SOME DIVISIBILITY PROPERTIES OF BINOMIAL COEFFICIENTS
AND THE CONVERSE OF WOLSTENHOLME’S THEOREM

Kevin A. Broughan
Department of Mathematics, University of Waikato, Hamilton, New Zealand
kab@waikato.ac.nz

Florian Luca
Instituto de Matemáticas, Universidad Nacional Autonoma de México, Ap. Postal
61-3 (Xangari), C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

Igor E. Shparlinski
Department of Computing, Macquarie University, Sydney, NSW 2109, Australia
igor@ics.mq.edu.au

Received: 6/9/09, Revised: 3/16/10, Accepted: 5/11/10, Published: 9/23/10

Abstract
We show that the set of composite positive integers $n \leq x$ satisfying the congruence
\[
\binom{2n-1}{n-1} \equiv 1 \pmod{n}
\]
is of cardinality at most $x \exp\left(-\frac{1}{\sqrt{2}} + o(1)\right)\sqrt{\log x \log \log x}$
as $x \to \infty$.

1. Introduction

We consider the sequence
\[
w_n = \binom{2n-1}{n-1} = \frac{1}{2} \binom{2n}{n}, \quad n \geq 1.
\]

By the Wolstenholme theorem [18], for each prime $p \geq 5$, we have
\[
w_p \equiv 1 \pmod{p^3}
\]
(see also [2, 7, 10]). It is a long standing conjecture that the converse to this theorem
is true, namely, that $w_n \not\equiv 1 \pmod{n^3}$ holds for all composite positive integers $n$ (see,
for example, [7, 9, 16, 17]). This has been verified numerically up to $10^9$ in [16], and
is easily verified for all even composite integers. Recently, Helou and Terjanian [11]
have investigated the distribution of $w_n$ modulo prime powers for composite values
of $n$.

Here, we show that the set of composite positive integers $n$ satisfying the more
relaxed congruence
\[
w_n \equiv 1 \pmod{n}
\]
is of asymptotic density zero. More precisely, if \( W(x) \) is defined to be the number of composite positive integers \( n \leq x \) which satisfy (2), then \( \lim_{x \to \infty} W(x)/x = 0 \).

In what follows, the implied constants in the symbol ‘\( O \)’ and in the equivalent symbol ‘\( \ll \)’ are absolute. The letter \( p \) is always used to denote a prime number.

**Theorem 1.** The estimate

\[
W(x) \leq x \exp \left( -(1/\sqrt{2} + o(1))\sqrt{\log x \log \log x} \right)
\]

holds as \( x \to \infty \).

Furthermore, let \( k \rem n \) denote the remainder of \( k \) on division by \( n \). The congruence (1) in particular implies that \( \{w_p \rem p : p \geq 5\} = \{1\} \). Furthermore, by [11, Corollary 5], we also have \( \{w_{p^2} \rem p^2 : p \geq 5\} = \{1\} \). However, we show that the set

\[
\mathcal{V}(x) = \{w_n \rem n : n \leq x\}
\]

is of unbounded size.

**Theorem 2.** We have

\[
\#\mathcal{V}(x) \gg x^{1/4}.
\]

It is also interesting to study the behavior of the sequence of numbers \( \gcd(n, w_n - 1) \). Let us define

\[
\text{li} x = \int_2^x \frac{dt}{\log t}.
\]

**Theorem 3.** The estimate

\[
\sum_{n \leq x} \gcd(n, w_n - 1) = \frac{1}{2} x \text{li} (x) + O \left( x^2 \exp \left( -(1/\sqrt{2} + o(1))\sqrt{\log x \log \log x} \right) \right)
\]

holds as \( x \to \infty \).

2. Preparations
2.1. Smooth Numbers

For a positive integer \( n \) we write \( P(n) \) for the largest prime factor of \( n \). As usual, we say that \( n \) is \( y \)-smooth if \( P(n) \leq y \). Let

\[
\psi(x, y) = \#\{1 \leq n \leq x : n \text{ is } y\text{-smooth}\}.
\]

The following estimate is a substantially relaxed and simplified version of Corollary 1.3 of [12] (see also [1, 8]).
Lemma 4. For any fixed $\varepsilon > 0$, uniformly over $y \geq \log^{1+\varepsilon} x$, we have
\[
\psi(x, y) = x \exp(-(1 + o(1))u \log u) \quad \text{as} \ u \to \infty,
\]
where $u = \log x / \log y$.

2.2. Distribution of $w_m$ in Residue Classes

We need some results about the distribution of $w_m$ in residue classes modulo primes. These results are either explicitly given in [4, 5, 6], or can be obtained from those results at the cost of merely minor typographical changes. More precisely, the results are obtained in [4, 5, 6] apply to middle binomial coefficients and Catalan numbers
\[
\binom{2m}{m} \quad \text{and} \quad \frac{1}{m + 1} \binom{2m}{m}, \quad m = 1, 2, \ldots,
\]
while the ones from [6] apply to the sequence of general term
\[
2^{-2m} \binom{2m}{m}, \quad m = 1, 2, \ldots,
\]
each of which is of the same type as the sequence with general term $w_m$.

In fact, the method of [4, 5, 6] which in turn is based on the arguments from [3, 15], can be applied to estimate the number of solutions of congruences
\[
H(m) \equiv a \pmod{p}, \quad 1 \leq m \leq M,
\]
uniformly in $a \in \{1, \ldots, p - 1\}$ for essentially all nontrivial “hypergeometric sequences” $H(m)$, that is, sequences of general term having the form
\[
H(m) = f(1) \cdots f(m), \quad m = 1, 2, \ldots,
\]
where $f(X) \in \mathbb{Q}(X)$ is a nonconstant rational function. Note that the original result of [3, 15] corresponds to the choice $f(m) = m$ for which $H(m) = m!$, while here we take $f(m) = 2(2m - 1)/m$ for which $H(m) = 2w_m$.

More precisely, let $\lambda$ be an integer and define $R_p(M, \lambda)$ to be the number of solutions to the congruence
\[
w_m \equiv \lambda \pmod{p}, \quad 0 \leq m \leq M - 1. \quad (3)
\]
We have the following estimate which follows immediately from [6, Lemma 5].
**Lemma 5.** Let $p$ be an odd prime and let $M$ be a positive integer. Then the estimate
\[ R_p(M, \lambda) \ll M^{2/3} + Mp^{-1/3} \]
holds uniformly over $\lambda \in \{1, \ldots, p-1\}$.

**Proof.** For $M \leq p$, the bound
\[ R_p(M, \lambda) \ll M^{2/3} \tag{4} \]
is equivalent to [6, Lemma 5]. Indeed, the congruence (3) is equivalent to
\[ \binom{2m}{m} \equiv 2\lambda \pmod{p}, \quad 0 \leq m \leq M - 1, \tag{5} \]
which by [6, Lemma 5] has $O(M^{2/3})$ solutions. We now assume that $M > p$. Write
\[ m = \sum_{j=0}^{s} m_j p^j, \tag{6} \]
with $p$-ary digits $m_j \in \{0, \ldots, p-1\}, j = 0, \ldots, s$. Then, by *Lucas’ Theorem* (see [14, Section XXI]), we have
\[ w_m = \frac{1}{2} \binom{2m}{m} \equiv \frac{1}{2} \prod_{j=0}^{s} \binom{2m_j}{m_j} \pmod{p}. \tag{7} \]

Every $m$ with $0 \leq m < M$ can be written as $m = ph + k$ with nonnegative integers $h < M/p$ and $k < p$.

Clearly, if $w_m \not\equiv 0 \pmod{p}$, then it follows from (7) that in the representation (6) we have
\[ m_j < p/2, \quad j = 0, \ldots, s. \]

We now see that for every $m = ph + k$ with $h < M/p$ and $k < p$, the congruence (7) implies that
\[ \binom{2k}{k} \equiv \lambda_h \pmod{p} \]
with some $\lambda_h \not\equiv 0 \pmod{p}$ depending only on $h$.

Therefore, by (4), we obtain $R_p(M, \lambda) \ll p^{2/3}(M/p) \ll Mp^{-1/3}$. \qed
We remark that for $\lambda \equiv 0 \pmod{p}$, the same bound also holds but only in the range $M < p/2$, and certainly fails beyond this range.

We also note that on average over $\lambda$ we have a better estimate.

**Lemma 6.** Let $p$ be an odd prime and let $M < p$ be a positive integer. Then

$$\sum_{\lambda=0}^{p-1} R_p(M, \lambda)^2 \ll M^{3/2}.$$

The above Lemma 6 follows from the equivalence between the congruences (3) and (5) and [5, Theorem 1] taken in the special case $\ell = 1$, a result which applies to middle binomial coefficients and Catalan numbers and easily extends to the sequence of general term $w_n$ (see also [4, Theorem 2]).

For large values $M$, we have a better bound which is based on some arguments of [4].

**Lemma 7.** Let $p$ be an odd prime and let $M \geq p^7$ be a positive integer. Then the estimate

$$R_p(M, \lambda) \ll M/p$$

holds uniformly over $\lambda \in \{1, \ldots, p-1\}$.

**Proof.** Every $m$ with $0 \leq m < M$ can be written as $m = p^7h + k$, with nonnegative integers $h < M/p^7$ and $k < p^7$.

Clearly, if $w_m \not\equiv 0 \pmod{p}$, then it follows from (7) that in the representation (6) we have

$$m_j < p/2, \quad j = 0, \ldots, s.$$  

We now see that for every $m = p^7h+k$ with $h < M/p^7$ and $k < p^7$, the congruence (7) implies that

$$\binom{2k}{k} \equiv \lambda_h \pmod{p}$$

holds with some $\lambda_h \not\equiv 0 \pmod{p}$ depending only on $h$. It now follows from [4, Equation (13)], that the asymptotic

$$R_p(p^7, \lambda) = (2^{-7} + o(1))p^6$$

holds as $p \to \infty$ uniformly over $\lambda \not\equiv 0 \pmod{p}$ (see also the comment at the end of [4, Section 2]). Therefore,

$$R_p(M, \lambda) \leq (2^{-7} + o(1))p^6(M/p^7) \quad \text{as} \quad p \to \infty,$$

yielding the desired conclusion $R_p(M, \lambda) \ll M/p$. \qed
3. Proofs of the Main Results

3.1. Proof of Theorem 1

We let $x$ be a large positive real number and we fix some real parameters $y > 3$ and $z \geq 1$ depending on $x$ to be chosen later.

Let $\mathcal{N}$ be the set of composite $n \leq x$ which satisfy (2). We note that, again by Lucas’ Theorem, for any prime $p$ and positive integer $m$ we have

$$\left(\frac{2mp}{mp}\right) \equiv \left(\frac{2m}{m}\right) \pmod{p}.$$ 

Hence, if $n = mp \in \mathcal{N}$, then

$$w_m \equiv w_n \equiv 1 \pmod{p}. \tag{8}$$

Let $\mathcal{E}_1$ be the set of $y$-smooth integers $n \in \mathcal{N}$ and let $\mathcal{N}_1$ be the set of remaining integers, that is, $\mathcal{N}_1 = \mathcal{N} \setminus \mathcal{E}_1$.

By Lemma 4,

$$\#\mathcal{E}_1 \leq x \exp\left(-(1 + o(1))u \log u\right) \quad \text{as} \quad u \to \infty, \tag{9}$$

where $u = \log x / \log y$, provided that $y > (\log x)^2$, which will be the case for us.

Next, we define the set

$$\mathcal{E}_2 = \{n \in \mathcal{N}_1 : P(n) > z\}.$$

For $n \in \mathcal{E}_2$, we write $n = mp$, where $p = P(n) \geq z$ and $m \leq x/z$. We see from (8) that each $p$ which appears as $p = P(n)$ for some $n \in \mathcal{E}_2$ must divide

$$Q = \prod_{2 \leq m \leq x/z} (w_m - 1) = \exp\left(O\left((x/z)^2\right)\right).$$

Observe that $Q$ is nonzero because $m = 1$ is not allowed in the product since $n$ is not prime. Therefore such $p$ can take at most $O(\log Q) = O\left((x/z)^2\right)$ possible values. Since $m$ takes at most $x/z$ possible values, we obtain

$$\#\mathcal{E}_2 \ll (x/z)^3. \tag{10}$$

Let $\mathcal{N}_2$ be the set of remaining $n \in \mathcal{N}_1$, that is

$$\mathcal{N}_2 = \mathcal{N}_1 \setminus \mathcal{E}_2.$$
We see from (8) that
\[ \#N_2 \leq \sum_{y \leq p \leq z} R_p([x/p], 1). \]

Using Lemma 5 for \( x^{1/8} < p \leq z \) and Lemma 7 for \( p \leq x^{1/8} \) and choosing
\[ z = x^{7/8}, \]
we derive
\[
\#N_2 \ll \sum_{x^{1/8} < p \leq z} \left( |x/p| p^{-1/3} + |x/p|^2/3 \right) + \sum_{y \leq p \leq x^{1/8}} \frac{|x/p|}{p} \\
\ll x \sum_{x^{1/8} < p \leq z} p^{-4/3} + x^{2/3} \sum_{x^{1/8} < p \leq z} p^{-2/3} + x \sum_{y \leq p \leq x^{1/8}} p^{-2} \\
\ll x^{23/24} + x^{2/3} z^{1/3} + xy^{-1}.
\]

The above estimates together with the given choice for \( z \) lead to the estimate
\[ \#N_2 \ll x^{23/24} + xy^{-1}. \quad (11) \]

Collecting (9), (10) and (11), we obtain
\[ \#N \ll x \exp(-(1 + o(1))u \log u) + x^{23/24} + xy^{-1}. \]

Choosing next
\[ \log y = \sqrt{\frac{1}{2} \log x \log \log x}, \quad (12) \]

to match the first and third terms, we conclude the proof.

### 3.2. Proof of Theorem 2

Let \( x \) be large and let us fix a prime \( x^{1/2} < p \leq 2x^{1/2} \). Define \( M_p = [x/p] \). We now consider integers \( n = np \) for which we have \( w_m \equiv w_n \pmod{p} \). Therefore,
\[ \#V(x) \geq \#\{\lambda \in \{0, \ldots, p - 1\} : R_p(M, \lambda) > 0\}. \]

We see that by the Cauchy-Schwartz inequality
\[ \left( \sum_{\lambda=0}^{p-1} R_p(M, \lambda) \right)^2 \leq \#V(x) \sum_{\lambda=0}^{p-1} R_p(M, \lambda)^2. \]

Using the trivial identity
\[ \sum_{\lambda=0}^{p-1} R_p(M, \lambda) = M_p \]
and Lemma 6, we conclude the proof.
3.3. Proof of Theorem 3

We follow the same approach as in the proof of Theorem 1. In particular, we let \( x \) be large and we fix some real parameter \( y > 3 \) depending on \( x \) to be chosen later.

Let \( \mathcal{R} \) be the set of integers \( n \leq x \) which are not \( y \)-smooth and for which
\[
P(n) \mid \gcd(n, w_n - 1).
\]

We see that (8) holds with \( p = P(n) \) and \( m = n/p \). Since this property is the only one used in the proof of the upper bound on \( \#\mathcal{N} \), we obtain the same bound on \( \#\mathcal{R} \), that is
\[
\#\mathcal{R} \ll x^{23/24} + xy^{-1}.
\]

For those \( n \leq x \) which are \( y \)-smooth and for \( n \in \mathcal{R} \), we estimate \( \gcd(n, w_n - 1) \) trivially as \( x \). For all the remaining composite integers \( n \leq x \), we have
\[
\gcd(n, w_n - 1) \leq n/P(n) \leq x/y.
\]

Therefore,
\[
\sum_{\substack{n \leq x \\ n \text{ composite}}} \gcd(n, w_n - 1) \ll x\psi(x, y) + (x^{23/24} + xy^{-1})x + x^2/y.
\]

Choosing \( y \) as in (12) and recalling Lemma 4, we obtain
\[
\sum_{\substack{n \leq x \\ n \text{ composite}}} \gcd(n, w_n - 1) \leq x^2 \exp \left( - \left( \frac{1}{\sqrt{2}} + o(1) \right) \sqrt{\log x \log \log x} \right),
\] (13)
as \( x \to \infty \).

Now, by (1), we see that
\[
\sum_{\substack{p \leq x \\ p \text{ prime}}} \gcd(p, w_p - 1) = \sum_{\substack{p \leq x \\ p \text{ prime}}} p.
\]

Using the Prime Number Theorem in the form given, for example, in [13, Theorem 8.30], as well as partial summation, we easily derive that the estimate
\[
\sum_{\substack{p \leq x \\ p \text{ prime}}} p = \frac{1}{2} x \ln x + O \left( x^2 \exp \left( -C(x)^{3/5}(\log \log x)^{-1/5} \right) \right)
\]
holds with some positive constant \( C \), which combined with (13) concludes the proof.
4. Comments

It follows from [11, Corollary 5] that if $n = p^2$ for some prime $p$, then $n$ satisfies the congruence $w_n \equiv 1 \pmod{n}$. In particular, by the Prime Number Theorem, we get that $W(x) \geq (1/2+o(1))\sqrt{x}/\log x$ as $x \to \infty$. There are perhaps very few positive integers $n$ with at least two distinct prime factors satisfying this congruence. There are only two such $n \leq 10^9$, namely $n = 27173 = 29 \times 937$ and $n = 2001341 = 787 \times 2543$, and one more example beyond this range (see [16, Section 3]).

There is little doubt that the bound of Theorem 2 is not tight and, based on somewhat limited numerical tests, we expect that the estimate $#V(x) = (c + o(1))x$ holds as $x \to \infty$ with $c \approx 0.355$. Studying the distribution of the fractional parts $\{w_n/n\}$ or maybe the easier question about the fractional parts $\{w_n/P(n)\}$ is of interest as well. A natural way to treat this question is to estimate the exponential sums

$$\sum_{n \leq x} \exp\left(2\pi i k \frac{w_n}{n}\right) \quad \text{and} \quad \sum_{n \leq x} \exp\left(2\pi i k \frac{w_n}{P(n)}\right),$$

which may be of independent interest.

It follows from [4, Theorem 3], that if $p$ is large and $M_p = \lfloor p^{13/2}(\log p)^6 \rfloor$, then there are $(1 + o(1))M_p/p$ positive integers $2 \leq m \leq M_p$ such that $w_m \equiv 1 \pmod{p}$ as $p \to \infty$. Clearly, only $O(1)$ of them are powers of $p$. Taking $n = mp$ for such an $m$ which is not a power of $p$, we conclude that there are infinitely many $n$ with at least two distinct prime factors such that the inequality $\gcd(n, w_n - 1) \geq n^{2/15+o(1)}$ holds as $n \to \infty$. Further investigation of the distribution of the numbers $\gcd(n, w_n - 1)$ for composite positive integers $n$ is of ultimate interest.

Acknowledgements We thank the anonymous referees for useful comments which improved the quality of this paper. Research of F. L. was supported in part by Grant SEP-CONACyT 79685, and that of I. S. was supported in part by ARC Grant DP0881473.

References


3299–3304.


Number Theory: Lattices, Number Fields, Curves, and Cryptography, Cambridge University 


128 (2008), 475–499.

Nombres de Bordeaux, 5 (1993), 411–484.


(1878), 184–240 and 289–321.

37 (2005), 809–817.

389.