Necklaces, Periodic Points and Permutation Representations

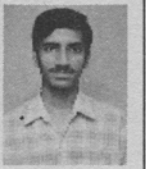
Fermat's Little Theorem

Somnath Basu, Anindita Bose, Sumit Sinha and Pankaj Vishe

S Basu

A Bose

S Sinha



P Vishe

The authors are second year students of BMath Hons in the Indian Statistical Institute, Bangalore.

One of the most ubiquitous results of elementary number theory is the so-called Fermat's little theorem. Pierre de Fermat was an amateur mathematician to whom are credited several deep and stunning results in mathematics. Let us list a few. He proved the impossibility of solving in non-zero integers the equation $X^4 + Y^4 = Z^2$ and, in the process, discovered the now-famous method of descent. He proved that a prime of the form $4n+1$ is a sum of two squares of integers. He is also credited, along with Newton and Leibniz, with the discovery of the integral calculus. The 'little theorem' of Fermat referred to above states that *for a prime number* p *and a number* a *which is not a multiple of* p, $a^{p-1} \equiv 1 \mod p$. The last phrase means that $a^{p-1} - 1$ is a multiple of p. This is also evidently equivalent to the statement that $a^p - a$ *is a multiple of* p *for any* a. This elementary proposition in classical number theory has, in fact, found applications to as applied a subject as cryptology. There are many ways to prove Fermat's little theorem. We discuss a number of them and go on to generalise one of them. Fix a prime p. We aim to prove that $a^p - a$ is a multiple of p for any integer a or, equivalently, that $a^{p-1} - 1$ is a multiple of p for all a coprime to p.

Induction

The method of mathematical induction provides perhaps the easiest proof. Assuming that a is any natural number satisfying $a^p - a$ is a multiple of p, let us look at $(a+1)^p - (a+1)$. This is simply $a^p - a$ upto a sum in which each term is a multiple of some binomial coefficient $\binom{p}{r}$, $1 \leq r < p$. It is an easy exercise to prove that each such binomial coefficient is a multiple of p (this is where we require p to be a prime). As the induction is trivially started with $a = 1$, and as the passage from a positive integer a to its negative $-a$ is evident, the proof is complete.

An attractive – one might say ornamental – proof of Fermat's little theorem can be obtained by counting necklaces.

Counting Necklaces

Here is an attractive - one might even say ornamental(!) - proof. Suppose there are a different bead colours which can be used to form a necklace of n beads. The only condition is that adjacent beads have different colours. First, let us just make a string (with two ends) of n beads with adjacent beads having different colours. As there are a choices of colours for the first bead, $a - 1$ choices for the second and $a - 1$ choices the next onwards, the number of arrangements possible is $a(a-1)^{n-1}$. Now, the first and the n-th bead may or may not have the same colour. If d_n (respectively, s_n) denotes the number of arrangements, where the first and the n-th bead have different colours (respectively, the same colour), then it is obvious that $s_n = d_{n-1}$. Moreover, $s_n + d_n = a(a-1)^{n-1}$ is the total number of arrangements. Therefore,

$$\begin{aligned} d_n + d_{n-1} &= a(a-1)^{n-1} \\ d_{n-1} + d_{n-2} &= a(a-1)^{n-2} \\ d_{n-2} + d_{n-3} &= a(a-1)^{n-3} \\ &\cdots\cdots\cdots\cdots\cdots \\ d_2 + d_1 &= a(a-1). \end{aligned}$$

Alternately adding and subtracting and noting that $d_1 = 0$, we get

$$d_n = (a-1)^n + (-1)^n(a-1).$$

If $n = p$, a prime, then each arrangement leads to p distinct cyclic combinations. Note that this is not so if

n is not a prime. Thus, we get that $(a-1)^p-(a-1)$ is a multiple of p.

The Euclidean algorithm tells us that the GCD of a and b can be expressed as $ax+by$.

Lagrange's Theorem – Some Group Theory

This is perhaps the most natural proof of this theorem. The set $\mathbf{Z_p} = \{0,1,\cdots,p-1\}$ has the structure of a ring viz., there are two operations – addition modulo p and multiplication modulo p. The special fact about primes that is relevant here is that all the numbers $1,2,\cdots,p-1$ are coprime to p. The Euclidean algorithm familiar from high school tells us that the greatest common divisor (GCD) of any a and p is of the form $ax+py$ for certain integers x,y. In particular, if $1 \leq i \leq p-1$, there is a corresponding j between 1 and $p-1$ so that $ij+py=1$ for some y. In terms of the operation of multiplication modulo p, this just means that i has a multiplicative inverse j. In other words, $\{1,2,\cdots,p-1\}$ form a group under the operation of multiplication modulo p. The famous theorem of Lagrange alluded to above implies that in any finite group G with n elements, every element g satisfies $g^n=e$, the identity element. Applying this to our group above, we get that a^{p-1} equals 1 modulo p for $a=1,2,\cdots,p-1$. But, clearly, Fermat's little theorem is an assertion on a only modulo p i.e., if it holds for some a, it holds for $a+$ (a multiple of p) as seen by using the binomial expansion.

Cauchy's Theorem – More Group Theory

A famous theorem of Lagrange implies that in a finite group G with n elements, every element of g satisfies $g^n=e$, the identity element.

Here is a variant where instead of using facts on the order of subgroups, one essentially uses the cardinality of orbits under a group action. Compare this proof with the one we got by counting necklaces. This approach has the additional attraction of providing us with a proof of the so-called theorem of Cauchy which asserts that *in a finite group with n elements, there are elements of every prime order which divides n.* To see both proofs simultaneously, look at an arbitrary finite group G with

n elements. Let p be a prime and, at present, we have not made any coprimality assumption about n and p. Consider the following subset of p-tuples of G:

$$S = \{(g_1, \cdots, g_p) : g_1 \cdots g_p = 1\},$$

where we have written 1 for the identity element of G. Evidently, $|S| = n^{p-1}$. Look at all cyclic permutations of p-tuples. In group-theoretic language, this corresponds to an action of the (cyclic) group of order p. Clearly, there are two types of orbits – those which have p elements and those which are singletons. This is easy to see from the first principles (where is it used that p is a prime?) but it also follows from the fact that the cardinality of an orbit of a finite group action is a divisor of the cardinality of the group itself – here $\mathbf{Z}_p$ acts on S by translations of the subscripts. Moreover, an orbit is a singleton if, and only if, it is of the form $(g, g, \cdots, g) \in S$. In particular, $g^p = 1$. There are two mutually exclusive and exhaustive cases – either n is coprime to p or n is a multiple of p.

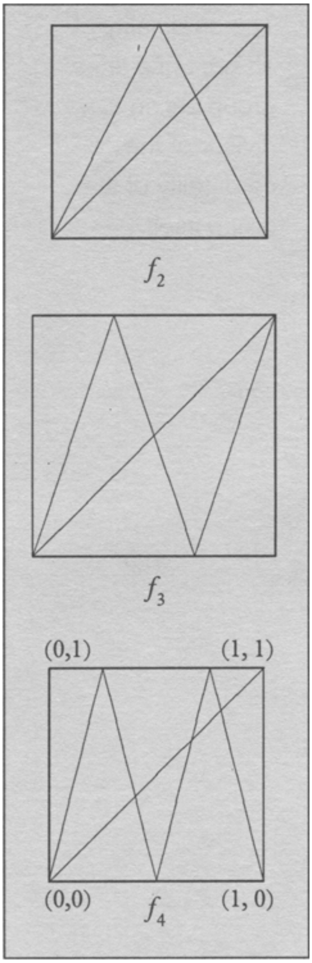
The cardinality of an orbit of a finite group action is a divisor of the cardinality of the group itself.

Look at the first case. Suppose $(g, g, \cdots, g) \in S$ is a singleton orbit. Then $g^p = 1$. As $g^n = 1$ from Lagrange's theorem and as $(n, p) = 1$, one obtains $g = 1$. Thus, there is exactly one singleton orbit in this case viz., $(1, 1, \cdots 1)$. As all other orbits have cardinality p, it follows that $n^{p-1} - 1$ is a multiple of p – Fermat's little theorem.

Now, we look at the second case when p divides n. Then, if there are d singleton orbits, one obtains $n^{p-1} - d$ as a multiple of p. As p divides n, it must divide d also. Clearly $d \geq 1$ since $(1, 1, \cdots, 1)$ is a singleton orbit. So, $d \geq p$. In particular, one obtains not only Cauchy's theorem but also the stronger fact that the number of elements g in G which satisfy $g^p = 1$ is a multiple of p.

In any finite group G, the number of elements which satisfy $g^p = 1$ for some prime p, dividing the order of g, is a multiple of p.

The general statement that for any divisor d of n, the number of elements g satisfying $g^d = 1$ is a multiple



of d, is also true. However, this result of Frobenius is very difficult and uses the representation theory of finite groups.

Periodic Points

Here is a geometric argument which possibly many readers would find more appealing.in comparison with the previous group-theoretic discussion. However, it is seen to bear a striking resemblance to the previous proof. We shall look at the graph of some functions $f : [0,1] \to [0,1]$. For any natural number $a \geq 2$, divide the unit interval into a parts and consider the piece-wise linear function f_a which connects, in turn, the point $(0,0)$ to $(1/a, 1)$, $(1/a, 1)$ to $(2/a, 0)$, etc. until we reach the right edge of the unit square. This will end up at $(1,0)$ or at $(1,1)$ depending on whether a is even or odd. For example, here are the graphs of f_2, f_3 and f_4. The graph of f_a intersects the diagonal joining $(0,0)$ to $(1,1)$ at exactly a points. These are the points (x,x), where $x = 2l/(a+1)$ with $0 \leq l \leq [a/2]$ and $x = 2l/(a-1)$ with $1 \leq l \leq [(a-1)/2]$. In fact, the line joining $((2l-1)/a, 1)$ to $(2l/a, 0)$ is $y = -ax + 2l$, which intersects the diagonal at the point with $x = 2l/(a+1)$. The line joining $(2l/a, 0)$ to $((2l+1)/a, 1)$ is $y = ax - 2l$ which intersects the diagonal at the point, where $x = 2l/(a-1)$.

Next, one notices that if $f_a^n = f_a(f_a(\cdots f_a)\cdots)$ is the function f_a iterated n times, then $f_a^n = f_{a^n}$. This can be routinely proved by induction. Indeed, one has $f_a(f_b(x)) = f_{ab}(x)$ for any a, b. Let us see how this relates to Fermat's little theorem. For any function $f : [0,1] \to [0,1]$ one has the disjoint union $\mathrm{Fix}(f^n) = \sqcup_{d/n} \mathrm{Ord}(f,d)$, where the left hand side is the set of fixed points of f^n and $\mathrm{Ord}(f,d) = \{x : f^d(x) = x, f^r(x) \neq x \;\forall\; 1 \leq r < d\}$. If one considers only functions which have finitely many fixed points, one has $|\mathrm{Fix}(f^n)| = \sum_{d/n} |\mathrm{Ord}(f,d)|$.

Now, observe that $|\mathrm{Ord}(f,d)|$ is a multiple of d for the following reason. If $x \in \mathrm{Ord}(f,d)$, then $x, f(x), \cdots,$

$f^{d-1}(x)$ are distinct points in Ord (f,d). This is the orbit of x under f and evidently, if $y \in \mathrm{Ord}(f,d)$, then its orbit is either identical with or disjoint from that of x. Note also that $\mathrm{Ord}(f,1) = \mathrm{Fix}(f)$.

Fixed points of iterates of a suitable function can be used to prove that a prime of the form 4*n*+1 is a sum of two squares which is another assertion of Fermat.

For instance, if $f = f_a$, we obtain $a^n = |\mathrm{Fix}(f_{a^n})| = |\mathrm{Fix}(f_a^n)|$. So, if $n = p$, a prime, then

$a^p = |\mathrm{Ord}(f_a,1)| + |\mathrm{Ord}(f_a,p)| = a + |\mathrm{Ord}(f,p)| \equiv a \bmod p$.

Thus, once again we have obtained Fermat's little theorem. Incidentally, if we write $a_n = |\mathrm{Ord}(f_a,n)|$ for a general n, then the Mobius inversion formula applied to $a^n = \sum_{d/n} a_d$ gives us the exact formula $a_n = \sum_{d/n} a^d \mu(n/d)$ for the number of points of order n for f_a. Observe that the fact $|\mathrm{Ord}(f_a,n)| \equiv 0 \bmod n$ simply means that for any natural numbers a, n one has

$$\sum_{d/n} a^d \mu(n/d) \equiv 0 \mod n.$$

It is an interesting exercise to prove this independently by elementary number theory. We mention in passing the delightful coincidence that fixed points of iterates of a suitable function can be used (see [1]) to prove that a prime of the form $4n+1$ is a sum of two squares which is another assertion of Fermat.

Permutation Representations

Now, we move on to a group-theoretic generalisation of Fermat's little theorem due to Strunkov. Let us recall some basic notions from finite group theory.

The permutation group (also called the symmetric group) on n letters is denoted by S_n. See [2] for some details on permutations and permutation groups. For our purpose, we only recall the following basic fact: *Any finite group G can be represented as a subgroup of S_n, where n is the order of G.* This is proved easily by looking at any element g of G as the corresponding permutation

In the more sophisticated language of representation theory of finite groups, the counting done here amounts to counting the multiplicity of the representation induced on G by the trivial representation of H in the representation of G on A^n.

obtained by multiplying (by g on the left) any element of *the underlying set G*. This is called the left regular representation of G and is an example of a permutation representation. Thus, a permutation representation of G is simply a homomorphism from G to some S_n. Similarly, if H is any subgroup of G, there is a permutation representation of G on the set of left cosets of H in G. Note that if $\alpha : G \to S_n$ is any permutation representation, then for any finite set A with a elements, say, there is a permutation representation $\beta : G \to S_{a^n}$ obtained in the following natural manner. Regard S_{a^n} as the symmetric group on the a^n elements which are all the n-tuples $(a_1, \cdots, a_n)$ with $a_i \in A$. Then, an element g corresponds to the permutation

$$\beta(g) : (a_1, \cdots, a_n) \to (a_{\alpha(g)(1)}, \cdots, a_{\alpha(g)(n)}).$$

Let H be a subgroup of G. Our aim is to count the number of subsets S of A^n, which admit a G-bijection with the set G/H of left cosets of H. Here, of course, G acts by left multiplication on G/H and, by a G-bijection, one means bijection which respects the permutation actions of G on both sides. Those who know the basic language of representation theory of finite groups would realise that this amounts to *counting the multiplicity of the representation induced on G by the trivial representation of H in the representation A^n of G.*

Let $S \subset A^n$ be such a subset and let $\pi : G/H \to S$ be a G-bijection. As G acts transitively on G/H (i.e., any two left cosets are permuted by some element of G), this must hold for the subset S also. This means that S is simply a G-orbit in A^n, say, $S = G \cdot (a_1, \cdots, a_n)$. We may assume that $\pi(H) = (a_1, \cdots, a_n)$, where H is the identity coset. Now, $\pi(gH) = g(a_1, \cdots, a_n)$ for all $g \in G$. Since π is well-defined, it follows that $(a_1, \cdots, a_n)$ is an H-invariant element. Furthermore, if $g(a_1, \cdots, a_n) = (a_1, \cdots, a_n)$ for some $g \in G$, then $gH = H$ i.e., $g \in H$. In other words, one might say that $(a_1, \cdots, a_n)$ is *exactly*

H-invariant. Conversely, suppose that $(a_1, \cdots, a_n)$ is exactly H-invariant. Then, if we define

$$\theta : G/H \to G \cdot (a_1, \cdots, a_n), \quad gH \mapsto g(a_1, \cdots, a_n),$$

then θ is well-defined as $(a_1, \cdots, a_n)$ is H-invariant and it is 1-1 since $(a_1, \cdots, a_n)$ is exactly H-invariant. Evidently, θ is onto.

To summarise, we have shown that a subset S of A^n admits a G-bijection with G/H if, and only if, $S = G \cdot (a_1, \cdots, a_n)$ and $(a_1, \cdots, a_n)$ is exactly H-invariant. We wish to count the number of these G-orbits. Thus, we have to scrutinise the exactly H-invariant elements and decide when two such give the same G-orbit. Suppose, $(a_1, \cdots, a_n)$ and $(b_1, \cdots, b_n) = x(a_1, \cdots, a_n)$ are both exactly H-invariant for some $x \in G$. Since $gH \mapsto g(b_1, \cdots, b_n) = gx(a_1, \cdots, a_n)$ is well defined, we must have $ghx(a_1, \cdots, a_n) = gx(a_1, \cdots, a_n)$ for all $h \in H$. In other words, $x^{-1}hx$ fixes $(a_1, \cdots, a_n)$ for any $h \in H$. As $(a_1, \cdots, a_n)$ is exactly H-invariant, one obtains $x^{-1}hx \in H \ \forall \ h \in H$ i.e. $x^{-1}Hx = x$ i.e, $x \in N_G(H)$, the normaliser of H in G. Of course, for $x, y \in N_G(H)$, we have $x(a_1, \cdots, a_n) = y(a_1, \cdots, a_n)$ if and only if, $y \in xH$. Hence, the number of subsets of A^n which admit a G-bijection with G/H is $e(H)\frac{|H|}{|N_G(H)|}$, where $e(H)$ denotes the number of exactly H-invariant elements in A^n. In particular, the number $e(H)\frac{|H|}{|N_G(H)|}$ is a positive integer(!)

This could give us useful information if we can compute $e(H)$ explicitly. First, note that since each H-invariant element is exactly K-invariant for a unique subgroup K containing H, one has $\sum_{K \supseteq H} e(K) = i(H)$, the total number of H-invariant elements. Also $i(G) = e(G)$. Now, it is easy to compute $i(K)$ for any subgroup K of G as follows. If the orbits of K in A^n correspond to the sets of subscripts $I_1, I_2, \cdots, I_{c(K)}$, then $\{1, 2, \cdots, n\} = \bigcup_{j=1}^{c(K)} I_j$ and $c(K)$ is the number of K-orbits. Clearly, then $(a_1, \cdots, a_n) \in A^n$ is K-invariant if, and only if, for each j, all the a_i's for $i \in I_j$ are equal. Thus, $i(K) = |A|^{c(K)} = a^{c(K)}$.

Now, the expressions $\sum_{K \supseteq H} e(K) = i(H)$ for any H, yield, by the inclusion-exclusion principle that $e(H) = \sum_{K \supseteq H} {}^{\mu}_{H}(K) i(K)$ where the *Mobius* function μ_H is defined on subgroups containing H by $\mu_H(H) = 1$ and $\sum_H \subseteq K \subseteq L \, {}^{\mu}_{H}(K) = 0$ for any subgroup $L \supset H$. Therefore, we have obtained that

$$e(H) = \sum_{K \supseteq H} {}^{\mu}_{H}(K) a^{c(K)}.$$

Combining this with the observation above that the number $e(H)\frac{|H|}{|N_G(H)|}$ is a positive integer, we obtain Strunkov's result:

$$\sum_{K \supseteq H} {}^{\mu}_{H}(K) a^{c(K)} \equiv 0 \mod \frac{|N_G(H)|}{|H|}.$$

Number-theoretic consequences can be derived from the above result. For instance, when G is $\mathbf{Z}_p$, H is trivial and $\alpha : G \to S_p$ is the regular representation, the result is just Fermat's little theorem. More generally, if G is cyclic of prime power order p^r and H is trivial, the result (once again for the regular representation) gives $a^{p^r} - a^{p^{r-1}} \equiv 0 \mod p^r$. Thus, if $(a, p) = 1$, we get $a^{\phi(p^r)} \equiv 1 \mod p^r$ where $\phi(n)$ is Euler's totient function, which counts the number of $m \leq n$, which are coprime to n. Thus, for any n and $(a, n) = 1$, one obtains by writing n as a product of prime powers that $a^{\phi(n)} \equiv 1 \mod n$. This is known as Euler's theorem.

Suggested Reading

[1] **B Bagchi, *Resonance*, Vol.4, No.7, pp.59-67, 1999.**

[2] **Jyoti Ramakrishnan, *Resonance*, Vol.5, No.11, p.88, 2000.**

[3] **S P Strunkov, *Math USSR Izvestiya*, Vol.38, pp.199-203, 1992.**

Address for Correspondence
Somnath Basu, Anindita Bose
Sumit Sinha and Pankaj Vishe
Indian Statistical Institute
8th Mile Mysore Road
Bangalore 560 059, India.