

# Fermat's Little Theorem: A Proof by Function Iteration

LIONEL LEVINE  
Harvard University  
Cambridge, MA 02138

It is a beautiful property of prime numbers, first proved more than three centuries ago by Fermat, that  $k^p \equiv k \pmod{p}$  for all prime numbers  $p$  and all integers  $k$ . Here we present a simple proof of Fermat's "little" theorem by considering iterates of the function  $f(z) = z^k$  on the complex plane. The method of proof has the advantage of generalizing the theorem to composite exponents: for every  $n$  we find a degree- $n$  polynomial, with coefficients  $\pm 1$ , that is always divisible by  $n$ . This is different from Euler's generalization ( $k^{\phi(n)} \equiv 1 \pmod{n}$  for  $k$  and  $n$  coprime). The method of proof is potentially more general still, since it is easily adapted to other functions  $f$ . Indeed, for any set  $S$ , every function  $f: S \rightarrow S$  satisfying a certain property corresponds to a divisibility result similar to Fermat's little theorem.

Let  $k$  be a positive integer and  $p$  be prime. Consider the function  $f(z) = z^k$  for complex  $z$ . The  $p$ th iterate of  $f$  is evidently  $f^p(z) = z^{k^p}$ . Let  $P_p$  be the set of those  $z$  that are fixed under  $f^p$  but not under  $f$  itself. Then  $|P_p| = k^p - k$ . But if  $z \in P_p$ , then  $f^i(z) \in P_p$  for every  $i = 0, 1, \dots, p-1$ ; and since  $p$  is prime, the  $p$  values  $z, f(z), \dots, f^{p-1}(z)$  are all distinct. Hence, we can partition  $P_p$  into equivalence classes, each containing  $p$  elements, obtaining

$$p \mid k^p - k, \quad (1)$$

Fermat's little theorem! The advantage to such an unusual approach is that it allows us to see a generalization that we might have missed otherwise. In general, if  $f^n(z) = z$ , then there must be some least positive integer  $d$  such that  $f^d(z) = z$ . Then  $d \mid n$ . Call this  $d$  the *order* of  $z$ . Let  $P_n$  be the set of all  $z$  of order  $n$ . As before, if  $z \in P_n$  then  $f^i(z) \in P_n$  for all  $i = 0, 1, \dots, n-1$ ; and the  $n$  values  $z, f(z), \dots, f^{n-1}(z)$  are all distinct because  $n$  is the *least* positive integer such that  $f^n(z) = z$ . Hence

$$n \mid |P_n| \quad (2)$$

for all positive integers  $n$ . In the case when  $n$  is prime, (2) reduces to (1), Fermat's little theorem. But when  $n$  is composite, (2) gives a different degree- $n$  polynomial, instead of  $k^n - k$ , that  $n$  must divide.

To illustrate what happens for general  $n$ , consider first the case  $n = pq$ , where  $p$  and  $q$  are distinct primes. There are  $k^{pq}$  values of  $z$  fixed under  $f^{pq}$ , and each such  $z$  has order  $d$  for exactly one  $d$  dividing  $pq$ . So

$$|P_{pq}| + |P_p| + |P_q| + |P_1| = k^{pq}.$$

Substituting  $|P_p| = k^p - k$ ,  $|P_q| = k^q - k$ , and  $|P_1| = k$ , and solving for  $|P_{pq}|$ , we get

$$|P_{pq}| = k^{pq} - k^p - k^q + k,$$

so by (2),  $pq \mid k^{pq} - k^p - k^q + k$ . So the polynomial  $k^{pq} - k^p - k^q + k$  in the case  $n = pq$  is the counterpart of the Fermat polynomial  $k^p - k$  in the case  $n = p$ . For

general  $n$ , there are  $k^n$  values of  $z$  fixed under  $f^n$  and every such  $z$  has order  $d$  for exactly one  $d$  dividing  $n$ , so

$$\sum_{d|n} |P_d| = k^n. \tag{3}$$

In their current form, the equations (3)—there is one equation for each  $n = 1, 2, 3, \dots$ —give an explicit formula for  $k^n$  in terms of the values  $|P_d|$ . What we'd like to do is “invert” (3) into an explicit formula for each  $|P_n|$  in terms of the powers of  $k$ . By (2), this will yield for each  $n$  a polynomial in  $k$  that is always divisible by  $n$ . The technique that accomplishes this task is called *Mobius inversion*:

Given two sequences  $\{a_n\}_{n \geq 1}$  and  $\{b_n\}_{n \geq 1}$  such that  $\sum_{d|n} a_d = b_n$ , Mobius inversion says that  $a_n = \sum_{d|n} \mu\left(\frac{n}{d}\right)b_d$ , where the function  $\mu$  is defined by  $\mu(p) = -1$  for  $p$  prime,  $\mu(p^m) = 0$  for  $m \geq 2$ , and  $\mu(ab) = \mu(a)\mu(b)$  for  $a, b$  coprime. (For further explanation of the Mobius function  $\mu$  and a proof of Mobius inversion, see [2].) Letting  $a_n = |P_n|$  and  $b_n = k^n$  in (3), we get  $|P_n| = \sum_{d|n} \mu\left(\frac{n}{d}\right)k^d$ , so by (2), we obtain our main result:

**THEOREM** (generalized form of Fermat's little theorem). *For all positive integers  $n$  and  $k$ ,  $n \mid \sum_{d|n} \mu\left(\frac{n}{d}\right)k^d$ .*

The method we used to prove this theorem can also be used to prove other such results. We applied the equation  $n \mid |P_n|$  to the particular function  $f(z) = z^k$ ; but in fact, the same argument shows that  $n \mid |P_n|$  holds whenever  $P_n$  is the set of points of order  $n$  for *any* function  $f$ . Let  $f$  be any function from a set  $S$  to itself such that  $f^n$  has finitely many fixed points for every  $n$ . If  $T(n)$  is the number of points fixed under  $f^n$ , then

$$n \mid \sum_{d|n} \mu\left(\frac{n}{d}\right)T(d) \tag{4}$$

for all positive integers  $n$ .

A final question: We have shown that (4) is a necessary condition for the sequence  $\{T(n)\}_{n \geq 1}$  to be of the form  $T(n) = |\{z \in S \mid f^n(z) = z\}|$  for some function  $f: S \rightarrow S$ . Is (4) a sufficient condition as well? In other words, given any sequence of nonnegative integers  $\{T(n)\}_{n \geq 1}$  satisfying (4), does there exist a function  $f: S \rightarrow S$  such that  $f^n$  has  $T(n)$  fixed points for every positive integer  $n$ ?

REFERENCES

1. R. Devaney, *A First Course in Chaotic Dynamical Systems*, Addison-Wesley, Reading, MA, 1992.
2. W. LeVeque, *Fundamentals of Number Theory*, Dover, Mineola, NY, 1977.