

Some Divisibility Properties of Binomial Coefficients and Wolstenholme's Conjecture

KEVIN A. BROUGHAN

Department of Mathematics

University of Waikato

Private Bag 3105, Hamilton, New Zealand

`kab@waikato.ac.nz`

FLORIAN LUCA

Instituto de Matemáticas

Universidad Nacional Autónoma de México

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`

7th January 2009

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 $x \exp\left(-\left(\frac{1}{\sqrt{2}} + o(1)\right)\sqrt{\log x \log \log x}\right)$ as $x \rightarrow \infty$.

Keywords: Wolstenholme's conjecture, binomial congruences, density zero.

AMS Subject Classification: 11A41, 11B05, 11B25, 11B83.

1 Introduction

We consider the sequence

$$w_n = \binom{2n-1}{n-1} = \frac{1}{2} \binom{2n}{n}, \quad n \geq 2.$$

where we also define $w_1 = 1$.

By the Wolstenholme theorem [18], for each prime $p \geq 5$, we have

$$w_p \equiv 1 \pmod{p^3} \tag{1}$$

(see also [2, 7, 10]). It is a long standing conjecture that the converse to this theorem is true, namely, that $w_n \not\equiv 1 \pmod{n^3}$ 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 integers satisfying the more relaxed congruence

$$w_n \equiv 1 \pmod{n} \tag{2}$$

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 $\lim_{x \rightarrow \infty} W(x)/x = 0$.

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. *We have*

$$W(x) \ll x \exp\left(-\left(1/\sqrt{2} + o(1)\right)\sqrt{\log x \log \log x}\right)$$

as $x \rightarrow \infty$.

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

$$\mathcal{V}(x) = \{w_n \bmod n : n \leq x\}$$

is of unbounded size.

Theorem 2. *We have*

$$\#\mathcal{V}(x) \gg (1 + o(1))x^{1/4}$$

as $x \rightarrow \infty$.

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

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

Theorem 3. *We have*

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

as $x \rightarrow \infty$.

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*

$$\psi(x, y) = x \exp(-(1 + o(1))u \log u) \quad \text{as } u \rightarrow \infty,$$

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 at the cost of merely minor typographical changes. More precisely, these results are obtained in [4, 5, 6] for middle binomial coefficients and Catalan numbers

$$\binom{2m}{m} \quad \text{and} \quad \frac{1}{m+1} \binom{2m}{m}, \quad m = 1, 2, \dots,$$

and in [6] for the sequence

$$2^{-2m} \binom{2m}{m}, \quad m = 1, 2, \dots,$$

each of which is of the same type as the sequence with general term w_m .

In fact, the method of [4, 5, 6] is in turn based on the arguments in [3, 15], which can be applied to estimate the number of solutions of congruences

$$H(m) \equiv a \pmod{p}, \quad 1 \leq m \leq M,$$

for essentially all nontrivial “hypergeometric sequences” $H(m)$, that is, sequences of the form

$$H(m) = f(1) \cdots f(m), \quad m = 1, 2, \dots,$$

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 λ be an integer and define $R_p(M, \lambda)$ to be the number of solutions to the congruence

$$w_m \equiv \lambda \pmod{p}, \quad 0 \leq m \leq M - 1.$$

We have the following analogue of [6, Lemma 5].

Lemma 5. *Let p be an odd prime and let M be a positive integer. Then*

$$R_p(M, \lambda) \ll M^{2/3} + Mp^{-1/3}$$

uniformly over $\lambda \in [1, \dots, p - 1]$.

Proof. For $M \leq p$, the bound

$$R_p(M, \lambda) \ll M^{2/3} \tag{3}$$

is a full analogue of [6, Lemma 5].

We now assume that $M > p$.

Write

$$m = \sum_{j=0}^s m_j p^j, \tag{4}$$

with p -ary digits $m_j \in \{0, \dots, p - 1\}$, $j = 0, \dots, s$. Then, by the *Lucas theorem* (see [14, Section XXI]), we have

$$w_m = \frac{1}{2} \binom{2m}{m} \equiv \frac{1}{2} \prod_{j=0}^s \binom{2m_j}{m_j} \pmod{p}, \tag{5}$$

where as usual, we define

$$\binom{0}{0} = 1.$$

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

Clearly, if $w_m \not\equiv 0 \pmod{p}$, then it follows from (5) that in the representation (4) we have.

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

We now see that for every $m = k + ph$ with $h < M/p$ and $k < p$, the congruence (5) 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 (3), we obtain $R_p(M, \lambda) \ll p^{2/3}(M/p) \ll Mp^{-1/3}$. \square

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 λ we have a better estimate which is a full analogue of [5, Theorem 1] (taken in the special case $\ell = 1$) which applies to middle binomial coefficients and Catalan numbers and easily extends to the sequence w_n , see also [4, Theorem 2].

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

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

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*

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

uniformly over $\lambda \in [1, \dots, p-1]$.

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

Clearly, if $w_m \not\equiv 0 \pmod{p}$, then it follows from (5) that in the representation (4) we have

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

We now see that for every $m = k + p^7 h$ with $h < M/p^7$ and $k < p^7$, the congruence (5) 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

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

uniformly over $\lambda \not\equiv 0 \pmod{p}$ (see also the comment at the end of [4, Section 2]). Therefore, $R_p(M, \lambda) \leq (2^{-7} + o(1))p^6 (M/p^7) \ll M/p$. \square

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 the *Lucas theorem*, for any prime p and an integer $m \geq 1$, we have

$$\binom{2mp}{mp} \equiv \binom{2m}{m} \pmod{p}.$$

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

$$w_m \equiv w_n \equiv 1 \pmod{p}. \tag{6}$$

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,

$$\#\mathcal{E}_1 \leq x \exp(-(1 + o(1))u \log u) \quad \text{as } u \rightarrow \infty, \tag{7}$$

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 (6) that each p which appears as $p = P(n)$ for some $n \in \mathcal{E}_2$ must divide

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

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 $O(\log Q) = O((x/z)^2)$ possible values. Since m takes at most x/z possible values, we obtain

$$\#\mathcal{E}_2 \ll (x/z)^3. \quad (8)$$

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 (6) that

$$\#\mathcal{N}_2 \leq \sum_{y \leq p \leq z} R_p(\lceil x/p \rceil, 1).$$

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

$$\begin{aligned} \#\mathcal{N}_2 &\ll \sum_{x^{1/8} < p \leq z} \left(\lfloor x/p \rfloor p^{-1/3} + \lfloor x/p \rfloor^{2/3} \right) + \sum_{y \leq p \leq x^{1/8}} \frac{\lfloor x/p \rfloor}{p} \\ &\ll x \sum_{x^{1/8} < p \leq z} p^{-4/3} + x^{2/3} \sum_{x^{1/8} < p \leq z} p^{-2/3} + x \sum_{y \leq p \leq x^{1/8}} p^{-2} \\ &\ll x^{23/24} + x^{2/3} z^{1/3} + xy^{-1}. \end{aligned}$$

This, with the given choice for z , leads to the estimate

$$\#\mathcal{N}_2 \ll x^{23/24} + xy^{-1}. \quad (9)$$

Collecting (7), (8) and (9), we obtain

$$\#\mathcal{N} \ll x \exp(-(1+o(1))u \log u) + x^{23/24} + xy^{-1}.$$

Choosing next

$$\log y = \sqrt{\frac{1}{2} \log x \log \log x}, \quad (10)$$

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

3.2 Proof of Theorem 2

Let us fix a prime $x^{1/2} < p \leq (1+o(1))x^{1/2}$ as $x \rightarrow \infty$ and define $M_p = \lfloor x/p \rfloor$. We now consider integers $n = mp$ for which we have $w_m \equiv w_n \pmod{p}$. Therefore,

$$\#\mathcal{V}(x) \geq \#\{\lambda \in \{0, \dots, p-1\} : R_p(M_p, \lambda) > 0\}.$$

We see that by the Cauchy-Schwartz inequality

$$\left(\sum_{\lambda=0}^{p-1} R_p(M_p, \lambda) \right)^2 \leq \#\{\lambda \in \{0, \dots, p-1\} : R_p(M_p, \lambda) > 0\} \sum_{\lambda=0}^{p-1} R_p(M_p, \lambda)^2.$$

Using the trivial identity

$$\sum_{\lambda=0}^{p-1} R_p(M_p, \lambda) = M_p$$

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 which are not y -smooth and for which

$$P(n) \mid \gcd(n, w_n - 1).$$

We see that (6) 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 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 (10) and recalling Lemma 4, we obtain

$$\sum_{\substack{n \leq x \\ n \text{ composite}}} \gcd(n, w_n - 1) \ll x^2 \exp\left(-\left(\frac{1}{\sqrt{2}} + o(1)\right)\sqrt{\log x \log \log x}\right). \quad (11)$$

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

$$\sum_{\substack{p \leq x \\ p \text{ prime}}} p = \frac{1}{2}x\text{li}(x) + O\left(x^2 \exp(-C(\log x)^{3/5}(\log \log x)^{-1/5})\right)$$

for some constant $C > 0$, which combined with (11) 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 $x \rightarrow \infty$. There are perhaps very few positive integers n with at least two distinct prime factors satisfying this congruence. A computation with Mathematica showed that there is only one such $n \leq 10^5$, namely $n = 27173 = 29 \times 937$.

There is little doubt that the bound of Theorem 2 is not tight and, based on somewhat limited numerical tests, we expect that $\#\mathcal{V}(x) = (c + o(1))x$ as $x \rightarrow \infty$ with $c \approx 0.355$. Studying the distribution of fractional parts $\{w_n/n\}$ or maybe the easier question about fractional parts $\{w_n/P(n)\}$ is of interest as well. A natural way to treat these 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 are 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 $w_m \equiv 1 \pmod{p}$. 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 $n \rightarrow \infty$. Further investigation of the distribution of $\gcd(n, w_n - 1)$ for composite n is of ultimate interest.

Acknowledgements

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

- [1] E. R. Canfield, P. Erdős and C. Pomerance, ‘On a problem of Oppenheim concerning “Factorisatio Numerorum”’, *J. Number Theory*, **17** (1983),

1–28.

- [2] L. E. Dickson, *History of the theory of numbers, Vol 1*, Chelsea, 1996.
- [3] M. Z. Garaev, F. Luca and I. E. Shparlinski, ‘Character sums and congruences with $n!$ ’, *Trans. Amer. Math. Soc.*, **356** (2004), 5089–5102.
- [4] M. Z. Garaev, F. Luca, and I. E. Shparlinski, ‘Catalan and Apéry numbers in residue classes’, *J. Combin. Theory*, **113** (2006), 851–865.
- [5] M. Z. Garaev, F. Luca, and I. E. Shparlinski, ‘Exponential sums with Catalan numbers’, *Indag. Math.*, **18** (2007), 23–37.
- [6] M. Z. Garaev, F. Luca, I. E. Shparlinski, and A. Winterhof, ‘On the lower bound of the linear complexity over \mathbb{F}_p of Sidelnikov sequences’, *IEEE Trans. on Inform. Theory*, **52** (2006), 3299–3304.
- [7] A. Granville, ‘Arithmetic properties of binomial coefficients. I. Binomial coefficients modulo prime powers’, *Organic Mathematics, Burnaby, 1995*, Canadian Math. Soc. Conf. Proc., Vol. 20, Amer. Math. Soc., 1997, 253–275.
- [8] A. Granville, ‘Smooth numbers: Computational number theory and beyond’, *Algorithmic Number Theory: Lattices, Number Fields, Curves, and Cryptography*, Cambridge University Press, 2008, 267–322.
- [9] R. K. Guy, *Unsolved problems in number theory*, Springer-Verlag, New York, 2004.
- [10] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford Univ. Press, Oxford, 1979.
- [11] C. Helou and G. Terjanian, ‘On Wolstenholme’s theorem and its converse’, *J. Number Theory*, **128** (2008), 475–499.
- [12] A. Hildebrand and G. Tenenbaum, ‘Integers without large prime factors’, *J. de Théorie des Nombres de Bordeaux*, **5** (1993), 411–484.
- [13] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc., Providence, RI, 2004.

- [14] E. Lucas, ‘Théorie des fonctions numériques simplement périodiques’, *Amer. J. Math.*, **1** (1878), 184–240 and 289–321.
- [15] F. Luca and I. E. Shparlinski, ‘Prime divisors of shifted factorials’, *Bull. Lond. Math. Soc.*, **37** (2005), 809–817.
- [16] R. J. McIntosh, ‘On the converse of Wolstenholme’s theorem’, *Acta. Arith.*, **71** (1995), 381–389.
- [17] P. Ribenboim, *The new book of prime number records*, 3rd ed., Springer, 1996.
- [18] J. Wolstenholme, ‘On certain properties of prime numbers’, *Quart. J. Pure Appl. Math.*, **5** (1862), 35–39.