Proofs of Fermat's little theorem

This article collects together a variety of proofs of Fermat's little theorem, which states that

$$a^p \equiv a \pmod{p}$$

for every prime number p and every integer a (see modular arithmetic).

Contents

Simplifications

Some of the **proofs of Fermat's little theorem** given below depend on two simplifications.

The first is that we may assume that a is in the range $0 \le a \le p - 1$. This is a simple consequence of the laws of modular arithmetic; we are simply saying that we may first reduce a modulo p. This is consistent with reducing a^p modulo p, as one can check.

Secondly, it suffices to prove that

$$a^{p-1} \equiv 1 \pmod{p}$$

for a in the range $1 \le a \le p - 1$. Indeed, if the previous assertion holds for such a, multiplying both sides by a yields the original form of the theorem,

$$a^p \equiv a \pmod{p}$$

On the other hand, if $a = 0$, the theorem holds trivially.

Combinatorial proofs

Proof by counting necklaces

This is perhaps the simplest known proof, requiring the least mathematical background. It is an attractive example of a combinatorial proof (a proof that involves counting a collection of objects in two different ways).

The proof given here is an adaptation of Golomb's proof.[1]

To keep things simple, let us assume that a is a positive integer. Consider all the possible strings of p symbols, using an alphabet with a different symbols. The total number of such strings is a^p, since there are a possibilities for each of p positions (see rule of product).

For example, if $p = 5$ and $a = 2$, then we can use an alphabet with two symbols (say A and B), and there are $2^5 = 32$ strings of length 5:

AAAAA, *AAAAB*, *AAABA*, *AAABB*, *AABAA*, *AABAB*, *AABBA*, *AABBB*,
ABAAA, *ABAAB*, *ABABA*, *ABABB*, *ABBAA*, *ABBAB*, *ABBBA*, *ABBBB*,
BAAAA, *BAAAB*, *BAABA*, *BAABB*, *BABAA*, *BABAB*, *BABBA*, *BABBB*,
BBAAA, *BBAAB*, *BBABA*, *BBABB*, *BBBAA*, *BBBAB*, *BBBBA*, *BBBBB*.

We will argue below that if we remove the strings consisting of a single symbol from the list (in our example, *AAAAA* and *BBBBB*), the remaining $a^p - a$ strings can be arranged into groups, each group containing exactly p strings. It follows that $a^p - a$ is divisible by p.

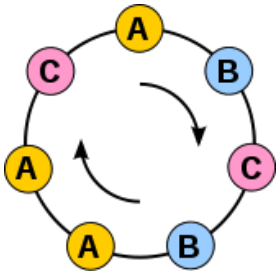
Necklaces

Let us think of each such string as representing a necklace. That is, we connect the two ends of the string together and regard two strings as the same necklace if we can rotate one string to obtain the second string; in this case we will say that the two strings are *friends*. In our example, the following strings are all friends:

AAAAB, *AAABA*, *AABAA*, *ABAAA*, *BAAAA*.

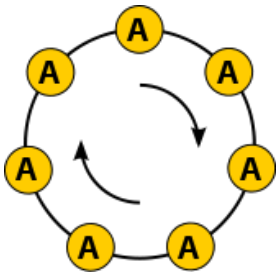
Similarly, each line of the following list corresponds to a single necklace.

AAABB, *AABBA*, *ABBAA*, *BBAAA*, *BAAAB*,
AABAB, *ABABA*, *BABAA*, *ABAAB*, *BAABA*,
AABBB, *ABBBA*, *BBBAA*, *BBAAB*, *BAABB*,
ABABB, *BABBA*, *ABBAB*, *BBABA*, *BABAB*,
ABBBB, *BBBBA*, *BBBAB*, *BBABB*, *BABBB*,
BAAAA, *AAAAB*, *AAABA*, *AABAA*, *ABAAA*,
AAAAA,
BBBBB.



Necklace representing seven different strings (*ABCBAAC*, *BCBAACA*, *CBAACAB*, *BAACABC*, *AACABCB*, *ACABCBA*, *CABCBAA*)

Notice that in the above list, each necklace with more than one symbol is represented by 5 different strings, and the number of necklaces represented by just one string is 2, i.e. is the number of distinct symbols. Thus the list shows very clearly why $32 - 2$ is divisible by 5.



Necklace representing only one string (*AAAAAAA*)

One can use the following rule to work out how many friends a given string S has:

> If S is built up of several copies of the string T, and T cannot itself be broken down further into repeating strings, then the number of friends of S (including S itself) is equal to the *length* of T.

For example, suppose we start with the string $S = ABBABBABBABB$, which is built up of several copies of the shorter string $T = ABB$. If we rotate it one symbol at a time, we obtain the following 3 strings:

ABBABBABBABB,
BBABBABBABBA,
BABBABBABBAB.

There aren't any others, because *ABB* is exactly 3 symbols long and cannot be broken down into further repeating strings.

Completing the proof

Using the above rule, we can complete the proof of Fermat's little theorem quite easily, as follows. Our starting pool of a^p strings may be split into two categories:

- Some strings contain p identical symbols. There are exactly a of these, one for each symbol in the alphabet. (In our running example, these are the strings $AAAAA$ and $BBBBB$.)
- The rest of the strings use at least two distinct symbols from the alphabet. If we can break up S into repeating copies of some string T, the length of T must divide the length of S. But, since the length of S is the prime p, the only possible length for T is also p. Therefore, the above rule tells us that S has exactly p friends (including S itself).

The second category contains $a^p - a$ strings, and they may be arranged into groups of p strings, one group for each necklace. Therefore, $a^p - a$ must be divisible by p, as promised.

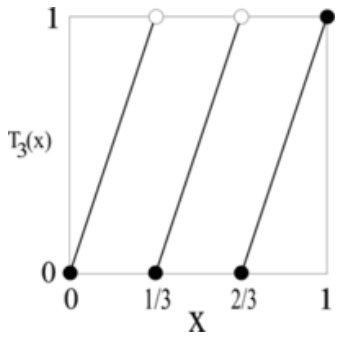
Proof using dynamical systems

This proof uses some basic concepts from dynamical systems.[2]

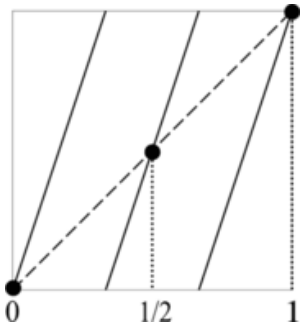
We start by considering a family of functions $T_n(x)$, where $n \geq 2$ is an integer, mapping the interval $[0, 1]$ to itself by the formula

$$T_n(x) = \begin{cases} \{nx\} & 0 \leq x < 1, \\ 1 & x = 1, \end{cases}$$

where $\{y\}$ denotes the fractional part of y. For example, the function $T_3(x)$ is illustrated below:



A number x_0 is said to be a *fixed point* of a function $f(x)$ if $f(x_0) = x_0$; in other words, if f leaves x_0 fixed. The fixed points of a function can be easily found graphically: they are simply the x coordinates of the points where the graph of $f(x)$ intersects the graph of the line $y = x$. For example, the fixed points of the function $T_3(x)$ are 0, 1/2, and 1; they are marked by black circles on the following diagram:



We will require the following two lemmas.

Lemma 1. For any $n \geq 2$, the function $T_n(x)$ has exactly n fixed points.

Proof. There are 3 fixed points in the illustration above, and the same sort of geometrical argument applies for any $n \geq 2$.

Lemma 2. For any positive integers n and m, and any $0 \leq x \leq 1$,

$$T_m(T_n(x)) = T_{mn}(x).$$

In other words, $T_{mn}(x)$ is the composition of $T_n(x)$ and $T_m(x)$.

Proof. The proof of this lemma is not difficult, but we need to be slightly careful with the endpoint $x = 1$. For this point the lemma is clearly true, since

$$T_m(T_n(1)) = T_m(1) = 1 = T_{mn}(1).$$

So let us assume that $0 \le x < 1$. In this case,

$$T_n(x) = \{nx\} < 1,$$

so $T_m(T_n(x))$ is given by

$$T_m(T_n(x)) = \{m\{nx\}\}.$$

Therefore, what we really need to show is that

$$\{m\{nx\}\} = \{mnx\}.$$

To do this we observe that $\{nx\} = nx - k$, where k is the integer part of nx; then

$$\{m\{nx\}\} = \{mnx - mk\} = \{mnx\},$$

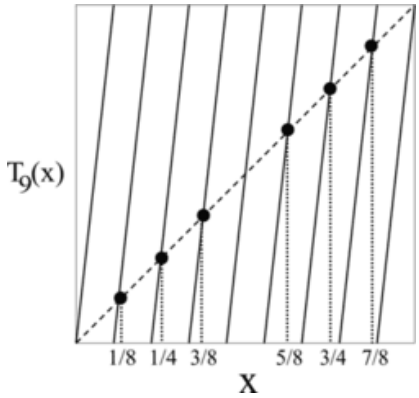
since mk is an integer.

Now let us properly begin the proof of Fermat's little theorem, by studying the function $T_{a^p}(x)$. We will assume that $a \ge 2$. From Lemma 1, we know that it has a^p fixed points. By Lemma 2 we know that

$$T_{a^p}(x) = \underbrace{T_a(T_a(\cdots T_a(x)\cdots))}_{p \text{ times}},$$

so any fixed point of $T_a(x)$ is automatically a fixed point of $T_{a^p}(x)$.

We are interested in the fixed points of $T_{a^p}(x)$ that are *not* fixed points of $T_a(x)$. Let us call the set of such points S. There are $a^p - a$ points in S, because by Lemma 1 again, $T_a(x)$ has exactly a fixed points. The following diagram illustrates the situation for $a = 3$ and $p = 2$. The black circles are the points of S, of which there are $3^2 - 3 = 6$.



The main idea of the proof is now to split the set S up into its orbits under T_a. What this means is that we pick a point x_0 in S, and repeatedly apply T_a(x) to it, to obtain the sequence of points

$$x_0, T_a(x_0), T_a(T_a(x_0)), T_a(T_a(T_a(x_0))), \ldots.$$

This sequence is called the orbit of x_0 under T_a. By Lemma 2, this sequence can be rewritten as

$$x_0, T_a(x_0), T_{a^2}(x_0), T_{a^3}(x_0), \ldots.$$

Since we are assuming that x_0 is a fixed point of $T_{a^p}(x)$, after p steps we hit $T_{a^p}(x_0) = x_0$, and from that point onwards the sequence repeats itself.

However, the sequence *cannot* begin repeating itself any earlier than that. If it did, the length of the repeating section would have to be a divisor of p, so it would have to be 1 (since p is prime). But this contradicts our assumption that x_0 is not a fixed point of T_a.

In other words, the orbit contains exactly p distinct points. This holds for every orbit of S. Therefore, the set S, which contains $a^p - a$ points, can be broken up into orbits, each containing p points, so $a^p - a$ is divisible by p.

(This proof is essentially the same as the necklace-counting proof given above, simply viewed through a different lens: one may think of the interval [0, 1] as given by sequences of digits in base a (our distinction between 0 and 1 corresponding to the familiar distinction between representing integers as ending in ".0000..." and ".9999..."). T_{a^n} amounts to shifting such a sequence by n many digits. The

fixed points of this will be sequences that are cyclic with period dividing n. In particular, the fixed points of T_{a^p} can be thought of as the necklaces of length p, with T_{a^n} corresponding to rotation of such necklaces by n spots.

This proof could also be presented without distinguishing between 0 and 1, simply using the half-open interval [0, 1); then T_n would only have $n - 1$ fixed points, but $T_{a^p} - T_a$ would still work out to $a^p - a$, as needed.)

Multinomial proofs

Proof using the binomial theorem

This proof, due to Euler,[3] uses induction to prove the theorem for all integers $a \geq 0$.

The base step, that $0^p \equiv 0 \pmod{p}$, is trivial. Next, we must show that if the theorem is true for $a = k$, then it is also true for $a = k + 1$. For this inductive step, we need the following lemma.

Lemma. For any integers x and y and for any prime p, $(x + y)^p \equiv x^p + y^p \pmod{p}$.

The lemma is a case of the freshman's dream. Leaving the proof for later on, we proceed with the induction.

Proof. Assume $k^p \equiv k \pmod{p}$, and consider $(k+1)^p$. By the lemma we have

$$(k+1)^p \equiv k^p + 1^p \pmod{p}.$$

Using the induction hypothesis, we have that $k^p \equiv k \pmod{p}$; and, trivially, $1^p = 1$. Thus

$$(k+1)^p \equiv k+1 \pmod{p},$$

which is the statement of the theorem for $a = k+1$. ∎

In order to prove the lemma, we must introduce the binomial theorem, which states that for any positive integer n,

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i,$$

where the coefficients are the binomial coefficients,

$$\binom{n}{i} = \frac{n!}{i!(n-i)!},$$

described in terms of the factorial function, $n! = 1\times2\times3\times\cdots\times n$.

Proof of Lemma. We consider the binomial coefficient when the exponent is a prime p:

$$\binom{p}{i} = \frac{p!}{i!(p-i)!}$$

The binomial coefficients are all integers. The numerator contains a factor p by the definition of factorial. When $0 < i < p$, neither of the terms in the denominator includes a factor of p (relying on the primality of p), leaving the coefficient itself to possess a prime factor of p from the numerator, implying that

$$\binom{p}{i} \equiv 0 \pmod{p}, \qquad 0 < i < p.$$

Modulo p, this eliminates all but the first and last terms of the sum on the right-hand side of the binomial theorem for prime p. ∎

The primality of p is essential to the lemma; otherwise, we have examples like

$$\binom{4}{2} = 6,$$

which is not divisible by 4.

Proof using the multinomial expansion

The proof, which was first discovered by Leibniz (who did not publish it)[4] and later rediscovered by Euler,[3] is a very simple application of the multinomial theorem which is brought here for the sake of simplicity.

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{k_1,k_2,\ldots,k_m} \binom{n}{k_1, k_2, \ldots, k_m} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}.$$

The summation is taken over all sequences of nonnegative integer indices k_1 through k_m such the sum of all k_i is n.

Thus if we express a as a sum of 1s (ones), we obtain

$$a^p = \sum_{k_1,k_2,\ldots,k_a} \binom{p}{k_1, k_2, \ldots, k_a}$$

Clearly, if p is prime, and if k_j is not equal to p for any j, we have

$$\binom{p}{k_1, k_2, \ldots, k_a} \equiv 0 \pmod{p}$$

and

$$\binom{p}{k_1, k_2, \ldots, k_a} \equiv 1 \pmod{p}$$

if k_j is equal to p for some j.

Since there are exactly a elements such that $k_j = p$, the theorem follows.

(This proof is essentially a coarser-grained variant of the necklace-counting proof given earlier; the multinomial coefficients count the number of ways a string can be permuted into arbitrary anagrams, while the necklace argument counts the number of ways a string can be rotated into cyclic anagrams. That is to say, that the nontrivial multinomial coefficients here are divisible by p can be seen as a consequence of the fact that each nontrivial necklace of length p can be unwrapped into a string in p many ways.

This multinomial expansion is also, of course, what essentially underlies the binomial theorem-based proof above)

Proof using power product expansions

An additive-combinatorial proof based on formal power product expansions was given by Giedrius Alkauskas.[5] This proof uses neither the Euclidean algorithm nor the binomial theorem, but rather it employs formal power series with rational coefficients.

Proof as a particular case of Euler's theorem

This proof,[3][6] discovered by James Ivory[7] and rediscovered by Dirichlet[8] requires some background in modular arithmetic.

Let us assume that a is positive and not divisible by p. The idea is that if we write down the sequence of numbers

$$a, 2a, 3a, \ldots, (p-1)a \qquad \text{(A)}$$

and reduce each one modulo p, the resulting sequence turns out to be a rearrangement of

$$1, 2, 3, \ldots, p-1. \qquad \text{(B)}$$

Therefore, if we multiply together the numbers in each sequence, the results must be identical modulo p:

$$a \times 2a \times 3a \times \cdots \times (p-1)a \equiv 1 \times 2 \times 3 \times \cdots \times (p-1) \pmod{p}.$$

Collecting together the a terms yields

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}.$$

Finally, we may “cancel out” the numbers $1, 2, \ldots, p-1$ from both sides of this equation, obtaining

$$a^{p-1} \equiv 1 \pmod{p}.$$

There are two steps in the above proof that we need to justify:

- Why the elements of the sequence (**A**), reduced modulo p, are a rearrangement of (**B**), and
- Why it is valid to "cancel" in the setting of modular arithmetic.

We will prove these things below; let us first see an example of this proof in action.

An example

If $a = 3$ and $p = 7$, then the sequence in question is

$$3, 6, 9, 12, 15, 18;$$

reducing modulo 7 gives

$$3, 6, 2, 5, 1, 4,$$

which is just a rearrangement of

$$1, 2, 3, 4, 5, 6.$$

Multiplying them together gives

$$3 \times 6 \times 9 \times 12 \times 15 \times 18 \equiv 3 \times 6 \times 2 \times 5 \times 1 \times 4 \equiv 1 \times 2 \times 3 \times 4 \times 5 \times 6 \pmod{7};$$

that is,

$$3^6(1 \times 2 \times 3 \times 4 \times 5 \times 6) \equiv (1 \times 2 \times 3 \times 4 \times 5 \times 6) \pmod{7}.$$

Canceling out $1 \times 2 \times 3 \times 4 \times 5 \times 6$ yields

$$3^6 \equiv 1 \pmod{7},$$

which is Fermat's little theorem for the case $a = 3$ and $p = 7$.

The cancellation law

Let us first explain why it is valid, in certain situations, to "cancel". The exact statement is as follows. If u, x, and y are integers, and u is not divisible by a prime number p, and if

$$ux \equiv uy \pmod{p}, \qquad \textbf{(C)}$$

then we may "cancel" u to obtain

$$x \equiv y \pmod{p}. \qquad \textbf{(D)}$$

Our use of this **cancellation law** in the above proof of Fermat's little theorem was valid, because the numbers $1, 2, \ldots, p - 1$ are certainly not divisible by p (indeed they are *smaller* than p).

We can prove the cancellation law easily using Euclid's lemma, which generally states that if a prime p divides a product ab (where a and b are integers), then p must divide a or b. Indeed, the assertion (**C**) simply means that p divides $ux - uy = u(x - y)$. Since p is a prime which does not divide u, Euclid's lemma tells us that it must divide $x - y$ instead; that is, (**D**) holds.

Note that the conditions under which the cancellation law holds are quite strict, and this explains why Fermat's little theorem demands that p is a prime. For example, $2 \times 2 \equiv 2 \times 5 \pmod{6}$, but it is not true that $2 \equiv 5 \pmod{6}$. However, the following generalization of the cancellation law holds: if u, x, y, and z are integers, if u and z are relatively prime, and if

$$ux \equiv uy \pmod{z},$$

then we may "cancel" u to obtain

$$x \equiv y \pmod{z}.$$

This follows from a generalization of Euclid's lemma.

The rearrangement property

Finally, we must explain why the sequence

$$a, 2a, 3a, \ldots, (p-1)a,$$

when reduced modulo p, becomes a rearrangement of the sequence

$$1, 2, 3, \ldots, p-1.$$

To start with, none of the terms a, $2a$, ..., $(p-1)a$ can be congruent to zero modulo p, since if k is one of the numbers $1, 2, \ldots, p-1$, then k is relatively prime with p, and so is a, so Euclid's lemma tells us that ka shares no factor with p. Therefore, at least we know that the numbers a, $2a$, ..., $(p-1)a$, when reduced modulo p, must be found among the numbers $1, 2, 3, \ldots, p-1$.

Furthermore, the numbers a, $2a$, ..., $(p-1)a$ must all be *distinct* after reducing them modulo p, because if

$$ka \equiv ma \pmod{p},$$

where k and m are one of $1, 2, \ldots, p-1$, then the cancellation law tells us that

$$k \equiv m \pmod{p}.$$

Since both k and m are between 1 and $p-1$, they must be equal. Therefore, the terms a, $2a$, ..., $(p-1)a$ when reduced modulo p must be distinct. To summarise: when we reduce the $p-1$ numbers a, $2a$, ..., $(p-1)a$ modulo p, we obtain distinct members of the sequence $1, 2, \ldots, p-1$. Since there are exactly $p-1$ of these, the only possibility is that the former are a rearrangement of the latter.

Applications to Euler's theorem

This method can also be used to prove Euler's theorem, with a slight alteration in that the numbers from 1 to $p-1$ are substituted by the numbers less than and coprime with some number m (not necessarily prime). Both the rearrangement property and the cancellation law (under the generalized form mentioned above) are still satisfied and can be utilized.

For example, if $m = 10$, then the numbers less than m and coprime with m are 1, 3, 7, and 9. Thus we have:

$$a \times 3a \times 7a \times 9a \equiv 1 \times 3 \times 7 \times 9 \pmod{10}.$$

Therefore,

$$a^{\varphi(10)} \equiv 1 \pmod{10}.$$

Proof as a corollary of Euler's criterion

Proofs using group theory

Standard proof

This proof[9] requires the most basic elements of group theory.

The idea is to recognise that the set $G = \{1, 2, \ldots, p-1\}$, with the operation of multiplication (taken modulo p), forms a group. The only group axiom that requires some effort to verify is that each element of G is invertible. Taking this on faith for the moment, let us assume that a is in the range $1 \le a \le p-1$, that is, a is an element of G. Let k be the order of a, that is, k is the smallest positive integer such that $a^k \equiv 1 \pmod{p}$. Then the numbers 1, a, a^2, ..., a^{k-1} reduced modulo p form a subgroup of G whose order is k and therefore, by Lagrange's theorem, k divides the order of G, which is $p-1$. So $p-1 = km$ for some positive integer m and then

$$a^{p-1} \equiv a^{km} \equiv (a^k)^m \equiv 1^m \equiv 1 \pmod{p}.$$

To prove that every element b of G is invertible, we may proceed as follows. First, b is coprime to p. Thus Bézout's identity assures us that there are integers x and y such that $bx + py = 1$. Reading this equality modulo p, we see that x is an inverse for b, since $bx \equiv 1 \pmod{p}$. Therefore, every element of G is invertible. So, as remarked earlier, G is a group.

For example, when $p = 11$, the inverses of each element are given as follows:

a	1	2	3	4	5	6	7	8	9	10
a^{-1}	1	6	4	3	9	2	8	7	5	10

Euler's proof

If we take the previous proof and, instead of using Lagrange's theorem, we try to prove it in this specific situation, then we get Euler's third proof, which is the one that he found more natural.[10][11] Let A be the set whose elements are the numbers $1, a, a^2, ..., a^{k-1}$ reduced modulo p. If $A = G$, then $k = p - 1$ and therefore k divides $p - 1$. Otherwise, there is some $b_1 \in G \backslash A$.

Let A_1 be the set whose elements are the numbers $b_1, ab_1, a^2b_1, \ldots, a^{k-1}b_1$ reduced modulo p. Then A_1 has k distinct elements, because otherwise there would be two distinct numbers $m, n \in \{0, 1, ..., k - 1\}$ such that $a^m b_1 \equiv a^n b_1 \pmod{p}$, which is impossible, since it would follow that $a^m \equiv a^n \pmod{p}$. On the other hand, no element of A_1 can be an element of A, because otherwise there would be numbers $m, n \in \{0, 1, \ldots, k - 1\}$ such that $a^m b_1 \equiv a^n \pmod{p}$, and then $b_1 \equiv a^n a^{k-m} \equiv a^{n+k-m} \pmod{p}$, which is impossible, since $b_1 \notin A$.

So, the set $A \cup A_1$ has $2k$ elements. If it turns out to be equal to G, then $2k = p - 1$ and therefore k divides $p - 1$. Otherwise, there is some $b_2 \in G \backslash (A \cup A_1)$ and we can start all over again, defining A_2 as the set whose elements are the numbers $b_2, ab_2, a^2b_2, ..., a^{k-1}b_2$ reduced modulo p. Since G is finite, this process must stop at some point and this proves that k divides $p - 1$.

For instance, if $a = 5$ and $p = 13$, then, since

- $5^2 = 25 \equiv 12 \pmod{13}$,
- $5^3 = 125 \equiv 8 \pmod{13}$,
- $5^4 = 625 \equiv 1 \pmod{13}$,

we have $k = 4$ and $A = \{1, 5, 8, 12\}$. Clearly, $A \neq G = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Let b_1 be an element of $G \backslash A$; for instance, take $b_1 = 2$. Then, since

- $2 \times 1 = 2$,
- $2 \times 5 = 10$,
- $2 \times 8 = 16 \equiv 3 \pmod{13}$,
- $2 \times 12 = 24 \equiv 11 \pmod{13}$,

we have $A_1 = \{2, 3, 10, 11\}$. Clearly, $A \cup A_1 \neq G$. Let b_2 be an element of $G \backslash (A \cup A_1)$; for instance, take $b_2 = 4$. Then, since

- $4 \times 1 = 4$,
- $4 \times 5 = 20 \equiv 7 \pmod{13}$,
- $4 \times 8 = 32 \equiv 6 \pmod{13}$,
- $4 \times 12 = 48 \equiv 9 \pmod{13}$,

we have $A_2 = \{4, 6, 7, 9\}$. And now $G = A \cup A_1 \cup A_2$.

Note that the sets A, A_1, and so on are in fact the cosets of A in G.

Notes

1. Golomb, Solomon W. (1956), "Combinatorial proof of Fermat's "Little" Theorem" (http://www.cimat.mx/~mmoreno/teaching/spring08/Fermats_Little_Thm.pdf) (PDF), *American Mathematical Monthly*, **63** (10): 718, doi:10.2307/2309563 (https://doi.org/10.2307%2F2309563), JSTOR 2309563 (https://www.jstor.org/stable/2309563)
2. Iga, Kevin (2003), "A Dynamical Systems Proof of Fermat's Little Theorem", *Mathematics Magazine*, **76** (1): 48–51, doi:10.2307/3219132 (https://doi.org/10.2307%2F3219132), JSTOR 3219132 (https://www.jstor.org/stable/3219132)

3. Dickson, Leonard Eugene (2005) [1919], "Fermat's and Wilson's theorems, generalizations, and converses; symmetric funcions of 1, 2, ..., $p-1$ modulo p", *History of the Theory of Numbers*, **I**, Dover, ISBN 978-0-486-44232-7, Zbl 1214.11001 (https://zbmath.org/?format=complete&q=an:1214.11001)
4. Vacca, Giovanni (1894), "Intorno alla prima dimostrazione di un teorema di Fermat", *Bibliotheca Mathematica*, 2nd series (in Italian), **8** (2): 46–48
5. Alkauskas, Giedrius (2009), "A Curious Proof of Fermat's Little Theorem", *American Mathematical Monthly*, **116** (4): 362–364, arXiv:0801.0805 (https://arxiv.org/abs/0801.0805), doi:10.4169/193009709x470236 (https://doi.org/10.4169%2F193009709x470236), JSTOR 40391097 (https://www.jstor.org/stable/40391097)
6. Hardy, G. H.; Wright, E. M. (2008), "Fermat's Theorem and its Consequences", *An Introduction to the Theory of Numbers* (6th ed.), Oxford University Press, ISBN 978-0-19-921986-5
7. Ivory, James (1806), "Demonstration of a theorem respecting prime numbers", *New Series of the Mathematical Depository*, **1** (II): 6–8
8. Lejeune Dirichlet, Peter Gustav (1828), "Démonstrations nouvelles de quelques théorèmes relatifs aux nombres", *Journal für die reine und angewandte Mathematik* (in French), **3**: 390–393
9. Weil, André; Rosenlicht, Maxwell (1979), "§ VIII", *Number Theory for beginners* (https://archive.org/details/numbertheoryforb0000weil), Springer-Verlag, doi:10.1007/978-1-4612-9957-8 (https://doi.org/10.1007%2F978-1-4612-9957-8), ISBN 978-0-387-90381-1, Zbl 0405.10001 (https://zbmath.org/?format=complete&q=an:0405.10001)
10. Weil, André (2007) [1984], "§ III.VI", *Number theory: An approach through history; from Hammurapi to Legendre*, Birkhäuser, ISBN 978-0-8176-4565-6, Zbl 1149.01013 (https://zbmath.org/?format=complete&q=an:1149.01013)
11. Euler, Leonhard (1761), "Theoremata circa residua ex divisione potestatum relicta" (http://math.dartmouth.edu/~euler/docs/originals/E262.pdf) (PDF), *Novi Commentarii Academiae Scientiarum Petropolitanae* (in Latin), **7**: 49–82
