Combinatorial Proofs for Some Number-Theoretic Facts

Koji Nuida

June 7, 2023 (first version)

Abstract

"Combinatorial proof" means a proof of equation for non-negative integers by counting the number of elements in some finite set in two different ways. In this note, we describe combinatorial proofs for some facts in number theory.

Notations

Let $\mathbb{Z}_{>0}$ denote the set of positive integers, and let $\mathbb{Z}_{\geq 0}$ denote the set of non-negative integers. For $n, m \in \mathbb{Z}$, we define $[n, m] := \{k \in \mathbb{Z} \mid n \leq k \leq m\}$. For a set S and $n \in \mathbb{Z}_{\geq 0}$, we write the set of all n-element subsets of S as $\binom{S}{n}$. That is, $\binom{S}{n} = \{T \subseteq S : |T| = n\}$. For $n \in \mathbb{Z}_{>0}$ and $a, b \in \mathbb{Z}$, we write $a \equiv_n b$ to mean $a \equiv b \pmod{n}$. Moreover, let $a \bmod n$ denote the remainder of $a \in \mathbb{Z}$ modulo $n \in \mathbb{Z}_{>0}$.

1 Warm-Up: Expression of Binomial Coefficients

First, as an example of the methodology of combinatorial proofs itself, we describe a proof for the explicit expression of binomial coefficients. Here, for non-negative integers $n, m \in \mathbb{Z}_{\geq 0}$, we define the binomial coefficient $\binom{n}{m}$ to be the number of the m-element subsets of an n-element set (e.g., $[1, n]$). By using the notation above, it can be expressed as $\binom{n}{m} = |\binom{[1,n]}{m}|$. This value is, by definition, a non-negative integer. We describe a combinatorial proof of the following well-known expression of binomial coefficients. We note that if $m > n$, then $\binom{n}{m} = 0$.

Proposition 1. If $n, m \in \mathbb{Z}_{\geq 0}$ and $m \leq n$, then $\displaystyle\binom{n}{m} = \frac{n!}{m!(n-m)!}$.

Proof. We enumerate the elements of the n-th symmetric group S_n in two ways. First, for $\sigma \in S_n$, there are n choices for $\sigma(1)$, there are $n-1$ choices for $\sigma(2)$, there are $n-2$ choices for $\sigma(3)$, and so on, and hence we have $|S_n| = n!$.

Secondly, we consider the following way of enumeration: (i) choose the set I of m numbers $\sigma(1), \dots, \sigma(m)$; (ii) determine the order of elements of I; and (iii) determine the order of the remaining elements not in I. There are $\binom{n}{m}$ choices for (i) by the definition of binomial coefficients. For each of them, there are $m!$ choices for (ii) and $(n-m)!$ choices for (iii). As these numbers are independent of I, the total number of elements of S_n is equal to $\binom{n}{m} \cdot m!(n-m)!$.

As a result, we have $n! = |S_n| = \binom{n}{m} \cdot m!(n-m)!$, which implies the claim by dividing both sides by $m!(n-m)!$. □

2 Fermat's Little Theorem

The statement of Fermat's Little Theorem is as follows (which is one of the equivalent formulations).

Theorem 1 (Fermat's Little Theorem)**.** Let p be a prime and $a \in \mathbb{Z}$. Then $a^p \equiv_p a$.

This is a famous result in elementary number theory, and some well-known proofs are one using the multiplicative group of the finite field $\mathbb{F}_p$ and one by mathematical induction using the expansion of $(a+1)^p$. Here we describe a combinatorial proof.

Proof. We may assume without loss of generality that $a > 0$, by adding some multiple of p to a if necessary. It suffices to show that $a^p - a$ is a multiple of p.

Let $X := [1, a]^p$, i.e., the set of sequences of length p on the set $\{1, 2, \dots, a\}$. We have $|X| = a^p$.

On the other hand, we consider the cyclic shift operation σ on sequences $x = (x_1, x_2, \dots, x_{p-1}, x_p) \in X$ defined by $\sigma(x) = (x_2, x_3, \dots, x_p, x_1) \in X$. This is a permutation on X with $\sigma^p = \mathsf{id}$. To analize the orbit decomposition of X by the action of the group $G := \langle \sigma \rangle$ of order p, we say that $x \in X$ is of type 1 if $\sigma(x) = x$, and of type 2 if $\sigma^k(x) \neq x$ for any $k \in [1, p-1]$ (note that both cannot be simultaneously satisfied, as $p \geq 2$). Now assume that there is an $x \in X$ not of type 1 nor type 2. As x is not of type 2, there is a $k \in [1, p-1]$ with $\sigma^k(x) = x$; we choose such a minimum k. As x is not of type 1 either, we have $2 \leq k \leq p - 1$. As p is prime, p is not a multiple of k, and by dividing p by k, we have $p = qk + r$ for some $q \in \mathbb{Z}_{\geq 0}$ and $r \in [1, k-1]$. Now $\sigma^k(x) = x$ and hence $\sigma^{qk}(x) = x$ by the choice of k, while $\sigma^p(x) = \sigma^{qk+r}(x) = x$. Comparing them implies that $\sigma^r(x) = x$, contradicting the minimality of k, as $1 \leq r < k$. Hence, each element of X is either of type 1 or of type 2.

For $x \in X$, being of type 1 is equivalent to that all components are equal, therefore the number of such elements of X is a. Hence the number of elements in X of type 2 is $a^p - a$. On the other hand, the set X_2 of elements in X of type 2 is invariant under the action of G, and each $x \in X_2$ has trivial fixing subgroup $G_x := \{\tau \in G \mid \tau(x) = x\} = \{\mathsf{id}\}$. Therefore X_2 is decomposed into the G-orbits each having cardinality $|G| = p$, implying that $|X_2| \equiv_p 0$. Hence we have $a^p - a \equiv_p 0$, as desired. □

3 On Divisors of Binomial Coefficients

Proposition 2. For $n, m \in \mathbb{Z}_{>0}$, if n is coprime to m, then $\binom{n}{m}$ is a multiple of n.

A special case of this proposition is a well-known fact that if p is prime and $k \in [1, p-1]$, then $\binom{p}{k}$ is a multiple of p. We note that the proof for Fermat's Little Theorem by mathematical induction using the expansion of $(a+1)^p$, briefly mentioned in Section 2, uses this fact, while our combinatorial proof above did not require this fact.

Proof. Let $X := \binom{[1,n]}{m}$. Then $|X| = \binom{n}{m}$ by the definition of binomial coefficients.

Let σ denote the cyclic permutation $(1\ 2\ \cdots\ n) \in S_n$ of length n. Then $G := \langle \sigma \rangle$ acts on X by $\sigma \cdot S = \{\sigma(s_1), \dots, \sigma(s_m)\}$ for $S = \{s_1, \dots, s_m\} \in X$. Each orbit of X by this action has order at most $|G| = n$. If each orbit has order precisely n, then the orbit decomposition implies that $|X| = \binom{n}{m}$ is a multiple of n, as desired. From now, we assume that there is an orbit in X with order less than n and deduce a contradiction. Let $S \in X$ be an element of this orbit.

By the choice of S, there is a $k \in [1, n-1]$ with $\sigma^k \cdot S = S$. We choose such a minimum k. Then $\sigma^k(a) \in S$ for each $a \in S$. Now by dividing n by k, we have $n = qk + r$ for some $q \in \mathbb{Z}_{\geq 0}$ and $r \in [0, k-1]$. For each $a \in S$, we have $\sigma^n(a) = a$ by the definition of σ, therefore $\sigma^{n+k-r}(a) = \sigma^{k-r}(a)$; while we have $n + k - r = (q+1)k$ and therefore $\sigma^k \cdot S = S$ by the choice of k, implying that $\sigma^{n+k-r}(a) \in S$. Hence we have $\sigma^{k-r}(a) \in S$. This implies that $\sigma^{k-r} \cdot S = S$, which contradicts the minimality of k if $r > 0$. Therefore we have $r = 0$ and k is a divisor of n. Let $\tau := \sigma^k$ and $d := n/k$. Then $\tau^d = \sigma^n = \mathsf{id}$. Moreover, for any $a \in S$ and $\ell \in [1, d-1]$, as $1 \leq k \cdot \ell < n$, we have $\tau^\ell(a) = \sigma^{k \cdot \ell}(a) \neq a$ by the definition of σ. This implies that each orbit of S by the action of $H := \langle \tau \rangle$ consists of precisely d elements, therefore $|S|$ is a multiple of d. However, now $|S| = m$ is coprime to n and d is a divisor of n with $d > 1$, a contradiction. This concludes the proof. □

We note that the converse of Proposition 2 does not hold; $\binom{10}{4} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$ gives a counterexample.

4 Lucas' Theorem

Lucas' Theorem [1] in elementary number theory is stated as follows. Here we describe a combinatorial proof.

Theorem 2 (Lucas' Theorem)**.** Let p be a prime and let $d \in \mathbb{Z}_{>0}$. Suppose that $n, m \in \mathbb{Z}_{\geq 0}$ can be expressed by d-digit p-ary expressions, say $n = (n_{d-1}n_{d-2}\cdots n_0)_p$, $m = (m_{d-1}m_{d-2}\cdots m_0)_p$ (where $n_i, m_i \in [0, p-1]$ and the most significant digits may be 0). Then we have

$$\binom{n}{m} \equiv_p \binom{n_{d-1}}{m_{d-1}}\binom{n_{d-2}}{m_{d-2}}\cdots\binom{n_0}{m_0} \ .$$

Proof. Let $X := \binom{[0,n-1]}{m}$. We have $|X| = \binom{n}{m}$ by the definition of binomial coefficients.

For $\ell \in [0, d-1]$ and $\alpha \in [0, n_\ell - 1]$, we define

$$Y(\ell, \alpha) := \{(n_{d-1}\cdots n_{\ell+1}\alpha *_{\ell-1}\cdots *_1 *_0)_p \in \mathbb{Z}_{\geq 0} \mid *_i \in [0, p-1] \text{ for any } i \in [0, \ell-1]\} \ .$$

They are disjoint and form a partition of $[0, n-1]$. For $k \in [0, d-2]$ and $x \in \mathbb{Z}_{\geq 0}$, we define $f_k(x)$ to be the number obtained by changing the k-th or lower digits x_k, …, x_1, x_0 to $p-1$ in the p-ary expression $x = (\cdots x_2x_1x_0)_p$. Moreover, for $x \in Y(\ell, \alpha)$, we define $\sigma_k(x)$ in a way that if $f_k(x) \leq n-1$, then $\sigma_k(x)$ is obtained by changing the k-th digit x_k of x to $x_k + 1 \bmod p$, and if $f_k(x) > n-1$, then $\sigma_k(x) = x$. Now for any $x \in Y(\ell, \alpha)$, if $k \leq \ell - 1$, then we have $f_k(x) \leq f_{\ell-1}(x) = (n_{d-1}\cdots n_{\ell+1}(\alpha+1)0\cdots 00)_p - 1 \leq n-1$, therefore x is not fixed by σ_k, and $\sigma_k(x) \in Y(\ell, \alpha)$ by the definition of $Y(\ell, \alpha)$. On the other hand, if $k \geq \ell$, then we have $f_k(x) \geq f_\ell(x) = (n_{d-1}\cdots n_{\ell+1}(p-1)\cdots(p-1)(p-1))_0 \geq n > n-1$, therefore $\sigma_k(x) = x$. This implies that the set $Y(\ell, \alpha)$ is invariant under any σ_k; each of $\sigma_\ell, \dots, \sigma_{d-2}$ fixes every element of $Y(\ell, \alpha)$, while each of $\sigma_0, \dots, \sigma_{\ell-1}$ fixes no element of $Y(\ell, \alpha)$. By this and the fact that $[0, n-1]$ is partitioned into the subsets $Y(\ell, \alpha)$, it follows that each σ_k is a permutation on $[0, n-1]$ with $\sigma_k^p = \mathsf{id}$.

We show that if $\ell_1 < \ell_2$, then σ_{ℓ_1} and σ_{ℓ_2} commute with each other. Indeed, for $x \in Y(\ell, \alpha)$, the argument in the previous paragraph implies the following:

- When $\ell > \ell_2$, we also have $\ell > \ell_1$. Therefore, both $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x)$ and $(\sigma_{\ell_2} \circ \sigma_{\ell_1})(x)$ are obtained by adding 1 (modulo p) to each of the ℓ_1-th and the ℓ_2-th digits of x, where they differ only in the order of the two additions. Hence we have $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) = (\sigma_{\ell_2} \circ \sigma_{\ell_1})(x)$.
- When $\ell_2 \geq \ell > \ell_1$, σ_{ℓ_2} fixes every element of $Y(\ell, \alpha)$, while $Y(\ell, \alpha)$ is invariant under the action of σ_{ℓ_1}. Hence we have $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) = \sigma_{\ell_1}(x) = (\sigma_{\ell_2} \circ \sigma_{\ell_1})(x)$.
- When $\ell_1 \geq \ell$, we also have $\ell_2 \geq \ell$, therefore both σ_{ℓ_1} and σ_{ℓ_2} fix x. Hence we have $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) = x = (\sigma_{\ell_2} \circ \sigma_{\ell_1})(x)$.

Hence we have $(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) = (\sigma_{\ell_2} \circ \sigma_{\ell_1})(x)$ in any case, therefore $\sigma_{\ell_1} \circ \sigma_{\ell_2} = \sigma_{\ell_2} \circ \sigma_{\ell_1}$. By this and the argument in the previous paragraph, the group G generated by $\sigma_0, \dots, \sigma_{d-2}$ is commutative and the map $(\mathbb{Z}/p\mathbb{Z})^{d-1} \to G$, $(e_0, e_1, \dots, e_{d-2}) \mapsto \sigma_0^{e_0}\sigma_1^{e_1}\cdots\sigma_{d-2}^{e_{d-2}}$ is a surjective group homomorphism. Hence by the isomorphism theorem for groups, the order $|G|$ of G is a divisor of $|(\mathbb{Z}/p\mathbb{Z})^{d-1}| = p^{d-1}$, which should be a power of the prime p.

We define the action of G on X by $\tau \cdot \{x_1, \dots, x_m\} := \{\tau(x_1), \dots, \tau(x_m)\}$. For the orbit decomposition of X by the action, each orbit has order equal to that of some quotient group of G, which is a power of the prime p as well as $|G|$. Hence, by considering the set $X_0 := \{S \in X \mid \tau \cdot S = S \text{ for any } \tau \in G\}$ of the fixed points by the action, any orbit in X involving an element of $X \setminus X_0$ has order divisible by p. Therefore we have $|X| \equiv_p |X_0|$. The remaining task is to show that $|X_0|$ is equal to the right-hand side of the statement.

Let $S \in X_0$. For $\ell \in [0, d-1]$ and $\alpha \in [0, n_\ell - 1]$, suppose that $S \cap Y(\ell, \alpha) \neq \emptyset$ and take its element x. By the argument above, each of $\sigma_0, \dots, \sigma_{\ell-1}$ fixes no element of $Y(\ell, \alpha)$. Therefore, by the definitions of these maps, all elements of $Y(\ell, \alpha)$ can be obtained by applying elements of G to x, and all of those elements belong to S, as $S \in X_0$. Therefore, either $S \cap Y(\ell, \alpha) = \emptyset$ or $Y(\ell, \alpha) \subseteq S$ holds. This implies that, by putting

$I_\ell := \{\alpha \in [0, n_\ell - 1] \mid Y(\ell, \alpha) \subseteq S\}$, we have $S = \bigcup_{\ell=0}^{d-1} \bigcup_{\alpha \in I_\ell} Y(\ell, \alpha)$. Conversely, by the argument above, each set $Y(\ell, \alpha)$ is invariant under the action of G, therefore any element $S \in X$ of this form belongs to X_0. This implies that an element of X_0 is determined solely by the choices of sets I_ℓ. Now put $c_\ell := |I_\ell|$. Then, as $|Y(\ell, \alpha)| = p^\ell$, the corresponding element $S \in X_0$ satisfies that $|S| = \sum_{\ell=0}^{d-1} c_\ell p^\ell = (c_{d-1} \cdots c_1 c_0)_p$. The latter value is equal to $|S| = m$ if and only if $c_\ell = m_\ell$ holds for every ℓ. This implies that $|X_0|$ is equal to the number of choices of m_ℓ elements for I_ℓ from the n_ℓ-element set $[0, n_\ell - 1]$ for all ℓ. The latter number is equal to the right-hand side $\binom{n_{d-1}}{m_{d-1}} \cdots \binom{n_1}{m_1} \binom{n_0}{m_0}$ of the claim, concluding the proof. □

References

[1] E. Lucas, "Théorie des Fonctions Numériques Simplement Périodiques", Amer. J. Math. **1**(3) (1878) 197–240