A TOUCHARD CONGRUENCE FOR THE m-BELL SEQUENCES

KOREA SUPERINTELLIGENCE LABS

Abstract. Fix a commutative ring, an integer $m \geq 1$ and a base c. Call a sequence a an *m-binomial-shift sequence with base c* if $a_{n+m} = \sum_{k \leq n} \binom{n}{k} c^{n-k} a_k$ for all $n \geq 0$; for $c = m$ over $\mathbb{Z}$ these are exactly the sequences that move m places to the left under m applications of the binomial transform, the *m-Bell sequences* recently introduced by Popov, whose $m = 1$ case are the Bell numbers. Popov asks for a systematic study of their congruence properties. We answer this. For every prime p, every $m \geq 1$ and every base c, the polynomial $(X^p - c^{p-1}X)^m - 1$ annihilates every such sequence modulo p; equivalently

$$\sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} c^{(p-1)(m-j)}\, a_{\,n+j(p-1)+m} \;\equiv\; a_n \pmod{p}.$$

At $m = c = 1$ this is Touchard's congruence $B_{n+p} \equiv B_n + B_{n+1}$. We prove the sharper statement that the *Touchard operator* $Ta = (a_{n+p} - c^{p-1} a_{n+1})_{n \geq 0}$ maps the m-dimensional solution space to itself with $T^m = \mathrm{id}$, so that the solution space is a module over $\mathbb{F}_p[y]/(y^m - 1)$ with y acting as T; every refinement below is a statement about where a sequence sits in that module. We then show that the composite m-Bell sequence can satisfy a strictly shorter recurrence than the general theorem gives: for every prime p it satisfies $B^{(p-1)}_{n+p} \equiv B^{(p-1)}_{n+1} - B^{(p-1)}_n \pmod{p}$, of degree p rather than $p(p-1)$, and we exhibit three further such cells. Finally we determine the least period of $B^{(m)} \bmod p$ in thirteen cells, including $B^{(4)} \bmod 7$, whose least period is $48 = 7^2 - 1$ against the $4N_7 = 549028$ that the neighbouring cells suggest. All theorems below are formally verified in Lean 4.

1. Introduction

1.1. **The sequences.** The *binomial transform* of a sequence $a = (a_n)_{n \geq 0}$ is the sequence with n-th term $\sum_{k=0}^{n} \binom{n}{k} a_k$; its m-th iterate has n-th term $\sum_{k=0}^{n} \binom{n}{k} m^{n-k} a_k$. The Bell numbers B_n are famously the sequence that moves one place to the left under one application of the transform, $B_{n+1} = \sum_{k \leq n} \binom{n}{k} B_k$. Popov [1] asks which sequences move m places to the left under m applications, i.e. which satisfy

$$a_{n+m} \;=\; \sum_{k=0}^{n} \binom{n}{k} m^{\,n-k} a_k, \qquad n \geq 0, \tag{1}$$

and calls their solutions the *m-Bell sequences*. The solution space of (1) over $\mathbb{Z}$ is free of rank m: the first m values may be prescribed arbitrarily and determine everything after them. Popov singles out the *composite* m-Bell sequence $B^{(m)}$, with initial vector $(1, 1, \ldots, 1)$, and the m *primitive* sequences $B^{(m,r)}$, $0 \leq r < m$, with initial vector $\delta_{n,r}$. For $m = 1$ the composite is the Bell numbers. For $m = 2$ the three sequences are recorded in the OEIS [12] as A007472 (composite), A351143 and A351028 (primitives); [1] records that for $m > 2$ they are not in the OEIS. The

first terms of the composite sequences for $m = 1, \dots, 4$ are

$$\begin{aligned} B^{(1)}: &\quad 1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, \dots \\ B^{(2)}: &\quad 1, 1, 1, 3, 9, 29, 105, 431, 1969, \dots \\ B^{(3)}: &\quad 1, 1, 1, 1, 4, 16, 64, 259, 1084, 4834, 23635, 128377, \dots \\ B^{(4)}: &\quad 1, 1, 1, 1, 1, 5, 25, 125, 625, 3129, 15745, 80265, \dots \end{aligned}$$

1.2. **The problem.** The arithmetic of the Bell numbers is a classical subject. Touchard [2] proved that for every prime p,

$$B_{n+p} \;\equiv\; B_n + B_{n+1} \pmod{p}. \tag{2}$$

Hall's 1934 abstract [4] states the same congruence and deduces from it that the residues of B_n modulo p are periodic with a period dividing $p^p - 1$; Williams [5] showed that the period in fact divides $N_p = (p^p - 1)/(p - 1)$. Whether N_p is the *least* period is a question with a long history: it is known to be least for all $p < 126$ and for $p = 137, 149, 157, 163, 167, 173$ [8, Theorem 1.1], building on [6, 7], and open in general [9]. Gessel [3] gives a short umbral proof of (2) and, in the same place, Carlitz's strengthening to higher powers of p.

For the m-Bell sequences with $m \geq 2$ nothing of this kind was recorded. The last section of [1] closes with a list of six open problems, of which the third reads, verbatim:

> The Bell numbers satisfy the Touchard congruence $B_{p+n} \equiv B_n + B_{n+1} \pmod{p}$ for every prime p. The m-Stirling triangles show visible m-adic structure (for instance $\left[\!\left[{n \atop m} \right]\!\right]_m = m^{\,n-m}$ for $n \geq m$, by the column generating function); a systematic study of congruence properties of the m-Bell sequences is open.

(The double brackets are Popov's notation for the m-Stirling numbers of the second kind, which play no further role here.) No conjecture is offered, so the task is to produce the statement, not to test one. This paper produces it.

1.3. **Results.** It costs nothing to decouple the base of the transform from the shift, so we work with the recurrence

$$a_{n+m} \;=\; \sum_{k=0}^{n} \binom{n}{k} c^{\,n-k} a_k, \qquad n \geq 0, \tag{3}$$

over an arbitrary commutative ring, with an arbitrary base c; Popov's sequences are $c = m$ over $\mathbb{Z}$. Keeping c free costs nothing in the proof and sharpens the negative controls, since the congruence sees c only through c^{p-1}.

Our main theorem (Theorem 3.4) is that over $\mathbb{F}_p$ the polynomial $(X^p - c^{p-1}X)^m - 1$ annihilates every solution of (3), that is,

$$\sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} c^{(p-1)(m-j)}\, a_{\,n+jp+(m-j)} \;=\; a_n \qquad \text{in } \mathbb{F}_p, \tag{4}$$

for every n. At $m = c = 1$ this is exactly (2). Since $jp + (m - j) = j(p-1) + m$, the shifted index in (4) is $n + j(p-1) + m$; when $p \nmid c$ all the coefficients $c^{(p-1)(m-j)}$ equal 1, and (4) says that the m-th finite difference of a along the arithmetic progression of step $p - 1$, read from $n + m$, equals a_n.

Theorem 3.4 is the m-th power of a single operator, and the operator statement (Theorem 4.3) is strictly stronger. Write

$$Ta \;=\; \bigl(a_{n+p} - c^{p-1}a_{n+1}\bigr)_{n\ge 0}.$$

Then T maps the solution space of (3) over $\mathbb{F}_p$ into itself and satisfies $T^m = \mathrm{id}$ on it. The whole reason is that $X^p - c^{p-1}X$ is invariant under the substitution $X \mapsto X + c$ over $\mathbb{F}_p$, whereas the recurrence (3) says precisely that multiplication by X^m and that substitution have the same effect on the umbral functional of a. Consequently the m-dimensional solution space is a module over $\mathbb{F}_p[y]/(y^m - 1)$, with y acting as T; this is the frame in which the finer results below are naturally read, since each of them says where a particular sequence sits in that module.

Section 5 is about sharpness. When $p \mid m$ the recurrence (1) itself degenerates to $a_{n+m} = a_n$ (Proposition 5.4), so the annihilator of degree mp is a p-fold redundant restatement of one of degree m; a computed grid shows that outside those cells the annihilator of the whole solution space is minimal. A separate phenomenon is that a *single* sequence in the space can do better than the space. We prove that for every prime p the composite $(p-1)$-Bell sequence satisfies

$$B^{(p-1)}_{n+p} \;\equiv\; B^{(p-1)}_{n+1} - B^{(p-1)}_{n} \pmod p$$

(Theorem 5.5), a recurrence of degree p where Theorem 3.4 supplies degree $p(p-1)$, strictly longer once $p \ge 3$; in operator terms $TB^{(p-1)} = -B^{(p-1)}$, so the composite is an eigenvector of T. At $p = 3$ this says that A007472 satisfies $a_{n+3} \equiv a_{n+1} - a_n \pmod 3$. Three further such cells are proved in Theorem 5.6, among them $(p, m) = (7, 4)$, where the composite satisfies $T^2 = -1$.

Section 6 is the analogue of Hall's question. The m-Touchard congruence forces a period of $B^{(m)} \bmod p$ for every m; the least period is a separate matter, and we determine it in thirteen cells (Theorems 6.1, 6.2, Proposition 6.3 and Theorem 6.4). The values do not follow the pattern $m\,N_p$ that the first three entries of the row $p = 7$ suggest. The extreme case is $(p, m) = (7, 4)$, where the least period of $B^{(4)} \bmod 7$ is $48 = 7^2 - 1$, four orders of magnitude below $4N_7 = 549028$ — and $(7, 4)$ is exactly the cell where the composite satisfies $T^2 = -1$. The deficiency table and the period table appear to be the same phenomenon seen twice, though we do not prove that either determines the other.

1.4. **How the theorem was found, and what is classical.** Honesty about provenance costs nothing here. The congruence (4) was found experimentally, not by insight: the sequences were computed modulo p from (3) and the Berlekamp–Massey algorithm [11] was run over $\mathbb{F}_p$ on the composite and on random $\mathbb{F}_p$-combinations of the primitives, on the grid $p \in \{2, 3, 5, 7, 11, 13\}$, $m \le 6$, with 900 terms in each cell. Every connection polynomial returned was supported on the arithmetic progression $mp, mp - (p-1), \dots, m, 0$ with binomial coefficients, and read off as $(X^p - c^{p-1}X)^m - 1$. The proof of Section 3 was found afterwards, and is Gessel's umbral argument for (2) generalised along the two new axes m and c. The closure of the solution space under T was likewise first observed numerically — 23 cells tested, $p \le 11$, $m \le 5$, 120 terms of each primitive, all closed — and then proved.

The case $m = 1$ of Theorem 3.4 is not new. For $c = 1$ it is Touchard's congruence; for general c it is the same argument, and the OEIS entry A004211 — the $m = 1$, $c = 2$ solution — records it verbatim in its formula section, "Touchard's congruence

holds: for odd prime p, $a(p+k) \equiv (a(k)+a(k+1)) \pmod p$ …(adapt the proof of Theorem 10.1 in Gessel)". What we have not found recorded anywhere is the range $m \geq 2$, the operator formulation, the deficiency statements of Section 5, and the period data of Section 6.

The nearest prior generalisations of (2) move along other axes. Gessel [3, §10] gives Carlitz's congruence $(B^p - B - 1)^k B^n \equiv 0 \bmod p^{\lceil k/2 \rceil}$, a statement about higher powers of p. Serafin [9] studies B_{n-p} and the r-Bell numbers; his classification [10, Theorem 3.1] determines exactly which Sheffer sequences satisfy a congruence of the shape $P_{n+p}(x) \equiv f_1 x^p P_n(x) + a P_{n+1}(x) \pmod{p\mathbb{Z}[x]}$. That shape involves only the shifts n, $n+1$ and $n+p$, so it is the analogue of the $m = 1$ case of Theorem 3.4; no congruence it classifies has the shape $(X^p - c^{p-1}X)^m - 1$ with $m \geq 2$. (A separate reason the m-Bell sequences sit outside the neighbouring classifications of triangular arrays is that the m-Stirling coefficient $m\lfloor k/m \rfloor$ is not affine in k — a point [1] itself makes.)

1.5. **Verification.** Every theorem and proposition stated below has been formally verified in Lean 4 against *Mathlib* [13]; Section 7 says exactly what is proved by argument and what is delegated to exhaustive evaluation, and lists the finite searches by size. Three displays below — the minimality grid of §5.4, the deficiency table of Remark 5.7 and the daggered rows of the period table in §6.1 — are computed data, are labelled as such, and carry no theorem.

2. Preliminaries

Throughout, R is a commutative ring, p is a prime, and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Definition 2.1. Let $m \geq 0$ and $c \in R$. A sequence $a \colon \mathbb{N} \to R$ is an *m-binomial-shift sequence with base c* if

$$a_{n+m} \;=\; \sum_{k=0}^{n} \binom{n}{k} c^{\,n-k}\, a_k \qquad \text{for every } n \geq 0.$$

Write $\mathcal{V}_{m,c}(R)$ for the set of such sequences; it is an R-submodule of $R^{\mathbb{N}}$.

For $R = \mathbb{Z}$ and $c = m$ this is (1), and $\mathcal{V}_{m,m}(\mathbb{Z})$ is the space of m-Bell sequences. If $\varphi \colon R \to S$ is a ring homomorphism and $a \in \mathcal{V}_{m,c}(R)$ then $\varphi \circ a \in \mathcal{V}_{m,\varphi(c)}(S)$; this is how an integer sequence is reduced modulo p, and it is the only compatibility we need.

The following proposition makes "the" composite and "the" primitive sequences well defined, and shows that the hypothesis of every theorem below is inhabited at every order and base.

Proposition 2.2 (Existence and rigidity)**.** *Let $m \geq 1$ and $c \in R$. For every $(v_0, \dots, v_{m-1})$ in R^m there is a sequence a in $\mathcal{V}_{m,c}(R)$ with $a_i = v_i$ for all $i < m$, and there is only one: two members of $\mathcal{V}_{m,c}(R)$ which agree at every index below m are equal.*

Proof. Existence is the evident construction: extend the initial segment one term at a time by the right-hand side of Definition 2.1, which refers only to terms already defined. Uniqueness is strong induction on n: for $n < m$ the values agree by hypothesis, and for $n \geq m$ the recurrence expresses a_n in terms of $a_0, \dots, a_{n-m}$. □

Definition 2.3. For $m \geq 1$, the *composite m-Bell sequence* $B^{(m)} \in \mathcal{V}_{m,m}(\mathbb{Z})$ is the one with $B_i^{(m)} = 1$ for $i < m$, and for $0 \leq r < m$ the *primitive* $B^{(m,r)} \in \mathcal{V}_{m,m}(\mathbb{Z})$ is the one with $B_i^{(m,r)} = \delta_{i,r}$ for $i < m$.

Proposition 2.4. *$B^{(1)}$ is the sequence of Bell numbers. The first terms of $B^{(1)}, B^{(2)}, B^{(3)}, B^{(4)}$ are as printed in §1, and*

$$\begin{aligned} B^{(2,0)} &: \quad 1,\, 0,\, 1,\, 2,\, 5,\, 16,\, 61,\, 258,\, 1177, \ldots \\ B^{(2,1)} &: \quad 0,\, 1,\, 0,\, 1,\, 4,\, 13,\, 44,\, 173,\, 792,\, 4009, \ldots \end{aligned}$$

These agree with the values printed in [1] *and with the OEIS entries A000110, A007472, A351143 and A351028.*

Proof. The identification $B^{(1)} = B$ is Proposition 2.2 applied to the two sequences $B^{(1)}$ and $(B_n)_n$, both of which lie in $\mathcal{V}_{1,1}(\mathbb{Z})$ and both of which begin with 1. The listed values are the recurrence evaluated; they were computed independently before the formal development and re-evaluated inside it. □

2.1. **The umbral functional.** The proofs in Section 3 use one device. For $a\colon \mathbb{N} \to R$ let

$$L_a\colon R[X] \longrightarrow R, \qquad L_a\Big(\sum_i b_i X^i\Big) = \sum_i b_i\, a_i,$$

the R-linear functional reading off the coefficients of a polynomial against the sequence; so $L_a(X^n) = a_n$. Everything rests on the following restatement of Definition 2.1, in which $f(X+c)$ means the composite $f \circ (X+c)$.

Lemma 2.5. *Let $m \geq 0$ and $c \in R$. A sequence a lies in $\mathcal{V}_{m,c}(R)$ if and only if*

$$L_a\big(X^m f\big) \;=\; L_a\big(f(X+c)\big) \qquad \text{for every } f \in R[X].$$

Proof. Both sides are additive in f, so it suffices to take $f = bX^n$. Then $X^m f = bX^{n+m}$ has L_a-value $b\,a_{n+m}$, while $f(X+c) = b\sum_{k\leq n}\binom{n}{k}c^{n-k}X^k$ has L_a-value $b\sum_{k\leq n}\binom{n}{k}c^{n-k}a_k$. These agree for all b and n exactly when $a \in \mathcal{V}_{m,c}(R)$. □

3. The m-Touchard congruence

Lemma 3.1 (Telescoping)**.** *Let $a \in \mathcal{V}_{m,c}(R)$. Then for every $N \geq 0$ and every $f \in R[X]$,*

$$L_a\Big(\prod_{k=0}^{N-1}(X-kc)^m \cdot f\Big) \;=\; L_a\big(f(X+Nc)\big).$$

Proof. Induction on N; the case $N = 0$ is trivial. Assume the identity for N and apply it to the polynomial $(X-Nc)^m f$ in place of f:

$$L_a\Big(\prod_{k=0}^{N}(X-kc)^m \cdot f\Big) = L_a\Big(\big((X-Nc)^m f\big)(X+Nc)\Big) = L_a\big(X^m \cdot f(X+Nc)\big),$$

because $(X+Nc)-Nc = X$. By Lemma 2.5 the last expression equals $L_a\big(f(X+Nc+c)\big) = L_a\big(f(X+(N+1)c)\big)$. □

Lemma 3.2. *For every prime p and every $c \in \mathbb{F}_p$,*

$$\prod_{k=0}^{p-1}\big(X-kc\big) \;=\; X^p - c^{p-1}X \qquad \text{in } \mathbb{F}_p[X].$$

Proof. If $c = 0$ both sides are X^p, since $c^{p-1} = 0$ for $p \geq 2$. If $c \neq 0$ then $k \mapsto kc$ is a bijection of $\mathbb{F}_p$, so the left-hand side is $\prod_{x \in \mathbb{F}_p}(X - x) = X^p - X$, the monic polynomial of degree p whose roots are exactly the elements of $\mathbb{F}_p$, each simple; and $c^{p-1} = 1$ by Fermat's little theorem. □

Proposition 3.3 (Annihilation)**.** *Let p be prime, $c \in \mathbb{F}_p$ and $a \in \mathcal{V}_{m,c}(\mathbb{F}_p)$. Then for every $f \in \mathbb{F}_p[X]$,*

$$L_a\big((X^p - c^{p-1}X)^m \cdot f\big) \;=\; L_a(f).$$

Proof. By Lemma 3.2, $(X^p - c^{p-1}X)^m = \prod_{k<p}(X - kc)^m$. By Lemma 3.1 with $N = p$, applying L_a to that product times f gives $L_a\big(f(X + pc)\big)$, and $pc = 0$ in $\mathbb{F}_p$. □

Theorem 3.4 (The m-Touchard congruence)**.** *Let p be a prime, $m \geq 1$, $c \in \mathbb{F}_p$, and let a be an m-binomial-shift sequence with base c over $\mathbb{F}_p$. Then for every $n \geq 0$,*

$$\sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} c^{(p-1)(m-j)}\, a_{\,n+jp+(m-j)} \;=\; a_n.$$

Equivalently, $(X^p - c^{p-1}X)^m - 1$ annihilates a over $\mathbb{F}_p$.

Proof. Take $f = X^n$ in Proposition 3.3, whose right-hand side is then a_n. For the left-hand side expand by the binomial theorem:

$$(X^p - c^{p-1}X)^m X^n = \sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} c^{(p-1)(m-j)}\, X^{\,pj+(m-j)+n},$$

and apply L_a term by term, using $L_a(bX^i) = b\, a_i$. □

Written out at small m, with $p \nmid c$ so that $c^{p-1} = 1$, Theorem 3.4 reads

$$\begin{aligned}
m = 1: &\quad a_{n+p} = a_n + a_{n+1},\\
m = 2: &\quad a_{n+2p} - 2a_{n+p+1} + a_{n+2} = a_n,\\
m = 3: &\quad a_{n+3p} - 3a_{n+2p+1} + 3a_{n+p+2} - a_{n+3} = a_n,
\end{aligned}$$

and so on with the coefficients of $(1 - 1)^m$ in the numerators.

Remark 3.5. Since $jp + (m - j) = j(p - 1) + m$ for $0 \leq j \leq m$, the indices occurring are $n + m$, $n + m + (p - 1)$, …, $n + m + m(p - 1)$: an arithmetic progression of step $p - 1$ starting at $n + m$. If $p \nmid c$ then $c^{p-1} = 1$ and the congruence reads $\sum_j (-1)^{m-j}\binom{m}{j} a_{\,n+j(p-1)+m} = a_n$, the m-th finite difference along that progression. If $p \mid c$ then $c^{p-1} = 0$ and only the term $j = m$ survives, so the congruence degenerates to $a_{n+mp} = a_n$; see Proposition 5.4 for what is really happening in that case.

The passage to integer sequences is immediate and is the form in which [1] poses the question.

Theorem 3.6 (Arithmetic form)**.** *Let p be a prime, $m \geq 1$, $c \in \mathbb{Z}$, and let $a\colon \mathbb{N} \to \mathbb{Z}$ satisfy $a_{n+m} = \sum_{k \leq n} \binom{n}{k} c^{n-k} a_k$ for all n. Then for every n,*

$$p \;\Big|\; \sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} c^{(p-1)(m-j)}\, a_{\,n+jp+(m-j)} \;-\; a_n.$$

Proof. Reduce a modulo p — the reduced sequence lies in $\mathcal{V}_{m,\bar{c}}(\mathbb{F}_p)$ — and apply Theorem 3.4. □

Corollary 3.7 (Touchard's congruence, recovered)**.** *Let p be a prime. Every $a \in \mathcal{V}_{1,1}(\mathbb{F}_p)$ satisfies $a_{n+p} = a_n + a_{n+1}$ for all n; in particular the Bell numbers satisfy $B_{n+p} \equiv B_n + B_{n+1} \pmod p$.*

Proof. Theorem 3.4 at $m = c = 1$ has two terms, $j = 0$ giving $-a_{n+1}$ and $j = 1$ giving a_{n+p}. The Bell numbers lie in $\mathcal{V}_{1,1}(R)$ for any R by Proposition 2.4. □

Corollary 3.7 is Touchard's theorem [2] and is stated here for one reason: it is the strongest available correctness check on Theorem 3.4. A mis-stated exponent, sign, binomial coefficient or shift index anywhere in the general statement would break it. In the formal development it is proved for *Mathlib*'s own definition of the Bell numbers, so that the check is against an independently written definition rather than against ours.

Corollary 3.8. *For every prime p, every $m \geq 1$, every $r < m$ and every $n \geq 0$,*

$$\sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} m^{(p-1)(m-j)}\, B^{(m)}_{n+jp+(m-j)} \;\equiv\; B^{(m)}_n \pmod p,$$

and the same congruence holds with $B^{(m)}$ replaced by $B^{(m,r)}$.

Proof. Theorem 3.6 with $c = m$, applied to the sequences of Definition 2.3; these are m-binomial-shift sequences with base m by Proposition 2.2. □

Corollary 3.9 ($m = 2$, explicitly)**.** *Let p be an odd prime and let $a = B^{(2)}$ be OEIS A007472. Then for every $n \geq 0$,*

$$a_{n+2p} - 2a_{n+p+1} + a_{n+2} \;\equiv\; a_n \pmod p.$$

Proof. Corollary 3.8 at $m = 2$, $c = 2$, with $2^{p-1} = 1$ because p is odd. □

The OEIS entry for A007472 carries no congruence and no modular comment, and we have not found the congruence of Corollary 3.9 recorded elsewhere. Before it was proved, it was checked against data this development did not produce: the OEIS *b*-file for A007472 (577 terms) yields 14 094 instances of the congruence over the 30 primes up to 113, all of which hold.

4. The Touchard operator

Theorem 3.4 is the m-th power of a single operator, and it is worth isolating that operator, because the refinements of Sections 5 and 6 are all statements about it.

Definition 4.1. For a prime p and $c \in \mathbb{F}_p$, the *Touchard operator* is the $\mathbb{F}_p$-linear map

$$T\colon \mathbb{F}_p^{\mathbb{N}} \to \mathbb{F}_p^{\mathbb{N}}, \qquad (Ta)_n \;=\; a_{n+p} - c^{p-1} a_{n+1}.$$

Equivalently, in the language of §2, $L_{Ta}(f) = L_a\big((X^p - c^{p-1}X)f\big)$ for every f: the operator T on sequences is multiplication by $y := X^p - c^{p-1}X$ on the polynomial side. Corollary 3.7 says exactly that T fixes the Bell numbers modulo p.

Lemma 4.2. *For every prime p and every $c \in \mathbb{F}_p$,*

$$\big(X^p - c^{p-1}X\big) \circ (X + c) \;=\; X^p - c^{p-1}X \qquad \text{in } \mathbb{F}_p[X].$$

Proof. $(X+c)^p = X^p + c^p$ by the Frobenius endomorphism, and $c^{p-1}(X+c) = c^{p-1}X + c^p$. Subtracting, the two c^p cancel. □

Theorem 4.3. *Let p be a prime, $m \geq 1$ and $c \in \mathbb{F}_p$. Then T maps $\mathcal{V}_{m,c}(\mathbb{F}_p)$ into itself, and $T^m = \mathrm{id}$ on $\mathcal{V}_{m,c}(\mathbb{F}_p)$.*

Proof. Let $a \in \mathcal{V}_{m,c}(\mathbb{F}_p)$. By Lemma 2.5 it suffices, for closure, to check $L_{Ta}(X^m f) = L_{Ta}(f(X+c))$ for every f. The left-hand side is $L_a\big(y\,X^m f\big) = L_a\big(X^m\,(yf)\big)$, which by Lemma 2.5 for a equals $L_a\big((yf)(X+c)\big)$. By Lemma 4.2,

$$(yf)(X+c) \;=\; y(X+c)\cdot f(X+c) \;=\; y\cdot f(X+c),$$

so the left-hand side equals $L_a\big(y\,f(X+c)\big) = L_{Ta}\big(f(X+c)\big)$.

For the second claim, iterating the identity $L_{Ta}(f) = L_a(yf)$ gives $L_{T^k a}(f) = L_a(y^k f)$ for every k; at $k = m$ and $f = X^n$, Proposition 3.3 gives $L_a(y^m X^n) = L_a(X^n) = a_n$, i.e. $(T^m a)_n = a_n$. □

Remark 4.4. Theorem 4.3 makes $\mathcal{V}_{m,c}(\mathbb{F}_p)$, which is m-dimensional over $\mathbb{F}_p$ by Proposition 2.2, a module over $\mathbb{F}_p[y]/(y^m-1)$, with y acting as T. Theorem 3.4 is the statement that $y^m - 1$ kills the module. Every refinement in Section 5 is the statement that some *proper* divisor of $y^m - 1$ kills a particular element: the module has torsion, and locating it is what "deficiency" means below.

5. Sharpness

5.1. **From a window to every index.** The congruence has one structural consequence that makes finite verification legitimate.

Proposition 5.1. *Let p be a prime, $m \geq 1$, $c \in \mathbb{F}_p$ and $a \in \mathcal{V}_{m,c}(\mathbb{F}_p)$. Then*

$$a_{n+mp} \;=\; a_n + \sum_{j=0}^{m-1} u_j\, a_{\,n+(jp+(m-j))}, \qquad u_j = -(-1)^{m-j}\binom{m}{j} c^{(p-1)(m-j)},$$

for every n, and every shift index occurring on the right satisfies $jp+(m-j) < mp$. In particular a obeys a monic linear recurrence of order mp.

Proof. Split off the term $j = m$ of Theorem 3.4, which is a_{n+mp}, and move the rest to the right. For the index bound, $jp + (m-j) < mp$ because $(m-j) \leq (m-j)p$ with equality only when $j = m$, which is excluded. □

Lemma 5.2 (Window bridge)**.** *Let $d \geq 1$ and let $a, b\colon \mathbb{N} \to R$ both satisfy $x_{n+d} = x_n + \sum_{j<e} u_j x_{n+t_j}$ with the same coefficients u_j and the same shifts $t_j < d$. If $a_i = b_i$ for all $i < d$ then $a = b$. Consequently, if a obeys such a recurrence and a linear combination $\sum_{i<f} v_i\, a_{n+i}$ vanishes for every $n < d$, then it vanishes for every n.*

Proof. The first claim is strong induction on n, using the recurrence to reduce an index $n \geq d$ to indices $n-d+t_j < n$ and $n-d < n$. For the second, note that $n \mapsto \sum_{i<f} v_i a_{n+i}$ obeys the same recurrence as a does, by linearity and reindexing, and that the zero sequence does too; apply the first claim to those two sequences. □

Proposition 5.1 and Lemma 5.2 together are what turns each of the finite computations below into a statement about all n: a check on the window $0 \leq n < mp$ *is* a check for every n.

5.2. **Every ingredient is load-bearing.** Before sharpening the congruence we record that it cannot be weakened.

Proposition 5.3. *Each of the following fails, in each case at $n = 0$ and by an explicit integer not divisible by the modulus.*

1. *The exponent cannot be lowered from m to $m-1$: the conclusion of Theorem 3.6 with m replaced by $m-1$ is false for $B^{(2)}$ at $p = 3$ and for $B^{(3)}$ at $p = 7$.*
2. *Primality of the modulus is needed: for $q = 4$ the congruence $B_{n+4} \equiv B_n + B_{n+1}$ fails, since $B_4 = 15$ and $B_0 + B_1 = 2$.*
3. *The binomial coefficient $\binom{m}{j}$ cannot be dropped: doing so breaks the congruence at $(p, m) = (7, 3)$.*
4. *The shift $+(m-j)$ cannot be dropped, i.e. replacing $a_{n+jp+(m-j)}$ by a_{n+jp} breaks the congruence at $(p, m) = (5, 2)$.*
5. *The factor $c^{(p-1)(m-j)}$ cannot be dropped: replacing it by 1 breaks the congruence at $p = m = 3$.*
6. *The stronger conclusion $a_{n+mp} \equiv a_n$ is false: it fails for $B^{(3)}$ at $p = 7$.*

Proof. Each item is the evaluation of one integer combination of finitely many terms of an explicit sequence and one division with remainder. The perturbation in (5) has to be placed at a cell with $p \mid c$: when $p \nmid c$ one has $c^{p-1} = 1$, so the factor is invisible and no counterexample exists. This is a fact about the theorem, not a gap in the checks: the congruence sees the base only through $c^{p-1} \in \{0, 1\}$. □

5.3. **The degenerate cells $p \mid m$.**

Proposition 5.4. *Let $a \in \mathcal{V}_{m,0}(R)$, i.e. let the base be 0. Then $a_{n+m} = a_n$ for every n. In particular, if $p \mid m$ then every m-Bell sequence reduced modulo p is periodic with period m, and the order-mp annihilator of Theorem 3.4 is a p-fold redundant restatement of the order-m one.*

Proof. In $\sum_{k \le n} \binom{n}{k} 0^{n-k} a_k$ every term with $k < n$ carries $0^{n-k} = 0$, so the sum is a_n. For an m-Bell sequence modulo p the base is $m \bmod p$, which is 0 exactly when $p \mid m$. □

5.4. **The minimality grid.** This subsection reports computation, not a theorem. For a cell (p, m) let $d(p, m)$ be the degree of the minimal polynomial annihilating the whole space $\mathcal{V}_{m,m}(\mathbb{F}_p)$, computed as the least common multiple of the Berlekamp–Massey connection polynomials of the m primitives. Theorem 3.4 says $d(p, m) \le mp$, and d was computed for all $p \in \{2, 3, 5, 7, 11, 13\}$ and $m \le 7$ with $mp \le 60$:

$d(p,m)$	$m=1$	2	3	4	5	6	7
$p=2$	2	**2**	6	**4**	10	**6**	14
$p=3$	3	6	**3**	12	15	**6**	21
$p=5$	5	10	15	20	**5**	30	35
$p=7$	7	14	21	28	35	42	**7**
$p=11$	11	22	33	44	55	—	—
$p=13$	13	26	39	52	—	—	—

Bold entries are those with $d(p, m) < mp$, and they are exactly the cells with $p \mid m$, where Proposition 5.4 explains them: $d(p, m) = m$ there. In every other computed cell the annihilator of Theorem 3.4 is minimal for the whole space. We emphasise that minimality is asserted here only as computed data: proving $d(p, m) = mp$

in a cell requires certifying the linear independence of $a, Sa, \dots, S^{mp-1}a$ for a suitable a, which is not part of the formal development. The half that *is* proved is the non-minimality half, which has witnesses: Proposition 5.4 and the theorems below.

5.5. **Deficiency of the composite.** A single sequence can satisfy a shorter recurrence than Theorem 3.4 supplies for its space — and, on the computed evidence of §5.4, shorter than the space itself admits. In the language of Remark 4.4, the sequence is killed by a proper divisor of $y^m - 1$ in $\mathbb{F}_p[y]/(y^m - 1)$. The cleanest instance holds for every prime at once.

Theorem 5.5. *Let p be a prime. Then for every $n \geq 0$,*

$$B^{(p-1)}_{n+p} \;\equiv\; B^{(p-1)}_{n+1} - B^{(p-1)}_{n} \pmod{p},$$

a linear recurrence of degree p, where Theorem 3.4 supplies degree $p(p-1)$ — strictly longer as soon as $p \geq 3$. Equivalently $TB^{(p-1)} = -B^{(p-1)}$ over $\mathbb{F}_p$: the composite is an eigenvector of the Touchard operator, killed by $y + 1$.

Proof. Put $m = p - 1$ and work over $\mathbb{F}_p$, where the base is $m \equiv -1$. Write a for $B^{(m)}$ reduced modulo p, so $a_i = 1$ for $i < m$. Also $a_m = a_0 = 1$, by the recurrence at $n = 0$. For $1 \leq k \leq m$ the recurrence at $n = k$ gives

$$a_{m+k} \;=\; \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} a_j \;=\; \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} \;=\; (1-1)^k \;=\; 0,$$

using $a_j = 1$ for $j \leq m$. Hence for $i < m$,

$$(Ta)_i = a_{i+p} - (-1)^{p-1} a_{i+1} = a_{m+(i+1)} - a_{i+1} = 0 - 1 = -a_i,$$

where $(-1)^{p-1} = 1$ by Fermat. Now Ta lies in $\mathcal{V}_{m,-1}(\mathbb{F}_p)$ by Theorem 4.3, and so does $-a$; they agree on $\{0, \dots, m-1\}$, so they are equal by Proposition 2.2. Reading $(Ta)_n = -a_n$ at a general n and using $(-1)^{p-1} = 1$ gives $a_{n+p} = a_{n+1} - a_n$. □

At $p = 2$ Theorem 5.5 is Touchard's congruence modulo 2. At $p = 3$ it says that A007472 satisfies $a_{n+3} \equiv a_{n+1} - a_n \pmod 3$, an order-3 recurrence for a sequence whose OEIS entry records no congruence at all. At $p = 5$ and $p = 7$ it accounts for the cells $(p, m) = (5, 4)$ and $(7, 6)$ of the deficiency table below.

Theorem 5.6 (Three sporadic deficient cells)**.** *For every $n \geq 0$:*

$$(p,m) = (2,3): \quad B^{(3)}_n + B^{(3)}_{n+1} + B^{(3)}_{n+4} \;\equiv\; 0 \pmod 2;$$

$$(p,m) = (3,4): \quad B^{(4)}_n + 2B^{(4)}_{n+1} + B^{(4)}_{n+2} + B^{(4)}_{n+4} + B^{(4)}_{n+6} + B^{(4)}_{n+9} \;\equiv\; 0 \pmod 3;$$

$$(p,m) = (7,4): \quad B^{(4)}_n + B^{(4)}_{n+2} + 5B^{(4)}_{n+8} + B^{(4)}_{n+14} \;\equiv\; 0 \pmod 7.$$

These are recurrences of degree 4, 9 and 14, against the degrees 6, 12 and 28 supplied by Theorem 3.4. In terms of the operator, the three relations say that the sequences are killed by $y^2 + y + 1$, $y^3 + y^2 + y + 1$ and $y^2 + 1$ respectively; the last reads $T^2 B^{(4)} = -B^{(4)}$ over $\mathbb{F}_7$, where Theorem 3.4 gives only $T^4 = \mathrm{id}$.

Proof. Each of the three displayed combinations is a linear combination of shifts of a sequence which, reduced modulo p, obeys the monic order-mp recurrence of Proposition 5.1. By Lemma 5.2 it therefore suffices to verify the combination for n in the window $0 \leq n < mp$, i.e. for 6, 12 and 28 values of n respectively; that is an exhaustive evaluation of finitely many residues, and it is what the formal

proof runs. The identification of the annihilators is the expansion of y^2+y+1, y^3+y^2+y+1 and y^2+1 at $y = X^p - c^{p-1}X$, using $c^{p-1}=1$ in each of the three cells. □

Remark 5.7. Berlekamp–Massey on the composite alone, over the same grid as §5.4, finds the following deficient cells with $p \nmid m$; the last column is the annihilator in $\mathbb{F}_p[y]/(y^m-1)$. This is computed data; the rows marked $\star$ are the ones proved above, either as instances of Theorem 5.5 or in Theorem 5.6.

(p,m)	mp	degree of $B^{(m)}$	annihilator
$(2,3)^\star$	6	4	y^2+y+1
$(2,7)$	14	8	y^4+y^2+y+1
$(3,2)^\star$	6	3	$y+1$
$(3,4)^\star$	12	9	y^3+y^2+y+1
$(5,4)^\star$	20	5	$y+1$
$(7,4)^\star$	28	14	y^2+1
$(7,6)^\star$	42	7	$y+1$

No cell with $p \in \{11,13\}$ in the computed range is deficient. In every deficient cell the annihilator is a polynomial in y dividing y^m-1, as Remark 4.4 predicts; which divisor occurs is not explained here.

6. Periods

That the Bell numbers are periodic modulo p is a consequence of (2) alone: Hall [4] deduced from it a period dividing p^p-1, and Williams [5] that the period divides $N_p = (p^p-1)/(p-1)$. Theorem 3.4 forces periodicity in the same way for every m: by Proposition 5.1 the sequence obeys a linear recurrence of order mp over $\mathbb{F}_p$ whose companion matrix is invertible, the coefficient of a_n being 1, so it is purely periodic with least period dividing the order of that matrix. What the least period actually is, is another matter; even for $m=1$ the question of whether N_p is least is open in general, though it is known to be least for all $p<126$ and for $p = 137, 149, 157, 163, 167, 173$ [8, Theorem 1.1].

We determine the least period of the composite $B^{(m)}$ modulo p in thirteen cells. In each case "least period T" means: $p \mid B^{(m)}_{n+T} - B^{(m)}_n$ for every n, and for every t with $0<t<T$ there is an n with $p \nmid B^{(m)}_{n+t} - B^{(m)}_n$. Leastness needs no separate divisibility argument: by Lemma 5.2, a period on the window $0 \le n < mp$ is a period outright, so failing on the window is failing outright.

Theorem 6.1. *The least period of* $B^{(m)} \bmod 2$ *is*

$$3,\ 1,\ 15,\ 1,\ 255,\ 1 \qquad \text{for } m = 1,2,3,4,5,6$$

respectively.

Theorem 6.2. *The least period of* $B^{(m)} \bmod 3$ *is*

$$13,\ 26,\ 1,\ 104,\ 1 \qquad \text{for } m = 1,2,3,4,6$$

respectively.

Proposition 6.3. *The least period of* $B^{(5)} \bmod 5$ *is* 1.

Theorem 6.4. *The least period of* $B^{(4)} \bmod 7$ *is* $48 = 7^2-1$.

Proof of 6.1, 6.2, 6.3, 6.4. All four rest on exhaustive computation, and it is worth being precise about how much. Fix a cell (p, m) and a candidate period T, and compute the residues of $B_i^{(m)}$ modulo p for all $i < mp + T$ directly from the recurrence. Then:

- T is a period if and only if $B_{n+T}^{(m)} \equiv B_n^{(m)}$ for all $n < mp$ — mp comparisons — by Proposition 5.1 and Lemma 5.2;
- no t with $0 < t < T$ is a period if and only if, for each such t, some $n < mp$ has $B_{n+t}^{(m)} \not\equiv B_n^{(m)}$ — at most $(T-1)\,mp$ comparisons.

So each cell is settled by at most $T \cdot mp$ comparisons of residues in a prefix of length $mp + T$. The largest searches carried out are $(p, m, T) = (2, 5, 255)$, a prefix of 265 residues and at most $255 \times 10 = 2550$ comparisons; the headline cell $(7, 4, 48)$, a prefix of 76 residues and at most $48 \times 28 = 1344$ comparisons; and $(3, 4, 104)$, a prefix of 116 residues and at most $104 \times 12 = 1248$ comparisons. For the cells with $p \mid m$ — namely $(2,2), (2,4), (2,6), (3,3), (3,6), (5,5)$ — the value $T = 1$ is also immediate from Proposition 5.4, since the composite has all its first m values equal to 1 and is then m-periodic, hence constant. □

6.1. **The table, and the absence of a formula.** Assembling the theorems above with cells computed but not formally verified (marked with a dagger; their windows exceed what the formal evaluator handles at reasonable cost) gives the following. The two entries $> 2 \cdot 10^6$ record a state-vector search truncated at $2 \cdot 10^6$ steps, not a computed period.

p	N_p	$m = 1$	2	3	4	5	6
2	3	3	1	15	1	255	1
3	13	13	26	1	104	$33215^\dagger$	1
5	781	$781^\dagger$	$1562^\dagger$	$1220703^\dagger$	$1562^\dagger$	1	$> 2 \cdot 10^{6\,\dagger}$
7	137257	$137257^\dagger$	$274514^\dagger$	$411771^\dagger$	**48**	$> 2 \cdot 10^{6\,\dagger}$	$274514^\dagger$

The row $p = 7$ begins N_7, $2N_7$, $3N_7$, suggesting the reading $m\,N_p$; that reading is broken three times. At $(5, 4)$ the period is $2N_5$, not $4N_5$. At $(2, 5)$ it is $255 = 85N_2$, not $5N_2$. And at $(7, 4)$ it collapses to $48 = 7^2 - 1$, against $4N_7 = 549028$: a factor of more than 10^4. That cell is exactly the one where the composite satisfies $T^2 = -1$ (Theorem 5.6), so the T-stable subspace the composite generates has dimension at most 2 rather than 4; the deficiency table and the period table are recording the same phenomenon twice. We do not prove that either one implies the other. We know of no formula for the least period, and none is recorded in the literature we have searched. The $m = 1$ column reproduces the classical least periods N_p for $p = 2, 3, 5, 7$ [8, Theorem 1.1].

7. Verification

Every theorem, proposition, lemma and corollary stated above has been formally verified in Lean 4 (toolchain `v4.32.0`) against *Mathlib*; the development contains no `sorry` and no additional axioms of our own. The three displayed items that are *not* theorems — the minimality grid of §5.4, the deficiency table of Remark 5.7 (whose starred rows are exactly the cells proved in Theorems 5.5 and 5.6) and the daggered entries of the period table in §6.1 — are labelled as computed data in the text and are outside the formal development. All of Sections 2–5, including Theorems 3.4, 3.6, 4.3, 5.5 and 5.6 and Propositions 2.2, 2.4, 5.1, 5.3 and 5.4, depends on no axioms beyond `propext`, `Classical.choice` and `Quot.sound`. What is delegated

to exhaustive evaluation is exactly the finite searches, and each is small: the seven checks of Proposition 5.3 are one integer division each; the three cells of Theorem 5.6 are 6, 12 and 28 residue checks; and the thirteen period cells of Section 6 are the residue comparisons counted in the proof there, at most 2550 in the largest case. The thirteen period statements are the only ones that use compiled evaluation (Lean's `native_decide`) rather than kernel reduction.

Three independent checks were run before any proof was written. Every numerical prefix printed in [1] was reproduced from the recurrence by a separate implementation and then again inside the formal development (Proposition 2.4); no printed value in [1] is wrong. The congruence of Theorem 3.4 was tested on 123 444 instances generated over $p \le 23$, $m \le 8$, with no violation, and separately on 14 094 instances built from the OEIS b-file for A007472 over the 30 primes up to 113 — data this work did not produce — again with no violation. The least periods of Section 6 were computed twice, by two independent implementations, before being formalised. Finally, Corollary 3.7 is proved for *Mathlib*'s own definition of the Bell numbers rather than for ours, which makes it a check of the general statement against an independently written definition.

8. Questions

Question 8.1. Which divisor of $y^m - 1$ is the annihilator of the composite $B^{(m)}$ in $\mathbb{F}_p[y]/(y^m - 1)$? Remark 5.7 lists the deficient cells found; the question is which cyclic submodule the all-ones initial vector generates, and it is a linear-algebra question over $\mathbb{F}_p$ rather than a search.

Question 8.2. Is $d(p, m) = mp$ whenever $p \nmid m$, as §5.4 suggests?

Question 8.3. Is there a formula, or even a divisibility rule, for the least period of $B^{(m)} \bmod p$? The table of §6.1 shows mN_p failing in three different ways. The primitive sequences $B^{(m,r)}$ have their own least periods, of which the period of the whole space is the least common multiple; nothing is proved about them here.

Question 8.4. Is there an m-analogue of Carlitz's congruence modulo $p^{\lceil k/2 \rceil}$ [3, §10]? The telescoping of Lemma 3.1 expanded one order further in p is the natural attack; we have not attempted it and have collected no data.

References

[1] V. Popov, *m-Bell and m-Stirling numbers: Iterated binomial transforms, hyper-Bessel functions, and moments of the Conway–Maxwell–Poisson distribution*, arXiv:2608.12011v1 (12 August 2026).

[2] J. Touchard, *Propriétés arithmétiques de certains nombres récurrents*, Ann. Soc. Sci. Bruxelles Sér. A **53** (1933), 21–31 (Zbl 0006.29102). This is the paper (2) is attributed to in [8, 9]; Gessel [3, §10] cites it together with Touchard's later *Nombres exponentiels et nombres de Bernoulli*, Canad. J. Math. **8** (1956), 305–320.

[3] I. M. Gessel, *Applications of the classical umbral calculus*, Algebra Universalis **49** (2003), no. 4, 397–434; arXiv:math/0108121v3.

[4] M. Hall, *Arithmetic properties of a partition function*, abstract 200, Bull. Amer. Math. Soc. **40** (1934), 387.

[5] G. T. Williams, *Numbers generated by the function $e^{e^x - 1}$*, Amer. Math. Monthly **52** (1945), 323–327.

[6] J. Levine and R. E. Dalton, *Minimum periods, modulo p, of first-order Bell exponential integers*, Math. Comp. **16** (1962), no. 80, 416–423.

[7] S. S. Wagstaff, Jr., *Aurifeuillian factorizations and the period of the Bell numbers modulo a prime*, Math. Comp. **65** (1996), no. 213, 383–391.
[8] P. L. Montgomery, S. Nahm and S. S. Wagstaff, Jr., *The period of the Bell numbers modulo a prime*, Math. Comp. **79** (2010), no. 271, 1793–1800.
[9] G. Serafin, *Backward Touchard congruence*, Bull. Belg. Math. Soc. Simon Stevin **28** (2022), no. 4, 603–614; arXiv:2110.06129.
[10] G. Serafin, *Congruences for Sheffer sequences*, Australas. J. Combin. **89** (2024), no. 1, 116–136.
[11] J. L. Massey, *Shift-register synthesis and BCH decoding*, IEEE Trans. Inform. Theory **15** (1969), no. 1, 122–127.
[12] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org`; entries A000110, A004211, A007472, A351028, A351143.
[13] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381.