

Binomial coefficients modulo prime powers

by

Andrew Granville *

(School of Mathematics at the Institute for Advanced Study)

Abstract: We study the value of binomial coefficients modulo given prime powers, and prove a wide variety of results, both old and new. Our main new theorem is a generalization of Lucas' Theorem to arbitrary prime powers, which allows us to evaluate $\binom{n}{m} \pmod{p^q}$ in $O(\log^2 n + q^4 \log n \log p + q^4 p \log^3 p)$ elementary operations. We also provide three quite different proofs of Lucas' Theorem, establish various extensions of Wolstenholme's Theorem, generalize a result of Morley, and quite a bit else besides. We have collected together some of the diverse directions taken in this subject, discussing connections with cellular automata, Fermat's Last Theorem and the prime recognition problem, providing proofs where it seems appropriate.

1. Introduction

Many great mathematicians of the nineteenth century considered problems involving binomial coefficients modulo a prime power (for instance Babbage, Cauchy, Cayley, Gauss, Hensel, Hermite, Kummer, Legendre, Lucas and Stickelberger — see [Di]). They discovered a variety of elegant and surprising Theorems which are often easy to prove. In this article we shall exhibit most of these results, extend them in a variety of ways, and give some new results. We start with a discussion of some of what is known and state selected parts of our new results:

In 1852 Kummer showed that the power of prime p that divides the binomial coefficient $\binom{n}{m}$ is given by the number of 'carries' when we add m and $n - m$ in base p . In 1878 Lucas gave a method to easily determine the value of $\binom{n}{m} \pmod{p}$: Let m_0 and n_0 be the least non-negative residues of m and $n \pmod{p}$, respectively. Then

$$(1) \quad \binom{n}{m} \equiv \binom{[n/p]}{[m/p]} \binom{n_0}{m_0} \pmod{p},$$

where, as usual, $[x]$ denotes the largest integer $\leq x$, and we use the convention $\binom{r}{s} = 0$ if $r < s$. Re-writing $n = n_0 + n_1p + n_2p^2 + \dots + n_dp^d$ and $m = m_0 + m_1p + \dots + m_dp^d$ in

* The author is supported, in part, by the National Science Foundation (grant number DMS-8610730)

base p (so that $0 \leq m_i, n_i \leq p-1$ for each i), this may also be expressed as

$$\binom{n}{m} \equiv \binom{n_0}{m_0} \binom{n_1}{m_1} \cdots \binom{n_d}{m_d} \pmod{p}.$$

Also, in this notation, Kummer's Theorem states that the power of p dividing $\binom{n}{m}$ is precisely the number of indices i for which $n_i < m_i$. We will give three very different proofs of Lucas' Theorem: the first, via number theory, in section 2; the second, via cellular automata, in section 5; and the third, via the combinatorics of power series, in section 6.

If p divides $\binom{n}{m}$ then (1) follows easily from Kummer's Theorem. However, if p^k is the exact power of p dividing $\binom{n}{m}$, then we might ask for the value of $\frac{1}{p^k} \binom{n}{m} \pmod{p}$. This is given by a result discovered by each of Anton (1869), Stickelberger (1890) and Hensel (1902) (and many others since!), which shows that

$$(2) \quad \frac{1}{p^k} \binom{n}{m} \equiv (-1)^k \left(\frac{n_0!}{m_0! r_0!} \right) \left(\frac{n_1!}{m_1! r_1!} \right) \cdots \left(\frac{n_d!}{m_d! r_d!} \right) \pmod{p},$$

where $r = n - m$. Numerous authors have asked whether there is an analogous formula, modulo p^q , for arbitrary $q \geq 1$. In section 2 we show the following:

For a given integer n define $(n!)_p$ to be the product of those integers $\leq n$ that are not divisible by p .

Theorem 1. *Suppose that prime power p^q and positive integers $m = n + r$ are given. Write $n = n_0 + n_1 p + \cdots + n_d p^d$ in base p , and let N_j be the least positive residue of $[n/p^j] \pmod{p^q}$ for each $j \geq 0$ (so that $N_j = n_j + n_{j+1} p + \cdots + n_{j+q-1} p^{q-1}$): also make the corresponding definitions for m_j, M_j, r_j, R_j . Let e_j be the number of indices $i \geq j$ for which $n_i < m_i$ (that is, the number of 'carries', when adding m and r in base p , on or beyond the j th digit). Then*

$$(3) \quad \frac{1}{p^{e_0}} \binom{n}{m} \equiv (\pm 1)^{e_{q-1}} \left(\frac{(N_0!)_p}{(M_0!)_p (R_0!)_p} \right) \left(\frac{(N_1!)_p}{(M_1!)_p (R_1!)_p} \right) \cdots \left(\frac{(N_d!)_p}{(M_d!)_p (R_d!)_p} \right) \pmod{p^q},$$

where (± 1) is (-1) except if $p = 2$ and $q \geq 3$.

Taking $q = 1$ in (3) gives (2). Note that (3) may be re-written in terms of factorials, as each $(k!)_p = k! / [k/p]! p^{[k/p]}$. Davis and Webb [DW] have also generalized Lucas' Theorem to prime powers, though their result is slightly more complicated, and they have the restriction that $q \leq d + 1 - e_0$ — Although it is 112 years since Lucas' proved (1), it is only within the last twelve months that it has been generalized to higher powers, and then in two independant papers!

Theorem 1 provides a quick way to compute the value of binomial coefficients modulo arbitrary prime powers, as it is straightforward to determine each of the $n_j, N_j, e_j, \dots$ etc.,

and then we need only determine the values of $(k!)_p \pmod{p^q}$ (in (3)) with $k < p^q$ (which would take $O(\log^2 n + p^q \log^2(p^q))$ elementary operations). By showing that we actually only need to determine the values of $(k!)_p \pmod{p^q}$ for $k < qp$, we will speed this up to $O(\log^2 n + q^4 \log n \log p + q^4 p \log^3 p)$ elementary operations, in section 3.

Wilson's Theorem (which was actually discovered by Leibnitz) states that $(p-1)! \equiv -1 \pmod{p}$ for all primes p . An easy consequence of this is that $\binom{np-1}{p-1} \equiv 1 \pmod{p}$ for all integers n . In 1819 Babbage noticed that, further, $\binom{2p-1}{p-1} \equiv 1 \pmod{p^2}$ for all primes $p \geq 3$, and Wolstenholme, in 1862, that

$$(4) \quad \binom{2p-1}{p-1} \equiv 1 \pmod{p^3}$$

for all primes $p \geq 5$. In 1952 Ljunggren generalized this to $\binom{np}{mp} \equiv \binom{n}{m} \pmod{p^3}$; and Jacobsthal to

$$(5) \quad \binom{np}{mp} \bigg/ \binom{n}{m} \equiv 1 \pmod{p^q}$$

for any integers $n > m > 0$ and prime $p \geq 5$, where q is the power of p dividing $p^3 nm(n-m)$ (this exponent q can only be increased if p divides B_{p-3} , the $(p-3)$ rd Bernoulli number).

These results, as well as many other similar congruences with larger exponents q , follow easily from Proposition 5 below. For example, if prime $p \geq 7$ then

$$(6) \quad \binom{3p}{2p} \bigg/ \binom{2p}{p}^3 \equiv \binom{3}{2} \bigg/ \binom{2}{1}^3 \pmod{p^5}.$$

Ljunggren's result above may be re-written as $((x+2)p!)_p (xp!)_p / ((x+1)p!)_p^2 \equiv 1 \pmod{p^3}$ for any integer $x \geq 0$ and $p \geq 5$. Proposition 5 implies the generalization

$$(7) \quad \prod_{j=0}^{2r} ((x+j)p!)_p^{\binom{2r}{j} (-1)^{2r-j}} \equiv 1 \pmod{p^{2r+1}},$$

unless $p^r = 2$, or $2r+1 = p$ or p^2 , when the congruence holds $\pmod{p^{2r}}$.

Another generalization of Wolstenholme's congruence is given by

Theorem 2. Suppose that prime p and positive integers u and r are given. Then

$$(8) \quad (up!)_p \equiv \pm \prod_{j=1}^r (jp!)_p^{\beta_j} \pmod{p^{2r+1}},$$

except if $p^r = 2$ or $2r + 1 = p$ or p^2 , when the congruence holds $(\bmod p^{2r})$, where ‘ $\pm$ ’ can only be ‘ $-$ ’ if $p = 2$ (which is easily determined by evaluating both sides of (8) modulo 4), and the integer

$$(9) \quad \beta_j (= \beta_{r,j}(u)) := \frac{u}{j} \prod_{\substack{1 \leq i \leq r \\ i \neq j}} \left(\frac{u^2 - i^2}{j^2 - i^2} \right).$$

Note how Theorem 2 allows us to express any $(up!)_p (\bmod p^q)$ in terms of $(jp!)_p$ with $j \leq [q/2]$. In Theorem 3 below we prove a similar result for any factorial, which allows us to compute such factorials very rapidly.

In 1899 Glaisher observed that the number of odd entries in any given row of Pascal’s Triangle is a power of 2. This follows from Kummer’s Theorem by noting that $\binom{n}{m}$ is odd if and only if there are no carries when adding m and $n - m$ in base 2; in other words that the digits ‘1’ in the binary expansion of m are a subset of those in the binary expansion of n . Clearly if there are k digits ‘1’ in the binary expansion of n , then there are 2^k possible subsets of these ‘1’s, and each corresponds to a value of m — thus there are 2^k odd entries in the n th row of Pascal’s Triangle.

Larry Roberts also has an elegant (unpublished) result, depending on Kummer’s Theorem: Let z_n be the binary number whose m th digit is $\equiv \binom{n}{m} \bmod 2$; in other words, the integer formed by reading the n th row of Pascal’s Triangle, modulo 2, from left to right. Then $z_n = \sum 2^m$, where the sum is over those values of m , for which the digits ‘1’ in its binary expansion are a subset of those of n . Thus if $S_n := \{i : n_i = 1\}$ then

$$(10) \quad z_n = \sum_{I \subseteq S_n} \prod_{i \in I} 2^{2^i} = \prod_{i \in S_n} F_i$$

where $F_i := 2^{2^i} + 1$ is the i th *Fermat number*.

In section 5 we give somewhat different proofs of these results, and of Lucas’ Theorem, using cellular automata.

In a recent paper [Gr], using methods from both elementary number theory and the theory of cellular automata, we extended this result of Glaisher’s to the entries in Pascal’s triangle that are 1 $(\bmod 4)$: specifically we showed that in any given row of Pascal’s triangle, the number of entries that are congruent to 1 $(\bmod 4)$ is either 0 or a power of 2. Similarly for 3 $(\bmod 4)$. We then extended this to 1 $(\bmod 8)$, 3 $(\bmod 8)$, 5 $(\bmod 8)$ and 7 $(\bmod 8)$. The likelihood of a general result of this type begins to emerge, and to find out more the reader is encouraged to look at [Gr].

In 1876 Hermite showed that if n is odd then the sum of the binomial coefficients $\binom{n}{m}$, over those positive integers m that are divisible by $p - 1$, is divisible by p . In 1899 Glaisher

generalized this by showing that for any given prime p and integers $1 \leq j, k \leq p-1$, we have

$$(11) \quad \sum_{\substack{1 \leq m \leq n \\ m \equiv j \pmod{p-1}}} \binom{n}{m} \equiv \binom{k}{j} \pmod{p}$$

for all positive integers $n \equiv k \pmod{p-1}$. We prove this in section 6. In 1953 Carlitz generalized Hermite's Theorem to prime powers: If p^{q-1} divides n , with $q \geq 1$ and $p \geq 3$, then

$$p + (p-1) \sum_{\substack{1 \leq m \leq n-1 \\ m \equiv 0 \pmod{p-1}}} \binom{n}{m} \equiv 0 \pmod{p^q}.$$

In 1913 Fleck gave the related result that for any given prime p and integers $0 \leq j \leq p-1 < n$, we have

$$(12) \quad \sum_{m \equiv j \pmod{p}} \binom{n}{m} (-1)^m \equiv 0 \pmod{p^q},$$

where $q = [(n-1)/(p-1)]$. In 1965 Bhaskaran showed that if p is an odd prime then $p+1$ divides n if and only if

$$(13) \quad \sum_{m \equiv j \pmod{p-1}} \binom{n}{m} (-1)^{(m-j)/(p-1)} \equiv 0 \pmod{p}$$

for $j = 1, 3, 5, \dots, p-2$. We present proofs of these two results in section 7.

In 1895 Morley [Mo] showed that for any prime $p \geq 5$,

$$(14) \quad (-1)^{\frac{p-1}{2}} \binom{p-1}{(p-1)/2} \equiv 4^{p-1} \pmod{p^3}.$$

His ingenious proof, which is based on an explicit form of De Moivre's Theorem, can be modified to show that (14) holds modulo p^4 if and only if p divides B_{p-3} ; however it cannot be modified to investigate other binomial coefficients, and so we use different method for this in section 9. Our first result is that

$$(15) \quad (-1)^{\frac{(p-1)(m-1)}{2}} \binom{p-1}{[p/m]} \binom{p-1}{[2p/m]} \cdots \binom{p-1}{[(m-1)p/m]} \equiv m^p - m + 1 \pmod{p^2},$$

for any $m \geq 2$: In fact, this product is $\equiv m^{m(p-1)} \pmod{p^3}$ whenever the $p-2$ nd Bernoulli polynomial vanishes $\pmod{p}$ at $1/m, 2/m, \dots, (m-1)/m$ (which is immediate for $m = 2$).

In 1938 Emma Lehmer [Le] related the values of $\binom{p-1}{[jp/m]} \pmod{p^2}$, for $1 \leq j < m \leq 6$, to Fermat's Last Theorem for exponent p . We shall show, in section 9, that recent, as yet

unpublished, results of Skula and Cui–Xiang imply that if the first case of Fermat’s Last Theorem is false for prime exponent p (that is, there exist integers a, b, c , not divisible by p , for which $a^p + b^p = c^p$) then

$$(16) \quad \binom{p-1}{[jp/m]} \equiv (-1)^{[jp/m]} \pmod{p^2}$$

for $1 \leq j < m \leq 12$; moreover, Skula’s approach implies that the ‘12’ here may be changed to any given number after a finite amount of computation.

There are many results in the literature that relate the value of binomial coefficients of the form $\binom{n(p-1)/d}{m(p-1)/d} \pmod{p}$, for a given, fixed $d > 0$ dividing $p-1$, to representations of the prime p by certain quadratic forms (see [HW]). The first such result, due to Gauss (1828), is for $d = 4$: Write any prime $p \equiv 1 \pmod{4}$ as $p = a^2 + b^2$, and choose the sign of a so that $a \equiv 1 \pmod{4}$; then $\binom{(p-1)/2}{(p-1)/4} \equiv 2a \pmod{p}$. Recently, Beukers conjectured that

$$\binom{(p-1)/2}{(p-1)/4} \equiv \left(1 + \frac{2^{p-1} - 1}{2}\right) \left(2a - \frac{p}{2a}\right) \pmod{p^2},$$

and this was proved in [CDE]. In 1846, Jacobi showed that if we write any prime $p \equiv 1 \pmod{3}$ as $4p = A^2 + 27B^2$, where the sign of A is chosen so that $A \equiv 1 \pmod{3}$, then

$$\binom{2(p-1)/3}{(p-1)/3} \equiv -A \pmod{p};$$

and this has now been shown to be $\equiv -A + p/A \pmod{p^2}$. These congruences, modulo p^2 , have only been discovered quite recently (in [CDE]) and there are presumably many others waiting to be found.

I’d like to thank Larry Roberts for allowing me to give here his unpublished result.

2. Elementary Number Theory and the Proof of Theorem 1.

Wilson’s Theorem, which states that $(p!)_p = (p-1)! \equiv -1 \pmod{p}$, may be generalized to prime powers as follows:

Lemma 1. *For any given prime power p^q we have*

$$(p^q!)_p \equiv \delta \pmod{p^q}$$

where $\delta = \delta(p^q) = -1$ unless $p = 2, q \geq 3$ when $\delta = 1$.

To show this we modify Gauss’s proof of Wilson’s Theorem: If we pair up each m in the product with its inverse $\pmod{p^q}$ then we see that $(p^q!)_p$ is congruent, modulo p^q ,

to the product of those $m \leq p^q$ that are not distinct from their inverses (mod p^q), that is those m for which $m^2 \equiv 1 \pmod{p^q}$. It is easy to show that the only such m are 1 and $p^q - 1$ unless $p^q = 2$ (when one only has $m = 1$) or $p = 2, q \geq 3$ when one has the additional solutions $2^{q-1} - 1$ and $2^{q-1} + 1$. The result follows.

In 1808 Legendre showed that the exact power of p dividing $n!$ is

$$(17) \quad [n/p] + [n/p^2] + [n/p^3] + \dots$$

Writing n in base p as above, we define the *sum of digits function* $\sigma(n) = \sigma_p(n) := n_0 + n_1 + n_2 + \dots + n_d$. Then (17) equals

$$(18) \quad (n - \sigma_p(n))/(p - 1).$$

These are both easily proved by an induction hypothesis: If $n < p$ (that is $n = n_0$) then clearly p does not divide $n!$ and both (17) and (23) equal zero. So, given $n \geq p$, note that the set of integers $m \leq n$ that are divisible by p are precisely the set of integers pk for $k \leq [n/p]$. Thus the power of p dividing $n!$ is exactly $[n/p]$ plus the power of p dividing $[n/p]!$. (17) then follows immediately from the induction hypothesis, and (18) after noting that $n_0 = n - p[n/p] = \sigma_p(n) - \sigma_p([n/p])$.

Kummer's Theorem follows easily from (18): Let $n = m + r$ and write each of n, m and r in base p . Let $\varepsilon_j = 1$ if there is a 'carry' in the j th digit when adding m and r in base p , let $\varepsilon_j = 0$ otherwise. Clearly then $n_0 = m_0 + r_0 - p\varepsilon_0$ and $n_j = m_j + r_j + \varepsilon_{j-1} - p\varepsilon_j$ for each $j \geq 1$. Thus the power of p dividing $\binom{n}{m}$ is, by (18),

$$\frac{\sigma_p(m) + \sigma_p(r) - \sigma_p(n)}{p - 1} = \sum_{j=0}^d \frac{m_j + r_j - n_j}{p - 1} = \frac{p\varepsilon_0 + \sum_{j=0}^d (p\varepsilon_j - \varepsilon_{j-1})}{p - 1} = \sum_{j=0}^{d-1} \varepsilon_j,$$

the total number of 'carries'.

We note here that, for each $j \geq 1$, we have

$$(19) \quad [n/p^j] - [m/p^j] - [r/p^j] = \varepsilon_{j-1}.$$

This can be seen by letting n', m', r' be the least residues, in absolute value, of $n, m, r \pmod{p^j}$, respectively, so that p^j times the left side of (19), plus $n' - m' - r'$ equals $n - m - r = 0$. However,

$$n' - m' - r' = \sum_{i=0}^{j-1} (n_i - m_i - r_i)p^i = -p\varepsilon_0 + \sum_{i=1}^{j-1} (\varepsilon_{i-1} - p\varepsilon_i)p^i = -p^j\varepsilon_{j-1}.$$

and (19) follows.

The improvement, (2), of Lucas' Theorem is easily deduced from the equation

$$(-1)^{[n/p]}(n!)_p \equiv n_0! \pmod{p},$$

which was discovered by Anton, Stickelberger and then Hensel. For an arbitrary prime power p^q , this may be generalized as follows: Define N_0 to be the least non-negative residue of $n \pmod{p^q}$, and $\delta = \delta(p^q)$. Then, writing each r in the product below as $ip^q + j$, we get

$$\begin{aligned} \delta^{[n/p^q]}(n!)_p &= \delta^{[n/p^q]} \prod'_{r \leq n} r = \left(\prod_{i=0}^{[n/p^q]-1} \delta \prod'_{1 \leq j \leq p^q} (ip^q + j) \right) \left(\prod'_{1 \leq j \leq N_0} ([n/p^q]p^q + j) \right) \\ (20) \quad &\equiv (\delta(p^q!)_p)^{[n/p^q]} (N_0!)_p \equiv (N_0!)_p \pmod{p^q} \end{aligned}$$

by Lemma 1, where $\prod'$ signifies, here and henceforth, a product over integers not divisible by p .

Now, with definitions as in Theorem 1, we have, for any given $j \geq 0$,

$$[n/p^j]! / p^{[n/p^{j+1}]} [n/p^{j+1}]! = ([n/p^j]!)_p \equiv \delta^{[n/p^{j+q}]} (N_j!)_p \pmod{p^q}$$

by (20). Multiplying together this congruence for each $j \geq 0$ we get

Proposition 1. *For any integer n and prime power p^q , we have*

$$n! / p^{\sum_{j \geq 1} [n/p^j]} \equiv \delta^{\sum_{j \geq q} [n/p^j]} \prod_{j \geq 0} (N_j!)_p \pmod{p^q},$$

where δ is defined as in Lemma 1 and each N_j as in Theorem 1.

Theorem 1 then follows from dividing the equation in Proposition 1 by the corresponding ones for m and r , and then using (19) to sort out the exponents of δ and p .

3. Fast computation of binomial coefficients $\pmod{p^q}$.

In (20) above we saw how any $(n!)_p$ may be expressed, modulo p^q , in terms of values of $(k!)_p \pmod{p^q}$ with $k < p^q$: this was the key fact behind Proposition 1 and Theorem 1. In this section we prove the following result which allows us to express $(n!)_p$, modulo p^q , in terms of $(k!)_p \pmod{p^q}$ with $k < qp$: Given integers $n \geq m \geq 0$ define $\binom{n}{m}_p := (n!)_p / (m!)_p (n-m!)_p$.

Theorem 3. Suppose that prime power p^q and non-negative integers u and v are given with $p - 1 \geq v \geq 0$. Then

$$(21) \quad \binom{up+v}{v}_p \equiv \prod_{j=1}^{q-1} \binom{jp+v}{v}_p^{\alpha_j} \pmod{p^q},$$

where the integer

$$\alpha_j (= \alpha_{q,j}(u)) := \frac{u}{j} \prod_{\substack{1 \leq i \leq q-1 \\ i \neq j}} \left(\frac{u-i}{j-i} \right).$$

Now, given $(up+v)!_p$, where $p-1 \geq v \geq 0$, $u \geq 0$, first write $(up+v)!_p = v!(up!)_p \binom{up+v}{v}_p$, and then compute $(up!)_p$ using Theorem 2, and $\binom{up+v}{v}_p$ using Theorem 3: We are thus able to express $(up+v)!_p \pmod{p^q}$ in terms of $(jp!)_p$ and $(jp+v)!_p$, with $0 \leq j \leq q-1$.

Notice that any $\binom{jp+v}{v}_p \equiv 1 \pmod{p}$, so that $\binom{jp+v}{v}_p^{p^{q-1}} \equiv 1 \pmod{p^q}$: Thus, in (21), we need only consider the value of $\alpha_j(u) \pmod{p^{q-1}}$ (and similarly $\beta_j(u) \pmod{p^{2r}}$ in (8)). Therefore, in order to compute $\frac{1}{p^k} \binom{n}{m} \pmod{p^q}$ rapidly (where k is as (2)), we suggest the following algorithm:

i) Use Theorem 1 to re-express $\frac{1}{p^k} \binom{n}{m} \pmod{p^q}$ as a product of integers of the form $(a!)_p$ with $a < p^q$.

ii) Write each such a as $up+v$ with $p-1 \geq v \geq 0$, and then use Theorems 2 and 3 to write each such $(up+v)!_p$ in terms of $(b!)_p$ with $b < qp$, to powers no larger than p^{q-1} .

iii) Compute each $(b!)_p \pmod{p^q}$ with $b < qp$, and then take each of these to the required power.

An elementary analysis reveals that (i) requires $O(\log^2 n)$, (ii) requires $O(q^4 \log n \log p)$ and (iii) requires $O(q^4 p \log^3 p)$ elementary operations, so that this algorithm typically produces enormous savings over just using Theorem 1.

In order to prove Theorem 3, we need the following

Proposition 2. For any given prime power p^q , integer u and rational y , whose denominator is not divisible by p , we have

$$(22) \quad (1 - upy) \equiv \prod_{j=1}^{q-1} (1 - jpy)^{\alpha_j(u)} \pmod{p^q}.$$

Theorem 3 then follows by taking the product of the equation in Proposition 2, with $y = -1/m$, for each $m \leq v$ that is not divisible by p .

Henceforth let $p' = p$ if $p \geq 3$ and $p' = 4$ if $p = 2$. Eisenstein (1850) and Kummer (1851) introduced the p -adic logarithm and p -adic exponential functions: Given a rational

number x , define p -adic numbers

$$\log_p(1-x) := -\sum_{n \geq 1} \frac{x^n}{n} \quad \text{when } p|x \quad \text{and} \quad \exp_p(x) := \sum_{n \geq 0} \frac{x^n}{n!} \quad \text{when } p' | x.$$

Various properties of these functions are discussed in section 5 of [Wa]: For instance if p divides both x and y then $\log_p((1-x)(1-y)) = \log_p(1-x) + \log_p(1-y)$. Moreover if p' divides x then $\log_p(\exp_p(x)) = x$ and $\exp_p(\log_p(1-x)) = 1-x$; also the highest power of p dividing x is the same as that dividing $\log_p(1-x)$. These properties will be vital in our proofs of Theorems 2 and 3.

The Proof of Proposition 2: If $p^q = 2$ then the result is trivial, so assume $p^q \geq 3$. Now take the p -adic logarithm of the quotient of the two sides of (22), so that the result is equivalent to proving that $\sum_{m \geq 1} p^m y^m H_m(u)/m \equiv 0 \pmod{p^q}$, where $H_m(u) := \sum_{j=1}^{q-1} \alpha_j(u) j^m - u^m$. We shall actually prove that each individual term, $p^m y^m H_m(u)/m$, of the sum is $\equiv 0 \pmod{p^q}$. For those terms with $1 \leq m \leq q-1$ we use

Lemma 2. *Given y_0 and distinct $y_1, y_2, \dots, y_n$, we have*

$$b_1 y_1^m + b_2 y_2^m + \dots + b_n y_n^m = y_0^m \quad \text{for } m = 0, 1, \dots, n-1$$

where

$$b_j = \prod_{\substack{1 \leq i \leq n \\ i \neq j}} \left(\frac{y_0 - y_i}{y_j - y_i} \right)$$

for each j , $1 \leq j \leq n$.

This may be verified in a number of ways, for instance by inverting the relevant Vandermonde determinant.

Now, taking $n = q-1$, $y_0 = u$ and each other $y_j = j$ in Lemma 2, we find that $b_j = j \alpha_j / u$ for each $j \geq 1$, and so $H_m(u) = 0$ for $1 \leq m \leq q-1$. Note that each $\alpha_j = (-1)^{q-1-j} \binom{u}{u-j} \binom{u-j-1}{u-q}$, and so is an integer.

For those terms with $q \leq m \leq 2q-1$ we use

Lemma 3. *Suppose that integers $b_0, b_1, \dots, b_n, y_0, y_1, \dots, y_n$ are given with $\sum_{j=0}^n b_j y_j^m = 0$ for $1 \leq m \leq r$. Then $\sum_{j=0}^n b_j y_j^m$ is divisible by m , whenever $r+1 \leq m \leq 2r+1$.*

Proof: If p^q divides m then $\phi(p^q)$ divides $m - m/p$ and $r \geq m/p \geq p^{q-1} \geq q$, so that $\sum_{j=0}^n b_j y_j^m \equiv \sum_{j=0}^n b_j y_j^{m/p} = 0 \pmod{p^q}$. The result follows from the Chinese Remainder Theorem.

Thus taking $b_0 = -1$ and, otherwise, the b_j 's and y_j 's as before, we find that $H_m(u)/m$ is an integer for $q \leq m \leq 2q-1$, by Lemma 3, and so $p^m y^m H_m(u)/m \equiv 0 \pmod{p^q}$.

Finally note that the power of p dividing m is (trivially) $\leq m/p$, so that the power of p dividing p^m/m is $\geq m/2$. Thus $p^m y^m H_m(u)/m \equiv 0 \pmod{p^q}$ whenever $m \geq 2q$ as each $y^m H_m(u)$ is an integer.

4. Recognizing the primes.

Gauss (Disquisitiones Arithmeticae, 1801, art. 329) wrote:

The problem of distinguishing prime numbers from composite numbers ... is known to be one of the most important and useful in arithmetic. ... The dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated.

In 1773 Lagrange observed that Wilson's Theorem could be used to identify primes by writing it in the form

An integer $n \geq 2$ is prime if and only if n divides $(n-1)! + 1$.

In connection with the solution of Hilbert's Tenth Problem, Matijasevič (1971) constructed an integer polynomial f (in many variables) such that the set of positive values of f is exactly the set of prime numbers (see [JSWW] for an elegant construction). The construction is based on Lagrange's reformulation of Wilson's Theorem. Professor J.P. Jones has asked whether a similar criteria to identify primes can be obtained from Wolstenholme's congruence: that is, whether it is true that (4) holds if and only if p is a prime ≥ 5 . (R. McIntosh has shown that the congruence can hold (mod p) for composite $p = 29 * 937$ and even (mod p^2) for $p = 16843^2$; thus the '3' is certainly necessary.) One might also ask the same question based on the generalization (5) of Wolstenholme's Theorem, or of (6).

As far back as 1819, Babbage gave an easily proved characterization of the primes, based on a number of simultaneous congruences:

An integer $n \geq 2$ is prime if and only if $\binom{n+m}{n} \equiv 1 \pmod{n}$ for all m , $0 \leq m \leq n-1$ (notice that the range for m may be shortened to $0 \leq m \leq \sqrt{n}$).

In 1972, Mann and Shanks [MS] came up with another characterization involving a number of simultaneous congruences:

An integer $n \geq 3$ is prime if and only if m divides $\binom{m}{n-2m}$ for all m , $0 \leq 2m \leq n$ (notice that the range for m may be shortened to $n - \sqrt{n} \leq 2m \leq n$). To prove this note first that if n is prime then $\binom{m}{n-2m} = \frac{m}{n-2m} \binom{m-1}{n-2m-1}$, which is clearly divisible by m , as $(m, n-2m) = 1$. On the other hand, if even n is composite take $m = n/2$ so that m does not divide $\binom{m}{0} = 1$, and if odd n is composite and divisible by prime p take $m = (n-p)/2$.

If p^r is the highest power of p that divides m (note that $r \geq 1$) then it is easy to show that p^{r-1} is the highest power of p that divides $\binom{m}{p}$, by Kummer's Theorem.

In 1915 Fleck gave an imaginative generalization of Wilson's Theorem, that can be used to characterize primes after a certain amount of trial division:

For any integers $a \geq 1$ and n , free of prime factors $\leq a$, we have that n is prime if and only if

$$(23) \quad \binom{a}{a} \binom{a+1}{a} \binom{a+2}{a} \cdots \binom{n-1}{a} \equiv (-1)^{\binom{a+1}{2}} \binom{a}{1} \binom{a}{2} \cdots \binom{a}{a-1} \pmod{n}.$$

(Actually Fleck did not include the condition that n is free of prime factors $\leq a$ but some such condition is essential as (23) holds for the example $a = 6$, $n = 15$.)

To see this, note first that if n is composite with prime factor q , $a < q < n$, then q divides the left side of (23) (as q divides $\binom{q}{a}$), but not the right side. On the other hand suppose that $n = p$ is prime. We prove (23) for $1 \leq a \leq p-1$ by induction on a : For $a = 1$ this is just Wilson's Theorem. Thereafter, the ratio of (23) for $a+1$ to (23) for a , is $(p-1-a)!/(a+1)^{p-1-a}$ on the left side and $(-1)^{a+1}(a+1)^a/a!$ on the right side, which are congruent, modulo p , by Fermat's Theorem and Wilson's Theorem.

5. Pascal's triangle via cellular automata.

A beautiful aspect of Pascal's triangle modulo 2 is that the 'pattern' inside any triangle of 1's is similar in design to that of any subtriangle, though larger in size. If we extend Pascal's triangle to infinitely many rows, and reduce the scale of our picture in half each time that we double the number of rows, then the resulting design is called *self-similar* – that is, our picture can be reproduced by taking any subtriangle and magnifying it. Such an approach to Pascal's triangle is taken in [Wo]; and many examples of self-similarity have been investigated by Mandelbrot [Ma].

We can study the value of entries in Pascal's triangle $(\bmod p)$ by such a 'pictorial approach':

From Kummer's Theorem we know that the p th row of Pascal's triangle $(\bmod p)$ has 1's on either end with 0's all the way in-between; and from the fact that any entry of the triangle is just the sum of the two adjacent entries on the line immediately above, we form a triangle underneath each of these 1's whose entries are the same as those of Pascal's triangle $(\bmod p)$. These two triangles stay independant of one another until they meet in the $2p$ th row. Thus, in that row, we have two copies of the p th row of Pascal's triangle, side-by-side, except that the middle term, has the corner terms of our two triangles overlayed: Therefore this row has a 1 on either end, a 2 in the middle, and 0's all the way in-between.

Again, underneath each of these 1's we form a triangle whose entries are the same as those of Pascal's triangle (mod p), while underneath the 2 we form a triangle whose entries are twice that in Pascal's triangle (mod p). These three triangles meet in the $3p$ th row, which thus has 1's on either end, 3's at one-third and two-thirds of the way across and 0's everywhere else. Now underneath each of the 1's we again form a triangle whose entries are the same as those of Pascal's triangle (mod p), while underneath the 3's we form a triangle whose entries are three times that in Pascal's triangle (mod p).

Continuing this process, we see that the np th row of Pascal's triangle (mod p) is a copy of the n th row, with $p - 1$ 0's placed between consecutive entries; and that the $p - 1$ rows immediately beneath the np th row are given by forming triangles underneath each non-zero entry of the np th row (say, $\binom{np}{mp} \equiv \binom{n}{m} \pmod{p}$), that are $\binom{n}{m}$ times Pascal's triangle (mod p). Thus $\binom{np+k}{mp+j} \equiv \binom{n}{m} \binom{k}{j} \pmod{p}$, so that Lucas' Theorem may be viewed as a result about automata with p possible states !

Wolfram gave an elegant proof of Glaisher's Theorem (that the number of odd entries in a given row of Pascal's triangle is a power of 2), via the following induction hypothesis: For each $n \geq 1$, rows 2^n to $2^{n+1} - 1$ modulo 2 are given by taking two copies of rows 0 to $2^n - 1$ of Pascal's triangle, modulo 2, side-by-side, and filling the space in-between with 0's; moreover Glaisher's result holds for each of these rows. For $n = 1$ we observe this by computation. For $n \geq 2$ note that row $2^n - 1$ must be all 1's so that row 2^n has 1's on either end with 0's all the way in-between. Thus, underneath each of these 1's we obtain a triangle whose entries are the same as those of Pascal's triangle, and the triangles don't meet until after the $(2^{n+1} - 1)$ th row. Therefore the $(2^n + r)$ th row ($0 \leq r \leq 2^n - 1$) modulo 2 is just two copies of the r th row modulo 2, with some 0's in-between, and so has twice as many odd entries as the r th row; this completes the proof.

Also, as row $k = 2^n + r$ is two copies of row r , whose first entries are separated by $2^n - 1$ 0's, thus Roberts' integer

$$z_k = \left(2^{2^n} + 1\right) z_r = \prod_{i \in \{n\} \cup S_r} F_i = \prod_{i \in S_k} F_i.$$

The above approach has a further pretty consequence (see also [Lo]): If we cut Pascal's triangle modulo p into subtriangles whose boundaries have p^k entries, in the obvious way (that is, with rows 0 to $p^k - 1$ in the first such triangle, then rows p^k to $2p^k - 1$ cut into three subtriangles, two outer and one inner inverted triangle, etc. etc.), then any given subtriangle is exactly the sum of the two adjacent subtriangles, in the row of subtriangles immediately above. In other words these subtriangles obey the same addition law as Pascal's triangle itself. The behaviour of Pascal's triangle modulo higher powers of p is somewhat more complicated, but still follows certain rules which are discussed in [Gr].

Finally we mention a result of Trollope (1968): Let $s(x) := \frac{1}{x} \sum_{n \leq x} \sigma_2(n)$. A typical integer $n \leq x$ has $\log x / \log 2$ digits, half of which one expects to be 1's, so that $s(x)$ should

be approximately $\log x / 2 \log 2$. Therefore, we compare $s(x)$ with $s(x')$, when $\log x / \log 2$ and $\log x' / \log 2$ have the same fractional part, by considering the function

$$f(x) = \lim_{k \rightarrow \infty} \left\{ s(2^k x) - \frac{\log(2^k x)}{2 \log 2} \right\},$$

for each x , $1 \leq x \leq 2$. One can easily show that this limit exists and that the function $f(x)$ is continuous. However Trollope proved the surprising result that $f(x)$ is nowhere differentiable. For more on such questions see [BCM].

6. Studying binomial coefficients through their generating function.

We start this section by giving another proof of Lucas' Theorem (due to Fine (1947)), based on the obvious generating function for $\binom{n}{m}$:

Start by noting that $(1 + X)^{p^j} \equiv 1 + X^{p^j} \pmod{p}$ as each $\binom{p^j}{i}$ is divisible by p , by Kummer's Theorem, unless $i = 0$ or p^j . Therefore, writing n in base p , we have

$$\begin{aligned} \sum_{m=0}^n \binom{n}{m} X^m &= (1 + X)^n = \prod_{j=0}^d \left((1 + X)^{p^j} \right)^{n_j} \\ &\equiv \prod_{j=0}^d \left(1 + X^{p^j} \right)^{n_j} = \prod_{j=0}^d \left(\sum_{m_j=0}^{n_j} \binom{n_j}{m_j} X^{m_j p^j} \right) \\ &= \sum_{m=0}^n \left(\prod_{j=0}^d \binom{n_j}{m_j} \right) X^m \pmod{p} \end{aligned}$$

and the result follows.

We can use the same approach to try to prove the analogue of Lucas' Theorem modulo p^2 , and arbitrary prime powers, but the details become much more complicated than in the proof given in section 2. We may also generalize this method to evaluate, modulo p , the coefficients of powers of any given polynomial:

Given an integer polynomial $f(X)$ of degree d , define $f(X)^n = \sum_{m=0}^{nd} \binom{n}{m}_f X^m$, and let $\binom{n}{m}_f = 0$ if $m < 0$ or $m > nd$ (note that $\binom{n}{m} = \binom{n}{m}_f$ when $f(X) = X + 1$). Clearly $f(X)^p \equiv f(X^p) \pmod{p}$ using Fermat's Theorem, and so

$$f(X)^n \equiv f(X^p)^{[n/p]} f(X)^{n_0} = \sum_{r,t \geq 0} \binom{[n/p]}{t}_f \binom{n_0}{r}_f X^{pt+r} \pmod{p}.$$

But if $m = pt + r$ then r is of the form $m_0 + kp$ and so we obtain the following generalization of (1):

$$(24) \quad \binom{n}{m}_f \equiv \sum_{k=0}^{d-1} \binom{[n/p]}{[m/p] - k}_f \binom{n_0}{m_0 + kp}_f \pmod{p}.$$

We use a similar approach in the

Proof of (11): By induction on n : For $1 \leq n \leq p-1$ we must have $n = k$ and the only possible value of m in the sum is j , so that the result is trivial. Now assume that $n \geq p$, and write m and n in base p . Then

$$(25) \quad m_0 + m_1 + \dots + m_d = \sigma_p(m) \equiv m \equiv j \pmod{p-1}$$

for each m in the sum in (11), as $p^i \equiv 1 \pmod{p-1}$ for each i . Thus, by Lucas' Theorem, the sum in (11) is congruent to

$$\sum \binom{n_0}{m_0} \binom{n_1}{m_1} \dots \binom{n_d}{m_d} \pmod{p},$$

where the sum is over all $(d+1)$ -tuples of integers $(m_0, m_1, \dots, m_d)$ satisfying (25) and not all zero. This is exactly the sum of the coefficients of

$X^j, X^{j+p-1}, X^{j+2(p-1)}, \dots$ in $(1+X)^{n_0}(1+X)^{n_1} \dots (1+X)^{n_d} = (1+X)^{\sigma_p(n)}$, which equals

$$\sum_{\substack{1 \leq r \leq \sigma_p(n) \\ r \equiv j \pmod{p-1}}} \binom{\sigma_p(n)}{r};$$

(11) then follows from the induction hypothesis as $1 \leq \sigma_p(n) \leq n-1$ and $\sigma_p(n) \equiv n \pmod{p-1}$.

7. A little algebraic number theory.

Proof of (12): Let ζ be a primitive p th root of unity and recall that $(p) = (1 - \zeta)^{p-1}$ as ideals in $\mathbf{Q}(\zeta)$. Define f_j to be the sum on the left side of (12) for each j , so that $g_i := \sum_{0 \leq j \leq p-1} f_j \zeta^{ij} = (1 - \zeta^i)^n$ which belongs to the ideal $(1 - \zeta)^n$, for $0 \leq i \leq p-1$. Therefore $f_j = \frac{1}{p} \sum_{0 \leq i \leq p-1} g_i \zeta^{-ij}$, belongs to $(1 - \zeta)^{n-p+1}$, but as each f_j is a rational integer, it is divisible by p^q where $(p-1)q$ is the smallest multiple of $(p-1)$, which is $\geq n-p+1$, and (12) follows immediately.

Proof of (13): Let d be a quadratic non-residue (mod p) and a, b and n any positive integers. Define the sequence $\{u_n\}_{n \geq 0}$ of integers by

$$u_0 = 0, u_1 = 1 \quad \text{and} \quad u_{n+2} = 2au_{n+1} - (a^2 - db^2)u_n \quad \text{for all } n \geq 0,$$

so that, from the binomial theorem,

$$(26) \quad u_n = \frac{(a + b\sqrt{d})^n - (a - b\sqrt{d})^n}{2b\sqrt{d}}$$

$$(27) \quad \equiv \sum_{\substack{j=1 \\ j \text{ odd}}}^{p-1} \nu_j a^{n-j} b^{j-1} d^{\frac{j-1}{2}} \pmod{p},$$

as $d^{(p-1)/2} \equiv -1 \pmod{p}$, where ν_j is the sum in (13).

Now, by Kummer's Theorem,

$$u_{p+1} \equiv \binom{p+1}{1} a^p b + \binom{p+1}{p} a b^p (-1) \equiv 0 \pmod{p};$$

and so, if $p+1$ divides n then p divides u_{p+1} , which divides u_n , by (26). So by selecting $a = b = 1$ and letting d run through all quadratic non-residues $\pmod{p}$, we have $\frac{p-1}{2}$ equations in the $\frac{p-1}{2}$ unknowns ν_j . Therefore each ν_j must be divisible by p as these equations give rise to a Vandermonde matrix whose determinant is not divisible by p .

On the other hand if (13) holds for all odd j then u_n is divisible by p for any admissible choices of a, b and d , by (27). Now fix d and select a and b so that $a + b\sqrt{d}$ is a primitive root modulo p in the field $\mathbf{Q}(\sqrt{d})$. Note that $(a + b\sqrt{d})^p \equiv a - b\sqrt{d} \pmod{p}$, so that $(a + b\sqrt{d})^{p-1} \equiv (a - b\sqrt{d})/(a + b\sqrt{d}) \pmod{p}$. By (26), we see that $\{(a - b\sqrt{d})/(a + b\sqrt{d})\}^n \equiv 1 \pmod{p}$, and so $(a + b\sqrt{d})^{n(p-1)} \equiv 1 \pmod{p}$. But $a + b\sqrt{d}$ is a primitive root modulo p and so $p^2 - 1$ divides $(p-1)n$, giving that $p+1$ divides n .

8. Bernoulli numbers and polynomials.

The *Bernoulli numbers*, $\{B_n\}_{n \geq 0}$, and the *Bernoulli polynomials*, $\{B_n(t)\}_{n \geq 0}$, are defined by the power series

$$\frac{X}{e^X - 1} = \sum_{n \geq 0} B_n \frac{X^n}{n!} \quad \text{and} \quad \frac{X e^{tX}}{e^X - 1} = \sum_{n \geq 0} B_n(t) \frac{X^n}{n!},$$

so that $B_n = B_n(0)$ and $B_n(t) = \sum_{k=0}^n \binom{n}{k} B_k t^{n-k}$. Some useful facts, that follow straight from these definitions, are that each B_n is a rational number, $B_n = 0$ if n is odd and ≥ 3 , and

$$(28) \quad m^n \sum_{i=1}^{m-1} \left(B_n \left(\frac{i}{m} \right) - B_n \right) = m B_n (1 - m^n)$$

for all integers $n \geq 0$.

In 1840 Clausen and Von Staudt showed that the denominator of B_n (n even) is precisely the product of those primes p for which $p-1$ divides n ; and further that $pB_n \equiv p-1 \pmod{p}$ for each such p (actually one also has $pB_{p^2-p} \equiv p-1 \pmod{p^2}$). In 1851 Kummer showed that $B_m/m \equiv B_n/n \pmod{p^r}$ for any even integers m and n , satisfying $m \geq n \geq r+1$, $m \equiv n \pmod{\phi(p^r)}$ and $n \not\equiv 0 \pmod{p-1}$; and one can use this in showing that

$$(29) \quad (B_m - B_n) \equiv 0 \pmod{p^r}$$

whenever $m \equiv n \pmod{\phi(p^{r+1})}$ and $m \geq n \geq r+2$.

For any positive integers t and n , we have

$$(30) \quad \sum_{j=0}^{t-1} j^n = \frac{1}{n+1} (B_{n+1}(t) - B_{n+1}).$$

From this, it is easy to deduce analogues of the Von Staudt–Clausen Theorem and Kummer’s congruences for Bernoulli polynomials. Consequently, if m divides $up+v$, for given integers $1 \leq u, v < m < p$, then, by combining the identity

$$B_n \left(\frac{up+v}{m} \right) - B_n \left(\frac{v}{m} \right) = \sum_{j=1}^n \binom{n}{j} B_{n-j} \left(\frac{v}{m} \right) \left(\frac{up}{m} \right)^j$$

with (30), we obtain

$$(31) \quad \begin{aligned} \sum_{0 \leq j \leq [up/m]} j^{n-1} &= \frac{1}{n} \left\{ B_n \left(\frac{v}{m} \right) - B_n \right\} + \sum_{i=1}^n \frac{1}{n} \binom{n}{i} B_{n-i} \left(\frac{v}{m} \right) \left(\frac{up}{m} \right)^i \\ &\equiv \frac{1}{n} \left\{ B_n \left(\frac{v}{m} \right) - B_n \right\} + \frac{up}{m} B_{n-1} \left(\frac{v}{m} \right) \pmod{p^2} \end{aligned}$$

for primes $p \geq 5$ provided $n \not\equiv 2 \pmod{p-1}$.

9. Theorems of Morley and Emma Lehmer and their generalizations.

Taking $n = p-1$ in (31) gives

$$(32) \quad \begin{aligned} (-1)^{[up/m]} \binom{p-1}{[up/m]} &= \prod_{j \leq up/m} \left(1 - \frac{p}{j} \right) \equiv 1 - p \sum_{1 \leq j \leq ((up+v)/m)-1} j^{p-2} \\ &\equiv 1 - p \left\{ \frac{B_{p-1} \left(\frac{v}{m} \right) - B_{p-1}}{p-1} \right\} \pmod{p^2}, \end{aligned}$$

which implies (15) after summing over each u , $1 \leq u \leq m-1$, applying (28) and then using Fermat's Theorem and the Von Staudt–Clausen Theorem.

In 1909, Wieferich showed that if the first case of Fermat's Last Theorem is false for prime exponent p then p^2 divides $2^{p-1} - 1$; thus $\binom{p-1}{(p-1)/2} \equiv (-1)^{(p-1)/2} \pmod{p^2}$ by (15). In 1914, Frobenius gave an induction hypothesis which should allow one to extend this to p^2 divides $m^{p-1} - 1$, for each successive integer m ; however, because of the enormous amount of computation required for each step of the hypothesis, this is currently known to hold only up to $m = 89$.

In 1938, Emma Lehmer used identities like (28) to show that if the first case of Fermat's Last Theorem is false for prime exponent p then $B_{p-1}(j/m) \equiv 0 \pmod{p}$ for $1 \leq j < m \leq 6$. Recently Skula has modified Frobenius's induction hypothesis so that the m th step might also show that $B_{p-1}(v/m) - B_{p-1} \equiv 0 \pmod{p}$ for $1 \leq v \leq m-1$: (32) would then give (16) for $1 \leq j = u \leq m-1$. (Skula has done this for $m \leq 10$; Zhong Cui–Xiang has obtained the same result for $\phi(m) \leq 6$, independantly.)

The left side of (15) is

$$(33) \quad \equiv 1 - p \sum_j \frac{1}{j} + \frac{p^2}{2} \left(\sum_j \frac{1}{j} \right)^2 - \frac{p^2}{2} \sum_j \frac{1}{j^2} \pmod{p^3},$$

where $\sum_j = \sum_{u=1}^{m-1} \sum_{1 \leq j \leq [up/m]}$.

Taking $n = p^2 - p$ and $n = p - 2$ in (31), and then using a number of the well-known congruences quoted in section 8 as well as (28), we obtain

$$\sum_j \frac{1}{j} \equiv m \left(\frac{1 - m^{p^2-p}}{p^2} \right) + \frac{p}{m} \sum_{v=1}^{m-1} u B_{p-2} \left(\frac{v}{m} \right) \pmod{p^2},$$

and $\sum_j 1/j^2 \equiv 0 \pmod{p}$. Substituting these equations into (33), and using the fact that $(m^{p^2-p} - 1)/p^2 \equiv q - pq^2/2 \pmod{p^2}$ for $q = (m^{p-1} - 1)/p$, we see that the left side of (15) is

$$\equiv m^{m(p-1)} - \frac{p^2}{m} \sum_{v=1}^{m-1} u B_{p-2} \left(\frac{v}{m} \right) \pmod{p^3}.$$

10. Some useful p -adic numbers.

In section 11 we apply the results from here to binomial coefficients. We start by proving

Proposition 3. *If x is divisible by prime p then*

$$(34) \quad \log_p(1-x) = \lim_{r \rightarrow \infty} \frac{(1-x)^{p^r} - 1}{p^r},$$

where the limit is taken p -adically.

Proof: Suppose that $r \geq q \geq 1$. If $j \geq r+q$ then p^q divides the numerator of both x^j/j and $\frac{1}{p^r} \binom{p^r}{j} (-x)^j$. If $1 \leq j \leq r+q-1$ and r is sufficiently large then

$$\frac{1}{p^r} \binom{p^r}{j} = \frac{1}{j} \prod_{i=1}^{j-1} \frac{p^r - i}{i} \equiv \frac{(-1)^{j-1}}{j} \pmod{p^q},$$

so that

$$\frac{(1-x)^{p^r} - 1}{p^r} = \sum_{j \geq 1} \frac{1}{p^r} \binom{p^r}{j} (-x)^j \equiv - \sum_{j=1}^{r+q-1} \frac{x^j}{j} \equiv \log_p(1-x) \pmod{p^q}.$$

Therefore, letting $r \rightarrow \infty$ and then $q \rightarrow \infty$, we obtain (34).

For each $n \geq 1$, define

$$B_{-n} := \lim_{r \rightarrow \infty} B_{\phi(p^r) - n},$$

where the limit here is taken p -adically: Note that this limit exists and is well defined by (29); moreover $B_{-n} = 0$ for all odd n . (Using Theorem 5.11 of [Wa], one can also show that $B_{-n} = nL_p(n+1, \omega^{-n})$, where $L_p(s, \chi)$ is the p -adic L -function, and $\omega(a)$ is that p -adic $(p-1)$ st root of unity for which $\omega(a) \equiv a \pmod{p}$.)

Our main result of this section is

Proposition 4. *For any integer x we have*

$$(35) \quad \log_p(\eta_x(px)!_p) = \lambda_p x + \sum_{\substack{k \geq 3 \\ k \text{ odd}}} \frac{B_{1-k}}{1-k} \frac{(px)^k}{k},$$

where $\eta_x (= \pm 1)$ is chosen so that $\eta_x(px)!_p \equiv 1 \pmod{p'}$, and

$$\lambda_p := \lim_{r \rightarrow \infty} \frac{pB_{\phi(p^r)} - (p-1)}{\phi(p^r)}.$$

(Using Theorem 5.11 of [Wa], one can also show that $\lambda_p = \lim_{s \rightarrow 1} \left\{ -pL_p(s, \chi_0) - \frac{p-1}{1-s} \right\}$, where $\chi_0 = 1$: note that $L_p(s, \chi_0)$ has a pole at $s = 1$.)

Proof: As $j^{\phi(p')} \equiv 1 \pmod{p'}$ whenever p does not divide j , we have

$$\phi(p') \log_p(\eta_x(px)!_p) = \sum'_{1 \leq j \leq px} \log_p(j^{\phi(p')}) = \lim_{r \rightarrow \infty} \sum'_{1 \leq j \leq px} \frac{j^{p^r \phi(p')} - 1}{p^r},$$

by Proposition 3. Now, fix $q \geq 1$ and take $n = p^r \phi(p')$ in (30), so that

$$\begin{aligned} \frac{1}{n} \sum'_{1 \leq j \leq px} (j^n - 1) &\equiv \frac{1}{n} \left(\sum_{j=1}^{px-1} j^n - (p-1)x \right) \pmod{p^q} \\ &\equiv \sum_{k=1}^{q+r+1} \frac{1}{n(n+1)} \binom{n+1}{k} B_{n+1-k}(px)^k - \frac{(p-1)x}{n} \pmod{p^q} \\ &\equiv x \left(\frac{pB_n - (p-1)}{n} \right) + \sum_{\substack{k=3 \\ k \text{ odd}}}^{q+r+1} \frac{B_{1-k}}{1-k} \frac{(px)^k}{k} \pmod{p^q}, \end{aligned}$$

for all sufficiently large r , as $\frac{1}{n(n+1)} \binom{n+1}{k} \equiv \frac{(-1)^k}{k(k-1)} \pmod{p^q}$ for those k in the sum. Therefore, letting $r \rightarrow \infty$ and then $q \rightarrow \infty$, we obtain (35).

11. Congruences modulo higher powers of primes.

The main result of this section is the most difficult of the paper:

Proposition 5. *Suppose that $1 \leq r \leq k-1$ and $a_1, a_2, \dots, a_k, x_1, x_2, \dots, x_k$, with each $x_j \geq 0$, are integers such that*

$$(36) \quad a_1 x_1^m + a_2 x_2^m + \dots + a_k x_k^m = 0$$

for each odd integer m , $1 \leq m \leq 2r-1$. Given prime p , let $\eta = 1$, unless $p = 2$ and $\frac{1}{2} \sum_{j: x_j \text{ is odd}} a_j$ is odd, in which case $\eta = -1$. Then

$$(37) \quad \eta \prod_{j=1}^k (px_j)!_p^{a_j} \equiv 1 \pmod{p^{2r+1}},$$

unless (i) $p^r = 2$; or

(ii) $2r+1 = p$ and p^2 does not divide $\sum_{j=1}^k a_j x_j^p$; or

(iii) $2r+1 = p^2$ and p does not divide $\sum_{p|x_j} a_j (x_j/p)$.

In each of these three cases the congruence in (36) holds $\pmod{p^{2r}}$.

(Note that η was chosen so that the left side of (37) is $\equiv 1 \pmod{p'}$.)

From this we can easily derive the

Proof of Theorem 2: Take $y_0 = u^2$ and each other $y_j = j^2$ in Lemma 2, so that $b_j = j\beta_j/x$ and thus $\beta_1 1^m + \beta_2 2^m + \dots + \beta_n n^m = u^m$ for each odd m , $1 \leq m \leq 2n-1$. Note that $\beta_j = (-1)^{r-j} \left(\binom{u+r}{u-j} \binom{u-j-1}{u-r-1} - \binom{u+r}{u+j} \binom{u+j-1}{u-r-1} \right)$, and so is an integer. The result follows from taking $r = n$ and $k = n+1$ in Proposition 5.

Note that (6) follows from Theorem 2 with $r = 2$ and $u = 3$. We can also give the **Proof of (7)**: Take $y_j = x + j$ and $n = 2r$ in Lemma 2 so that

$$\sum_{j=0}^{2r} \binom{2r}{j} (-1)^{2r-j} (x+j)^m = 0$$

for each odd integer m , $1 \leq m \leq 2r - 1$. The result then follows from taking $k = 2r + 1$ in Proposition 5.

Now assume that (36) holds for $m = 1$. Proposition 4 then implies that

$$(38) \quad \log_p \left(\eta \prod_{j=1}^k (px_j)!_p^{a_j} \right) = \sum_{\substack{m \geq 3 \\ m \text{ odd}}} \frac{B_{1-m}}{1-m} \frac{p^m}{m} \left(\sum_{j=1}^k a_j x_j^m \right).$$

The idea will be to apply the p -adic exponential function $\exp_p$ to both sides of this equation. For example, let $k = p + 1$, $a_{p+1} = -1$, $x_{p+1} = p^q$ and $a_j = 1$, $x_j = p^{q-1}$ for $1 \leq j \leq p$. Then (38) gives that

$$(39) \quad (p^{q+1}!)_p \equiv (p^q!)_p^p \pmod{p^{3q+1}}$$

for $p \geq 5$, and modulo p^{3q-1} for $p = 2, 3$ except if $p^q = 2$.

For another example let $k = 3$, $a_1 = 1$ and $a_2 = a_3 = -1$, so that $x_1 = x_2 + x_3$; note that this implies that $x_1^m - x_2^m - x_3^m$ is divisible by $x_1 x_2 x_3$ for all odd $m \geq 1$. Jacobsthal's result (5) then follows easily from (38), as well as a version for primes 2 and 3 ((5) holds if p^q divides $18mn(m-n)$).

We now proceed to the

Proof of Proposition 5: Start by noting that the proof of Lemma 3 is easily modified to show that $\sum_{j=1}^k a_j x_j^m$ is divisible by both $m-1$ and m for all odd m , $2r+1 \leq m \leq 4r-3$, given that (36) holds for all odd $m \leq 2r-1$. Therefore, as each pB_{1-m} is a p -adic unit, (38) implies that

$$(40) \quad \log_p \left(\eta \prod_{j=1}^k (px_j)!_p^{a_j} \right) \equiv p^{2r+1} B_{-2r} * \text{an integer} \pmod{p^{2r+2}},$$

other than in those few cases where the terms for $m \geq 4r-1$ (in (38)) are relevant (namely for $r = 1$, $m = 5$, $p = 2$ and 5, for $r = 2$, $m = 7$, $p = 2, 3$ and 7, and for $r = 2$, $m = 9$, $p = 2$; however they are all $\equiv 0 \pmod{p^{2r+1}}$, except $r = 1$, $m = 5$, $p = 2$).

Evidently the right side of (40) is $\equiv 0 \pmod{p^{2r+1}}$ unless p divides the denominator of B_{-2r} , in which case $p-1$ divides $2r$, by the Von Staudt–Clausen Theorem. For such cases we will prove

Lemma 4. *In addition to the hypothesis of Proposition 5, assume that $p - 1$ divides $2r$, and let p^q be the power of p that divides $2r(2r + 1)$. Then p^{q+1} divides $\sum_{j=1}^k a_j x_j^{2r+1}$, except in cases (ii) and (iii) of Proposition 5.*

From this we deduce that p divides the ‘integer’ in (40) whenever $p - 1$ divides $2r$, except in cases (ii) and (iii); and Proposition 5 follows immediately.

It remains to give a

Proof of Lemma 4: By hypothesis $2r + 1 \equiv t \pmod{\phi(p^{q+1})}$, where $t = 1$ or p^q . If $2r + 1 > t$ then $\sum_{j=1}^k a_j x_j^{2r+1} \equiv \sum_{j=1}^k a_j x_j^t = 0 \pmod{p^{q+1}}$, and we are done. Clearly $t \geq 3$ and so we may assume that $2r + 1 = p^q$, which implies that p is odd. If $q = 1$ we get case (ii) so assume that $q \geq 2$.

Now, if p does not divide a given integer x then $z^{p-1} \equiv 1 \pmod{p^{q-1}}$ for $z = x^{p^{q-2}}$, so that

$$z^{p^2} - z^p = z^p \{(1 + (z^{p-1} - 1))^p - 1\} \equiv zp(z^{p-1} - 1) = p(z^p - z) \pmod{p^{q+1}},$$

and thus $x_j^{2r+1} \equiv (p+1)x_j^{p^{q-1}} - px_j^{p^{q-2}} \pmod{p^{q+1}}$ whenever p does not divide x_j . This implies that

$$\sum_{j=1}^k a_j x_j^{2r+1} \equiv (p+1) \sum_{j=1}^k a_j x_j^{p^{q-1}} - p \sum_{j=1}^k a_j x_j^{p^{q-2}} + p \sum_{p|x_j} a_j x_j^{p^{q-2}} \pmod{p^{q+1}} :$$

The first two sums here are 0 by (36) and the last is $\equiv 0 \pmod{p^{q+1}}$ except in case (iii).

12. Concluding remarks.

There have been numerous papers over the last few years that have been concerned with sequences of integers for which a ‘Kummer-type’ Theorem, a ‘Lucas-type’ Theorem and/or a ‘Wolstenholme-type’ Theorem holds. One nice example is the *Apéry numbers*,

$$a_n := \sum_{m=0}^n \binom{n}{m}^2 \binom{n+m}{m}^2,$$

which were introduced in Apéry’s proof of the irrationality of $\zeta(3)$. At first, a few seemingly surprising congruences were found for these numbers, but in 1982, Gessel [Ge] showed that these were all consequences of the fact that the Apéry numbers satisfy ‘Lucas-type’ and ‘Wolstenholme-type’ Theorems (that is $a_{np+m} \equiv a_n a_m \pmod{p}$ and $a_{np} \equiv a_n \pmod{p^3}$) for all $n \geq 0$, $p - 1 \leq m \leq 0$ and primes $p \geq 5$). R. McIntosh has asked whether a

non-trivial sequence of integers, satisfying a ‘Lucas-type’ Theorem, can grow slower than $a_n = 2^n$?

One can also generalize the notion of binomial coefficients, as follows, and obtain ‘Kummer-type’ and ‘Lucas-type’ Theorems: Given a sequence $A := \{a_n\}_{n \geq 0}$ of integers, define $(n!)_A := a_n a_{n-1} \dots a_1$ and $\binom{n}{m}_A := (n!)_A / (m!)_A ((n-m)!)_A$, and ask what power of a prime p divides $\binom{n}{m}_A$, and also for the value of $\binom{n}{m}_A \pmod{p}$. The first of these questions is attacked systematically in [KW]. A nice example was given by Fray [Fr], who proved ‘Kummer-type’ and ‘Lucas-type’ Theorems for the sequence of ‘ q -binomial coefficients’ (where each $a_n = q^n - 1$).

There are a number of questions that have received a lot of attention in the literature which do not concern us here. Many require straightforward manipulations of some of the results given here (for instance, how many entries of a given row of Pascal’s triangle are not divisible by p), others easy generalizations (for instance to multinomial coefficients — most results in that area follow immediately from the fact that multinomial coefficients can be expressed as a product of binomial coefficients). People have also investigated the density of entries in Pascal’s triangle divisible by any given integer n , and strong estimates of the average (and various connections therein to fractals and cellular automata). For these questions, and some others that are not covered here, the reader should look at [BCM], [Si] and [St].

Some relevant references

We have not given a complete set of references as there are far too many ! Instead the reader is encouraged to look at [Di], [Si], [St], and [Gu] and [Lv], sections A08 and B64. We have, however, given more recent references particularly to results for which we have not supplied an alternative proof, or in which the exposition is particularly elegant.

- [BCM] D.W. Boyd, J. Cook and P. Morton, *On sequences of ± 1 ’s defined by binary patterns*, Diss. Math., **283**, 60pp.
- [CDE] S. Chowla, B. Dwork and R. Evans, *On mod p^2 determination of $\binom{(p-1)/2}{(p-1)/4}$* , J. of Number Theory, **24** (1986), 188–196.
- [DW] K.S. Davis and W.A. Webb, *Lucas’ Theorem for prime powers*, Europ. J. Combinatorics, **11** (1990), 229–233.
- [Di] L.E. Dickson, *Divisibility of Factorials and Multinomial Coefficients*, Chapter XI in ‘History of the Theory of Numbers’, Vol.I, (Chelsea, New York, 1919).
- [Fr] R.D. Fray, *Congruence Properties of Ordinary and q -Binomial Coefficients*, Duke Math. J., **34** (1967), 467–480.
- [Ge] I. Gessel, *Some Congruences for the Apéry numbers*, J. of Number Theory, **14** (1982), 362–368.

- [Gr] A. Granville, *Zaphod Beeblebrox's brain and the fifty-ninth row of Pascal's Triangle*, to appear in Amer. Math. Monthly.
- [Gu] R.K. Guy, *Reviews in Number Theory*, (Amer. Math. Soc., Rhode Island, 1984).
- [HW] R.H. Hudson and K.S. Williams, *Binomial Coefficients and Jacobi Sums*, Trans. Amer. Math. Soc., **281** (1984), 431–505.
- [JSWW] J.P. Jones, D. Sato, H. Wada and D. Wiens, *Diophantine representation of the set of prime numbers*, Amer. Math. Monthly, **83** (1976), 449–464.
- [KW] D. E. Knuth and H. S. Wilf, *The power of a prime that divides a generalized binomial coefficient*, J. reine angew. Math., **396** (1989), 212–219.
- [Le] E. Lehmer, *On congruences involving Bernoulli numbers and the quotients of Fermat and Wilson*, Annals of Math., **39** (1938), 350–360.
- [Lv] W.J. Leveque, *Reviews in Number Theory*, (Amer. Math. Soc., Rhode Island, 1974).
- [Lo] C. T. Long, *Pascal's triangle modulo p* , Fib. Quart., **19** (1981), 458–463.
- [Ma] B. Mandelbrot, *The Fractal Geometry of Nature*, (Freeman, San Francisco, 1982).
- [MS] H.B. Mann and D.S. Shanks, *A necessary and sufficient condition for primality, and its source*, J. of Comb. Theory Ser. A, **13** (1972), 131–134.
- [Mo] F. Morley, *Note on the congruence $2^{4n} \equiv (-)^n(2n)!/(n!)^2$, where $2n + 1$ is a prime*, Annals of Math., **9** (1895), 168–170.
- [Si] D. Singmaster, *Divisibility of binomial and multinomial coefficients by primes and prime powers*, in 'A collection of manuscripts related to the Fibonacci sequence', (Fib. Assoc., Santa Clara, 1980), 98–113.
- [St] K.B. Stolarsky, *Power and Exponential sums of digit sums related to binomial coefficient parity*, SIAM J. Appl. Math., **32** (1977), 717–730.
- [Wa] L.C. Washington, *Introduction to Cyclotomic Fields*, (Springer-Verlag, New York, 1982).
- [Wo] S. Wolfram, *Geometry of Binomial Coefficients*, Amer. Math. Monthly, **91** (1984), 566–571.

School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA.