. Recall that in this situation we have $c(g) = d$ if g is an element of order n/d in G, and so the coefficient of a^d in congruence (6) is exactly the sum of all primitive (n/d)th roots of unity in $\mathbb{C}$. By Lemma G, this coefficient is $\mu(n/d)$, and the result follows. $\square$

4 Binomial Coefficients

What remains is to prove Lemma B, and for this we use a standard trick: we view binomial coefficients as polynomials. For nonnegative integers m, write

$$\binom{X}{m} = \frac{X(X-1)(X-2)\cdots(X-m+1)}{m!},$$

and observe that this is a polynomial of degree m with rational coefficients. It follows that the set

$$\left\{\binom{X}{m} \;:\; 0 \le m \le d\right\}$$

is a basis for the $\mathbb{Q}$-vector space consisting of all rational polynomials of degree at most d.

Proof of Lemma B. Since the given rational polynomial $f(X)$ has degree d, we can write

$$f(X) = \sum_{m=0}^{d} a_m \binom{X}{m},$$

where the coefficients a_m are rational. We know that $\binom{n}{n} = 1$ and $\binom{n}{m} = 0$ when $m > n$. Accordingly, if n is an integer with $0 \le n \le d$, we can write

$$f(n) = a_0\binom{n}{0} + a_1\binom{n}{1} + \cdots + a_{n-1}\binom{n}{n-1} + a_n\,.$$

By hypothesis, $f(n)$ is an integer when n is in this range, and since the binomial coefficients are also integers, it follows by induction on n that all of the coefficients a_n are integers.

To show that $f(a)$ is an integer for every integer a, we see now that it suffices to establish that the binomial-coefficient polynomials have this property. But the fact that the (usual) binomial coefficient $\binom{a}{m}$ is an integer when $a \ge m$ tells us that $m!$ divides every product of m consecutive positive integers. We conclude that $m!$ divides *every* product of m consecutive integers, and the result follows. $\square$

Acknowledgment: This work was done while the second author was a Postdoctoral Research Associate at the School of Mathematics, Institute for Studies in Theoretical Physics and Mathematics (IPM). He would like to thank the IPM for their financial support.

References

[1] L. E. Dickson, "*History of the Theory of Numbers*", vol. 1, Chelsea, New York, 1971.

[2] P. M. Neumann, *A Lemma that is not Burnside's*, Math. Sci. **4** (1979), no. 2, 133-141.

[3] C. J. Smyth, *A Coloring Proof of a Generalization of Fermat's Little Theorem*, Amer. Math. Monthly **93** (1986), no. 6, 469-471.

The Authors' Addresses

I. M. Isaacs, Mathematics Department, University of Wisconsin, 480 Lincoln Drive, Madison, WI 53706, USA.

E-mail address: `isaacs@math.wisc.edu`

M. R. Pournaki, School of Mathematics, Institute for Studies in Theoretical Physics and Mathematics, P.O. Box 19395-5746, Tehran, Iran.

E-mail address: `pournaki@ipm.ir`