Combinatorial Proofs of Congruences

Ira M. Gessel

Department of Mathematics
Brandeis University

Joint Mathematics Meeting
January, 2010
In 1872 Julius Petersen published a proof of Fermat’s theorem \( a^p \equiv a \pmod{p} \), where \( p \) is a prime:
We take a wheel with \( p \) spokes and color each spoke in one of \( a \) colors. We call two colorings *equivalent* if one can be rotated into the other:
In 1872 Julius Petersen published a proof of Fermat’s theorem 
\[ a^p \equiv a \pmod{p}, \] 
where \( p \) is a prime:
We take a wheel with \( p \) spokes and color each spoke in one of \( a \) colors. We call two colorings *equivalent* if one can be rotated into the other:
An equivalence class of colorings has size 1 if and only if every spoke has the same color, so there are $a$ of these equivalence classes. Every other equivalence class contains $p$ different colorings, so

$$a^p = \text{the total number of colorings}$$

$$= \text{the number of colorings in equivalence classes of size } p$$

$$+ \text{ the number of colorings in equivalence classes of size 1}$$

$$\equiv a \pmod{p}$$
An equivalence class of colorings has size 1 if and only if every spoke has the same color, so there are $a$ of these equivalence classes. Every other equivalence class contains $p$ different colorings, so

$$a^p = \text{the total number of colorings}$$

$$= \text{the number of colorings in equivalence classes of size } p$$

$$+ \text{the number of colorings in equivalence classes of size 1}$$

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

How do we know that every equivalence class has size 1 or $p$?
An equivalence class of colorings has size 1 if and only if every spoke has the same color, so there are $a$ of these equivalence classes. Every other equivalence class contains $p$ different colorings, so

$$a^p = \text{the total number of colorings}$$

$$= \text{the number of colorings in equivalence classes of size } p$$

$$+ \text{the number of colorings in equivalence classes of size 1}$$

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

How do we know that every equivalence class has size 1 or $p$?

The equivalence classes are orbits under the action of a cyclic group of order $p$, and we know that the size of any orbit divides the order of the group.
In general if a group of order $p$, where $p$ is a prime, acts on a finite set $S$, then $|S|$ is congruent modulo $p$ to the number of fixed points.
In general if a group of order $p$, where $p$ is a prime, acts on a finite set $S$, then $|S|$ is congruent modulo $p$ to the number of fixed points.

Note that the same is true for a group of order $p^k$. 
In general if a group of order $p$, where $p$ is a prime, acts on a finite set $S$, then $|S|$ is congruent modulo $p$ to the number of fixed points.

Note that the same is true for a group of order $p^k$.

Another useful variation: If a group of order $n$ acts on a set $S$ then $|S|$ is congruent modulo $n$ to the number of elements in orbits of size $n$. 
We can get a variant of Petersen’s proof by using Burnside’s lemma (the Cauchy-Frobenius theorem) to count orbits:

\[ \frac{1}{p} \left( a^p + (p-1)a \right) \]

so this quantity must be an integer.

But the results we get from Burnside’s lemma aren’t in general as nice as those we get from counting fixed points.
We can get a variant of Petersen’s proof by using Burnside’s lemma (the Cauchy-Frobenius theorem) to count orbits: We find that the number of orbits is

\[ \frac{1}{p}(a^p + (p - 1)a), \]

so this quantity must be an integer.
We can get a variant of Petersen’s proof by using Burnside’s lemma (the Cauchy-Frobenius theorem) to count orbits: We find that the number of orbits is

\[ \frac{1}{p} \left( a^p + (p - 1)a \right), \]

so this quantity must be an integer.

But the results we get from Burnside’s lemma aren’t in general as nice as those we get from counting fixed points.
Petersen also gave a similar proof of Wilson’s theorem

\[(p - 1)! + 1 \equiv 0 \pmod{p}.
\]
Petersen also gave a similar proof of Wilson’s theorem 
\((p - 1)! + 1 \equiv 0 \pmod{p}\).

There are \((p - 1)!\) cyclic permutations of \(p\) points. The cyclic group \(C_p\) acts on them by conjugation, which we can view geometrically as rotation.
Petersen also gave a similar proof of Wilson’s theorem

\[(p - 1)! + 1 \equiv 0 \pmod{p}\].

There are \((p - 1)!\) cyclic permutations of \(p\) points. The cyclic group \(C_p\) acts on them by conjugation, which we can view geometrically as rotation

\[
\begin{align*}
\text{There are } p - 1 \text{ fixed cycles so } (p - 1)! &\equiv p - 1 \pmod{p}.
\end{align*}
\]
We can generalize the proof of Fermat’s to composite moduli.
We can generalize the proof of Fermat’s to composite moduli. If we take a wheel with $p^k$ spokes and color each spoke in one of $a$ colors, there are $a^{p^k}$ colorings. There are $a^{p^{k-1}}$ colorings that are in orbits of size less than $p^k$, so

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

a form of Euler’s theorem.
We can generalize the proof of Fermat’s to composite moduli.

If we take a wheel with $p^k$ spokes and color each spoke in one of $a$ colors, there are $a^{p^k}$ colorings. There are $a^{p^k-1}$ colorings that are in orbits of size less than $p^k$, so

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

a form of Euler’s theorem.

More generally, we can show that if we take a wheel with $n$ spokes, for any $n$, then the number of colorings in orbits of size $n$ is $\sum_{d|n} \mu(d) a^{n/d}$, so

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

(Gauss).
Another example of a combinatorial proof of a congruence is Lucas’s theorem:

If \( a = a_0 + a_1 p + \cdots + a_k p^k \) and \( b = b_0 + b_1 p + \cdots + b_k p^k \), where \( 0 \leq a_i, b_i < p \) then

\[
\binom{a}{b} \equiv \binom{a_0}{b_0} \binom{a_1}{b_1} \cdots \binom{a_k}{b_k} \pmod{p}.
\]
It’s convenient to prove a slightly different form of Lucas’s theorem: If $0 \leq b, d < p$ then

\[
\binom{ap + b}{cp + d} \equiv \binom{a}{c} \binom{b}{d} \pmod{p}.
\]
It’s convenient to prove a slightly different form of Lucas’s theorem: If $0 \leq b, d < p$ then
\[
\binom{ap + b}{cp + d} \equiv \binom{a}{c} \binom{b}{d} \pmod{p}.
\]

To prove this we take $ap + b$ boxes arranged in a $p \times a$ rectangle with an additional $b < p$ boxes.

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\hline
\hline
\hline
\end{array}
\]

\[p = 5\]
\[a = 4\]
\[b = 3\]

We choose $cp + d$ of the boxes, in $\binom{ap+b}{cp+d}$ ways.
We choose $cp + d$ of the boxes and mark them.

$p = 5$

$a = 4$

$b = 3$
We choose $cp + d$ of the boxes and mark them.

Now we rotate each of the $a$ columns of $p$ boxes independently. Each arrangement will be in an orbit of size divisible by $p$ except for those arrangements that consist only of full and empty columns. Since $b$ and $d$ are less than $p$, we must choose $d$ boxes from the $b$ additional boxes, and then choose $c$ whole columns from the $a$ columns, which can be done in \( \binom{a}{c} \binom{b}{d} \) ways.
We choose \( cp + d \) of the boxes and mark them.

Now we rotate each of the \( a \) columns of \( p \) boxes independently. Each arrangement will be in an orbit of size divisible by \( p \) except for those arrangements that consist only of full and empty columns. Since \( b \) and \( d \) are less than \( p \), we must choose \( d \) boxes from the \( b \) additional boxes, and then choose \( c \) whole columns from the \( a \) columns, which can be done in \( \binom{a}{c} \binom{b}{d} \) ways.
The same argument shows that \( \binom{ap}{cp} \equiv \binom{a}{c} \pmod{p^2} \), since if we are choosing \( cp \) boxes from the \( p \times a \) rectangle, if there is one incomplete column then there must be at least two incomplete columns.
The same argument shows that $\binom{ap}{cp} \equiv \binom{a}{c} \pmod{p^2}$, since if we are choosing $cp$ boxes from the $p \times a$ rectangle, if there is one incomplete column then there must be at least two incomplete columns.

In fact if $p \geq 5$ then $\binom{ap}{cp} \equiv \binom{a}{c} \pmod{p^3}$. The combinatorial approach reduces this to showing that $\binom{2p}{p} \equiv 2 \pmod{p^3}$. It’s probably impossible to prove this combinatorially, but here is a simple proof due to Richard Stanley.
\[
\binom{2p}{p} - 2 = \sum_{k=1}^{p-1} \binom{p}{k}^2 = \sum_{k=1}^{p-1} \left[ \frac{p}{k} \binom{p-1}{k-1} \right]^2 \\
= p^2 \sum_{k=1}^{p-1} \frac{1}{k^2} \binom{p-1}{k-1}^2
\]

Since \( \binom{p-1}{k-1} \equiv \binom{-1}{k-1} = (-1)^{k-1} \pmod{p} \), it’s enough to show that \( \sum_{k=1}^{p-1} 1/k^2 \) is divisible by \( p \). But

\[
\sum_{k=1}^{p-1} \frac{1}{k^2} \equiv \sum_{k=1}^{p-1} k^2 = \frac{1}{6} p(2p-1)(p-1) \equiv 0 \pmod{p}
\]

if \( p \neq 2 \) or \( 3 \).
The Catalan number \( C_n = \frac{1}{n+1} \binom{2n}{n} \) counts, among other things, binary trees with \( n \) internal vertices and \( n + 1 \) leaves. For example, if \( n = 3 \) one such tree is

![Binary tree diagram]
The Catalan number $C_n = \frac{1}{n+1} \binom{2n}{n}$ counts, among other things, binary trees with $n$ internal vertices and $n + 1$ leaves. For example, if $n = 3$ one such tree is

![Binary Tree Diagram]

When is $C_n$ odd?
A group of order $2^n$ acts on the binary trees counted by $C_n$: For each internal vertex we can switch the two subtrees rooted at its children:

![Diagram of binary trees](image.png)

The size of every orbit will be a power of two, and the only orbits of size 1 are for trees in which every leaf is at the same level: So there are $2^k$ leaves for some $k$, so $n = 2^k - 1$. Conversely, if $n = 2^k - 1$ then there is exactly one orbit of size 1, so $C_n$ is odd.
A group of order $2^n$ acts on the binary trees counted by $C_n$: For each internal vertex we can switch the two subtrees rooted at its children:

The size of every orbit will be a power of two, and the only orbits of size 1 are for trees in which every leaf is at the same level:
A group of order $2^n$ acts on the binary trees counted by $C_n$: For each internal vertex we can switch the two subtrees rooted at its children:

```
    ┌───┐
   ┤    ├
   └───┘
```

The size of every orbit will be a power of two, and the only orbits of size 1 are for trees in which every leave is at the same level:

```
    ┌───┐
   ┤    ├
    └───┘
```

So there are $2^k$ leaves for some $k$, so $n = 2^k - 1$. Conversely, if $n = 2^k - 1$ then there is exactly one orbit of size 1, so $C_n$ is odd.
Another class of applications of the combinatorial method is to sequences that counting “labeled objects” like permutations or graphs. For example, the derangement number $d_n$ is the number of permutations of $[n] = \{1, 2, \ldots, n\}$ with no fixed points:

<table>
<thead>
<tr>
<th>$n$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_n$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>44</td>
<td>265</td>
<td>1854</td>
<td>14833</td>
</tr>
</tbody>
</table>

We think of a derangement as a set of cycles, each of length greater than 1:

$(1 \ 3 \ 6) \ (2 \ 5) \ (4 \ 7)$

The cyclic group $C_n$ acts on the set of derangements of $[n + m]$ by cyclically permuting $1, 2, \ldots, n$:

For $n = 3$ a generator of $C_3$ takes $(1 \ 3 \ 6) \ (2 \ 5) \ (4 \ 7)$ to $(2 \ 1 \ 6) \ (3 \ 5) \ (4 \ 7)$
If a derangement has elements of $[n]$ and of $[m] + n = \{n + 1, n + 2, \ldots, n + m\}$ in the same cycle, then it will be in an orbit of size $n$. Thus $d_{m+n} - d_m d_n$ is divisible by $n$, i.e.,

$$d_{m+n} \equiv d_m d_n \pmod{n}.$$ 

For a prime modulus $p$, we have $d_p \equiv p - 1 \pmod{p}$, so

$$d_{m+p} \equiv (p - 1)d_m \equiv -d_m \pmod{p}.$$
The **Bell number** $B_n$ is the number of partitions of an $n$-element set.

<table>
<thead>
<tr>
<th>$n$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_n$</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>52</td>
<td>203</td>
<td>877</td>
<td>4140</td>
<td>21147</td>
</tr>
</tbody>
</table>

We will prove **Touchard’s congruence** $B_{n+p} \equiv B_{n+1} + B_n \pmod{p}$, where $p$ is a prime.
The cyclic group $C_p$ acts on partitions of $[p + m]$ by cyclically permuting $1, 2, \ldots, p$. Then $B_{n+p}$ is congruent modulo $p$ to the number of fixed partitions.
The cyclic group $C_p$ acts on partitions of $[p + m]$ by cyclically permuting $1, 2, \ldots, p$. Then $B_{n+p}$ is congruent modulo $p$ to the number of fixed partitions.

There are two kinds of fixed partitions:

1. those in which $1, 2, \ldots, p$ are all in the same block
2. those in which $1, 2, \ldots, p$ are each in singleton blocks
The cyclic group $C_p$ acts on partitions of $[p + m]$ by cyclically permuting $1, 2, \ldots, p$. Then $B_{n+p}$ is congruent modulo $p$ to the number of fixed partitions.

There are two kinds of fixed partitions:

1. those in which $1, 2, \ldots, p$ are all in the same block
2. those in which $1, 2, \ldots, p$ are each in singleton blocks

The number of partitions of type 1 is $B_{n+1}$ since we can think of $1, 2, \ldots, p$ as being replaced by a single point.
The cyclic group $C_p$ acts on partitions of $[p + m]$ by cyclically permuting $1, 2, \ldots, p$. Then $B_{n+p}$ is congruent modulo $p$ to the number of fixed partitions.

There are two kinds of fixed partitions:

1. those in which $1, 2, \ldots, p$ are all in the same block
2. those in which $1, 2, \ldots, p$ are each in singleton blocks

The number of partitions of type 1 is $B_{n+1}$ since we can think of $1, 2, \ldots, p$ as being replaced by a single point.

The number of type 2 is $B_n$. 
The cyclic group $C_p$ acts on partitions of $[p + m]$ by cyclically permuting $1, 2, \ldots, p$. Then $B_{n+p}$ is congruent modulo $p$ to the number of fixed partitions.

There are two kinds of fixed partitions:

1. those in which $1, 2, \ldots, p$ are all in the same block
2. those in which $1, 2, \ldots, p$ are each in singleton blocks

The number of partitions of type 1 is $B_{n+1}$ since we can think of $1, 2, \ldots, p$ as being replaced by a single point.

The number of type 2 is $B_n$.

So $B_{n+p} \equiv B_n + B_{n+1} \pmod{p}$. 