

Integrality of iterated averages and pointwise prime powers

GPT-6 Astra, edited and proofread by Darij Grinberg*

October 27, 2026

Abstract

Suppose that we repeatedly replace the values of an integer-valued function on $\mathbb{Z}$ by averages over p consecutive arguments, where p is prime. If the first ℓ rounds of averaging give integers, will this remain true when we begin with the p th powers of the original values? We prove that the answer is yes for $\ell \leq p$. For $p = 2$ and $p = 3$, the answer is yes for every ℓ , and therefore also when integrality is required at every stage. For larger primes, a known counterexample first fails at round $p + 1$. The proofs use finite differences and polynomial interpolation. We also explain connections with earlier work on divisibility of finite differences and Stirling numbers.

Manifest. This is an answer to MathOverflow question #68595 (which also gives a new answer to #66959). It was produced by GPT-6 Astra (Researcher subscription). The main proof was found in one 3-minute shot, and expanded and polished over several rounds of edits. I have proofread the whole text except for Section 8.

This text is in the public domain. –Darij

1 Statement and finite differences

Start with an integer-valued function on the integers. Replace each value by the average of that value and the next $p - 1$ values, where p is a prime. The new values need not be integers. If they are, we can repeat the operation and ask whether they remain integers at each stage. A classical olympiad problem, proposed by D. Tamarkin at the All-Russian Mathematical Olympiad 1993 [9], concerns this process for $p = 2$: It claims that if every stage is integer-valued, then the same is true when we start with the squares of the original values. In this note, we consider the case of general p and taking the entries to the p th power; we also ask what happens when we assume integrality for only finitely many stages. The results revealed are weaker but no less beautiful.

We begin by fixing the notation. We consider the $\mathbb{Q}$ -vector space $\mathbb{Q}^{\mathbb{Z}}$, which consists of all functions $f : \mathbb{Z} \rightarrow \mathbb{Q}$. These functions can be regarded as infinite sequences

$$(\dots, f(-2), f(-1), f(0), f(1), f(2), \dots)$$

*Section 8 has not been proofread.

(infinite in both directions). Such a function $f : \mathbb{Z} \rightarrow \mathbb{Q}$ is said to be *integral* if all its values $f(n)$ are integers, i.e., if $f \in \mathbb{Z}^{\mathbb{Z}}$. Sums, differences, and integer multiples of integral functions are again integral; in other words, $\mathbb{Z}^{\mathbb{Z}}$ is a $\mathbb{Z}$ -submodule of $\mathbb{Q}^{\mathbb{Z}}$.

We shall consider linear (that is, $\mathbb{Q}$ -linear) operators on $\mathbb{Q}^{\mathbb{Z}}$ such as the shift and finite-difference operators defined below. Here linearity means that $A(af + bg) = aAf + bAg$ for rational numbers a, b and functions f, g . If A is such an operator and $f \in \mathbb{Q}^{\mathbb{Z}}$ is a function, then we will write $Af(n)$ for the value $(Af)(n)$.

Fix a prime p . Define a linear operator $F_p : \mathbb{Q}^{\mathbb{Z}} \rightarrow \mathbb{Q}^{\mathbb{Z}}$ (called the *averaging operator*) by

$$(F_p f)(n) = \frac{1}{p} \sum_{r=0}^{p-1} f(n+r) \quad \text{for all } f \in \mathbb{Q}^{\mathbb{Z}} \text{ and } n \in \mathbb{Z}.$$

That is, F_p acts on a function f by replacing each value $f(n)$ by the average of p consecutive values (starting with $f(n)$). Furthermore, for a function $f : \mathbb{Z} \rightarrow \mathbb{Q}$, define the function $f^p : \mathbb{Z} \rightarrow \mathbb{Q}$ by

$$f^p(n) = f(n)^p \quad \text{for all } n \in \mathbb{Z}.$$

Operator powers mean repeated application: $F_p^2 f = F_p(F_p f)$, and $F_p^0 f = f$. These are different from the pointwise powers f^p .

A function $f : \mathbb{Z} \rightarrow \mathbb{Q}$ will be called *ℓ -average-integral* for a given nonnegative integer ℓ if $F_p^k f$ is integral for every $0 \leq k \leq \ell$. Note that an ℓ -average-integral function f is thus automatically integral (since we can take $k = 0$ and obtain the integrality of $F_p^k f = F_p^0 f = f$). Thus the parameter ℓ in “ ℓ -average-integral” records how many successive averaging steps we require to give integers.

We say that $f : \mathbb{Z} \rightarrow \mathbb{Q}$ is *average-integral* if $F_p^k f$ is integral for every $k \geq 0$. Of course, this is the same as being ℓ -average-integral for all $\ell \geq 0$ simultaneously. For example, when $p = 2$, the function $f : n \mapsto 2n$ is average-integral: its k th average-iterate $F_p^k f$ is $n \mapsto 2n + k$.

We will prove the following two properties of ℓ -average-integrality:

Theorem 1.1. *Let $0 \leq \ell \leq p$. If $f : \mathbb{Z} \rightarrow \mathbb{Z}$ is ℓ -average-integral, then f^p is ℓ -average-integral.*

Theorem 1.2. *Let $\ell \geq 0$ be arbitrary.*

1. *For $p = 2$, the ℓ -average-integral functions form a ring under pointwise addition and pointwise multiplication.*
2. *For $p = 3$, if f is ℓ -average-integral, then f^3 is ℓ -average-integral.*

Corollary 1.3. *The set of all average-integral functions is closed under pointwise products for $p = 2$ and under pointwise cubes (i.e., under the map $f \mapsto f^3$) for $p = 3$.*

Theorem 1.1 answers a question asked in [1]. Corollary 1.3 was already discussed in [2]. For primes $p > 3$, the bound in Theorem 1.1 cannot be increased: Proposition 6.1 gives a function whose p th power first fails at round $p + 1$.

Our proofs use several ideas suggested by Zarathustra Brady (“zeb” on MathOverflow) in his answer to [2]. They have two major ingredients in common. First we translate integrality of averages into divisibility of finite differences. Then we use polynomial interpolation to

study these divisibilities on a finite interval. This proves Theorem 1.1 in Section 3. To handle arbitrarily many rounds for $p = 2, 3$ (and thus prove Theorem 1.2 and Corollary 1.3), we will need a fuller description of the polynomials that remain integral under averaging.

Our main tool is the *finite difference operator* $\Delta : \mathbb{Q}^{\mathbb{Z}} \rightarrow \mathbb{Q}^{\mathbb{Z}}$ defined by

$$\Delta f(n) = f(n+1) - f(n) \quad \text{for all } f \in \mathbb{Q}^{\mathbb{Z}} \text{ and } n \in \mathbb{Z}.$$

Unlike averaging, taking a finite difference always preserves integrality; i.e., we have $\Delta f \in \mathbb{Z}^{\mathbb{Z}}$ for each $f \in \mathbb{Z}^{\mathbb{Z}}$. We will express the condition on repeated averages as a condition on the divisibility of repeated differences.

Define the *shift operator* $E : \mathbb{Q}^{\mathbb{Z}} \rightarrow \mathbb{Q}^{\mathbb{Z}}$ by

$$Ef(n) = f(n+1) \quad \text{for all } f \in \mathbb{Q}^{\mathbb{Z}} \text{ and } n \in \mathbb{Z}.$$

Thus, $\Delta = E - \text{id}$. All three operators F_p , E , and Δ are $\mathbb{Q}$ -linear endomorphisms of $\mathbb{Q}^{\mathbb{Z}}$. The latter two preserve $\mathbb{Z}^{\mathbb{Z}}$. The first few differences of a function $f : \mathbb{Z} \rightarrow \mathbb{Q}$ are

$$\Delta f(n) = f(n+1) - f(n), \quad \Delta^2 f(n) = f(n+2) - 2f(n+1) + f(n).$$

Set $q = p - 1$. All divisibility statements about functions below are pointwise: for example, $\Delta^j f \in p^k \mathbb{Z}^{\mathbb{Z}}$ means that p^k divides $\Delta^j f(n)$ for every integer n . The next lemma says that – as far as integrality is concerned – one averaging step “corresponds” to a block of $q = p - 1$ difference steps. For instance, when $p = 3$ and f is integral, two integral averaging steps are equivalent to the two conditions $3 \mid \Delta^2 f(n)$ and $9 \mid \Delta^4 f(n)$ for every n .

Lemma 1.4 (Difference criterion). *For a function $f : \mathbb{Z} \rightarrow \mathbb{Q}$ and any nonnegative integer $\ell \geq 0$, the following conditions are equivalent:*

1. *The function f is ℓ -average-integral.*
2. *We have $\Delta^{kq} f \in p^k \mathbb{Z}^{\mathbb{Z}}$ for all $0 \leq k \leq \ell$.*
3. *We have $\Delta^j f \in p^{\lfloor j/q \rfloor} \mathbb{Z}^{\mathbb{Z}}$ for all $0 \leq j \leq \ell q$.*

Proof. We assume that f is integral, since all three of the ostensibly equivalent conditions imply this¹.

We first compare one averaging step with q difference steps. If $A(X) = \sum_i a_i X^i$ is a polynomial, then $A(E)$ means the operator $\sum_i a_i E^i$. Since $(E^i f)(n) = f(n+i)$, we can rewrite the definition of F_p as

$$F_p = \frac{1}{p}(\text{id} + E + \cdots + E^{p-1}). \tag{1}$$

Hence,

$$pF_p = \text{id} + E + \cdots + E^{p-1}. \tag{2}$$

¹Indeed, the second condition, applied to $k = 0$, says precisely that $f \in \mathbb{Z}^{\mathbb{Z}}$, that is, f is integral. The third condition, applied to $j = 0$, likewise yields the same conclusion. And of course, the first condition yields integrality of f as well, as we already know.

Moreover, from $\Delta = E - \text{id}$ and $q = p - 1$, we obtain $\Delta^q = (E - \text{id})^{p-1}$. These formulas explain why the following polynomial congruence is useful:

$$1 + X + \cdots + X^{p-1} \equiv (X - 1)^{p-1} \pmod{p}. \quad (3)$$

To see this congruence, reduce modulo p and use $X^p - 1 = (X - 1)^p$. Dividing both sides by $X - 1$ in the polynomial ring $\mathbb{F}_p[X]$ gives the congruence (3).

That is, in $\mathbb{Z}[X]$, we have

$$1 + X + \cdots + X^{p-1} = (X - 1)^{p-1} + pB(X) \quad \text{for some } B(X) \in \mathbb{Z}[X].$$

Now substitute the shift operator E for X in this equality. This gives the operator identity

$$pF_p = \Delta^q + pB(E)$$

(since $pF_p = \text{id} + E + \cdots + E^{p-1}$ and $\Delta^q = (E - \text{id})^{p-1}$). Putting

$$D = \Delta^q/p \quad (4)$$

and dividing by p , we obtain

$$F_p = D + B(E), \quad (5)$$

so that $D = F_p - B(E)$ and $B(E) = F_p - D$. The important point is that $B(E)$ preserves $\mathbb{Z}^{\mathbb{Z}}$: it is an integer linear combination of shifts. For a given integral function f , the identity (5) therefore shows that $F_p f$ is integral if and only if Df is integral. For example, when $p = 3$, we have $B(X) = X$ and

$$F_3 f(n) = f(n+1) + \frac{f(n+2) - 2f(n+1) + f(n)}{3}.$$

The numerator in the fraction is $\Delta^2 f(n)$, making the one-step divisibility condition particularly transparent.

The same comparison works for every iterate. All three operators $F_p, D, B(E)$ are polynomials in E , so they commute, and the binomial theorem applies. If $F_p^i f$ is integral for $0 \leq i \leq k$, then (from $D = F_p - B(E)$) we see that

$$D^k f = (F_p - B(E))^k f = \sum_{i=0}^k \binom{k}{i} (-B(E))^{k-i} F_p^i f$$

is integral as well: each summand is obtained from an integral function (viz., $F_p^i f$) by applications of the operator $B(E)$ (which preserves $\mathbb{Z}^{\mathbb{Z}}$) and integer linear combinations. Conversely, suppose that $D^i f$ is integral for every $0 \leq i \leq k$. Using $F_p = D + B(E)$ and the commutativity of D and $B(E)$, we obtain

$$F_p^k f = (D + B(E))^k f = \sum_{i=0}^k \binom{k}{i} B(E)^{k-i} D^i f.$$

Each $D^i f$ is integral by assumption, and applying $B(E)$ any number of times preserves integrality. Multiplication by the integer $\binom{k}{i}$ also preserves integrality, as does summing the resulting functions. Thus $F_p^k f$ is integral. Altogether, we thus have shown that $F_p^k f$ is integral for every $0 \leq k \leq \ell$ if and only if $D^k f$ is integral for every $0 \leq k \leq \ell$. Since $D^k f = \Delta^{kq} f / p^k$, this proves the equivalence of conditions 1 and 2 in the lemma.

It remains to show that conditions 2 and 3 are equivalent. Clearly, condition 3 implies condition 2 (by applying it to $j = kq$). Conversely, consider a difference order j that is not necessarily a multiple of q . Dividing j by q with remainder, we write j as $j = kq + r$, where $0 \leq r < q$; then $k = \lfloor j/q \rfloor \leq \ell$. Condition 2 thus says that $\Delta^{kq} f = p^k g$ for an integral function g . Taking r more differences then gives $\Delta^j f = p^k \Delta^r g$ (since $j = kq + r$), which is still divisible by p^k at every integer. That is, $\Delta^j f \in p^k \mathbb{Z}^{\mathbb{Z}} = p^{\lfloor j/q \rfloor} \mathbb{Z}^{\mathbb{Z}}$. This yields condition 3. $\square$

2 The polynomial congruence

Our next step is a statement about polynomials. We will later use it on a polynomial that agrees with f on a finite interval. A small amount of notation lets us keep track of powers of p without being concerned about other prime factors.

Let $R = \mathbb{Z}_{(p)}$, the subring of $\mathbb{Q}$ consisting of fractions (of the form r/s with $r, s \in \mathbb{Z}$) whose denominators are not divisible by p . For example, if $p = 3$, then $1/2$ belongs to R , whereas $1/3$ does not. A statement such as $a \in p^k R$ (for a given nonnegative integer k) means that, after cancellation, the numerator of a has at least k factors of p and its denominator has none. When a is an integer, this is just the usual divisibility by p^k .

For a nonzero rational number a , write $v_p(a)$ for the exponent of p in its numerator minus the exponent in its denominator, and set $v_p(0) = +\infty$. Thus $a \in p^k R$ is equivalent to $v_p(a) \geq k$. We will use Legendre's formula (also known as de Polignac's formula)

$$v_p(t!) = \sum_{i \geq 1} \left\lfloor \frac{t}{p^i} \right\rfloor \quad \text{for all } t \geq 0. \quad (6)$$

Indeed, among $1, \dots, t$, each multiple of p contributes one factor of p , each multiple of p^2 contributes a further factor, and so on. The sum counts every factor exactly once.

We use the same operators E, Δ, F_p on polynomials as on functions:

$$(EH)(x) = H(x+1), \quad (\Delta H)(x) = H(x+1) - H(x), \quad (F_p H)(x) = \frac{1}{p} \sum_{r=0}^{p-1} H(x+r).$$

These definitions are compatible with evaluation at integers. More precisely, for any $H \in \mathbb{Q}[x]$, let $\text{ev}(H) \in \mathbb{Q}^{\mathbb{Z}}$ be the function $n \mapsto H(n)$. Then

$$\text{ev}(AH) = A \text{ev}(H) \quad \text{for } A = E, \Delta, F_p. \quad (7)$$

For example, both sides for $A = \Delta$, evaluated at n , equal $H(n+1) - H(n)$. Thus it makes no difference whether we first take a polynomial difference and then evaluate, or first evaluate

and then take the difference of the resulting function. The same holds for any number of iterations. We will use this compatibility when passing between polynomials and functions.

For finite differences, a convenient R -module basis of the polynomial ring $R[x]$ is given by the *falling factorials*

$$(x)_j = x(x-1) \cdots (x-j+1) \quad \text{for all } j \geq 0.$$

The first few are $(x)_0 = 1$, $(x)_1 = x$, $(x)_2 = x(x-1)$, and $(x)_3 = x(x-1)(x-2)$. Each has leading coefficient 1, with one polynomial of each degree. We can therefore expand any polynomial in $R[x]$ uniquely in this basis: subtract a multiple of the falling factorial of highest degree, then repeat with the remainder. For example,

$$x^2 = (x)_2 + (x)_1, \quad x^3 = (x)_3 + 3(x)_2 + (x)_1.$$

So each polynomial $P \in R[x]$ can be written uniquely as $\sum_{j \geq 0} c_j (x)_j$ with $c_j \in R$ (such that $c_j = 0$ for all but finitely many j). The coefficients c_j are called the *falling-factorial coefficients* of P . This holds equally well for any commutative ring in place of R .

The falling factorials also satisfy $\Delta(x)_0 = 0$ and

$$\Delta(x)_j = j(x)_{j-1} \quad \text{for all } j \geq 1. \quad (8)$$

This follows by factoring $(x)_{j-1}$ out of $(x+1)_j - (x)_j$; the remaining factor is $(x+1) - (x-j+1) = j$. Iterating (8) and setting $x = 0$ shows that, if $H(x) = \sum_{j \geq 0} c_j (x)_j$, then

$$\Delta^m H(0) = m! c_m \quad \text{for all } m \geq 0. \quad (9)$$

Indeed, terms $c_j (x)_j$ with $j < m$ have disappeared after m differences, and those with $j > m$ still have a factor x , so only the term for $j = m$ remains when we evaluate at $x = 0$.

We first isolate a fact about falling-factorial coefficients over $\mathbb{F}_p$, the field of residues modulo p . The falling factorials form a basis of $\mathbb{F}_p[x]$ by the same leading-coefficient argument as above. For $P \in \mathbb{F}_p[x]$, write $[(x)_i]P$ for the $(x)_i$ -coefficient of P in this basis. We set $[(x)_i]P = 0$ when $i < 0$; this is a convention for coefficients and does not define falling factorials with negative indices.

Multiplication by $x^p - x$ has a particularly simple effect on this basis: it shifts every index up by p .

Lemma 2.1. *Put $y = x^p - x \in \mathbb{F}_p[x]$. Then:*

1. *We have*

$$y(x)_i = (x)_{p+i} \quad \text{for all } i \geq 0. \quad (10)$$

2. *More generally, we have*

$$y^a(x)_j = (x)_{ap+j} \quad \text{for all } a, j \geq 0. \quad (11)$$

3. *For every $P \in \mathbb{F}_p[x]$, we have*

$$[(x)_i](yP) = [(x)_{i-p}]P \quad \text{for all integers } i. \quad (12)$$

4. More generally, for every $P \in \mathbb{F}_p[x]$, we have

$$[(x)_i](y^a P) = [(x)_{i-ap}]P \quad \text{for all } a \geq 0 \text{ and all integers } i. \quad (13)$$

Proof. We first recall the classical finite-field property

$$x(x-1) \cdots (x-p+1) = x^p - x,$$

which holds because both sides are monic polynomials of degree p whose roots are precisely the elements of $\mathbb{F}_p$. This identity is also known as Lagrange's congruence; see [8, Section 10], where it is used in an umbral proof of Touchard's congruence for Bell numbers. Now, we study the product $(x)_{p+i} = x(x-1) \cdots (x-p-i+1)$. Its first p factors have product

$$x(x-1) \cdots (x-p+1) = x^p - x = y.$$

Its remaining i factors are $x-p, x-p-1, \dots, x-p-i+1$, which (because we are in characteristic p) are precisely the same as $x, x-1, \dots, x-i+1$; thus their product is $(x)_i$. So we obtain $y(x)_i = (x)_{p+i}$. This proves (10) (including the case $i = 0$, when the remaining product is empty). Applying (10) successively a times yields (11).

To prove (12), write $P = \sum_{j \geq 0} c_j(x)_j$. Part (1) gives

$$yP = \sum_{j \geq 0} c_j(x)_{p+j}.$$

Thus the $(x)_i$ -coefficient is c_{i-p} if $i \geq p$, and is zero otherwise. This is exactly (12), with our convention for negative indices. To prove (13), apply (12) successively a times. $\square$

We now use this shift to describe which coefficients of a p th power can occur at indices less than p^2 .

Lemma 2.2. *Let $H \in \mathbb{F}_p[x]$, and let Q be the remainder when H is divided by $x^p - x$. Suppose that $\deg Q \leq d$, where $0 \leq d < p$, and write*

$$H(x)^p = \sum_{j \geq 0} \gamma_j(x)_j \quad \text{with } \gamma_j \in \mathbb{F}_p$$

(so that γ_j is the $(x)_j$ -coefficient of $H(x)^p$ for each $j \geq 0$). Then

$$\gamma_{ap+b} = 0 \quad \text{for all } 0 \leq a, b < p \text{ with } a+b > d.$$

Thus a nonzero $(x)_j$ -coefficient at an index $j < p^2$ can occur only when the quotient and remainder of j upon division by p sum to at most d .

Proof. Put $y = x^p - x$ and write $H = Q + yT$. We use the Frobenius endomorphism of $\mathbb{F}_p[x]$, namely the map $U \mapsto U^p$. This map preserves sums and products: in particular, $(U + V)^p = U^p + V^p$, by the binomial theorem in characteristic p . Also, it fixes every coefficient in $\mathbb{F}_p$, so $Q(x)^p = Q(x^p)$. Thus, from $H = Q + yT$, we obtain

$$H(x)^p = Q(x)^p + y^p T(x)^p = Q(x^p) + y^p T(x)^p = Q(x+y) + y^p T(x)^p \quad (14)$$

(since $x^p = x + y$). We examine the falling-factorial coefficients of the two terms separately; each time we will show that the $(x)_{ap+b}$ -coefficient is 0 whenever $0 \leq a, b < p$ with $a + b > d$.

For the second term, apply the coefficient-shift identity (13) with $a = p$ and $P = T(x)^p$. It gives

$$[(x)_i](y^p T(x)^p) = [(x)_{i-p^2}](T(x)^p) = 0 \quad \text{for all } 0 \leq i < p^2.$$

The last equality holds because $i - p^2 < 0$. Thus the second term contributes nothing at any of the indices under consideration.

It remains to examine $Q(x + y)$. Expand it first as a polynomial in two formal variables x and y , before making the substitution $y = x^p - x$. Because $\deg Q \leq d$, every term is a multiple of $x^b y^a$ with $a + b \leq d$. Express x^b as a linear combination of $(x)_t$ for $0 \leq t \leq b$. Multiplication by y^a then turns these into $(x)_{ap+t}$, by (11). Thus, $Q(x + y)$ is expanded as an $\mathbb{F}_p$ -linear combination of falling factorials $(x)_{ap+t}$ whose indices $ap + t$ satisfy

$$a, t \geq 0 \text{ and } a + t \leq d.$$

Since $d < p$, we have $a, t < p$, so a and t are precisely the quotient and remainder when the index is divided by p . This proves the claimed vanishing. $\square$

The next lemma extracts the divisibility we need from just one condition: the sum of P over p consecutive arguments is divisible by p .

Lemma 2.3. *Suppose $P \in R[x]$ satisfies*

$$\sum_{r=0}^{p-1} P(r) \in pR. \tag{15}$$

Then

$$\Delta^{k(p-1)}(P^p)(0) \in p^k R \quad \text{for all } 1 \leq k \leq p.$$

Proof. Write $P(x)^p = \sum_{j \geq 0} c_j(x)_j$ with $c_j \in R$. For $m = k(p - 1)$ and $1 \leq k \leq p$, we have $m < p^2$ and thus, by (6), we have

$$v_p(m!) = \left\lfloor \frac{m}{p} \right\rfloor = k - 1. \tag{16}$$

(For the last equality, note that $m/p = k - k/p$. Since $1 \leq k \leq p$, we have $0 < k/p \leq 1$, and hence $k - 1 \leq m/p < k$. Thus $\lfloor m/p \rfloor = k - 1$.) By (9), the difference we want to study is

$$\Delta^m(P^p)(0) = m! c_m.$$

We need this product to be divisible by p^k . Formula (16) says that $m!$ already supplies exactly $k - 1$ factors of p . Thus we need just one more factor from c_m : our remaining task is to show $c_m \in pR$ for $m = k(p - 1)$.

This is a question modulo p . Fractions in R can be reduced modulo p by inverting their denominators in $\mathbb{F}_p$ (there is a canonical ring isomorphism $R/pR \cong \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$), so we may now work in $\mathbb{F}_p[x]$. Here $\mathbb{F}_p$ is the field of residues modulo p , and $\bar{P}$ denotes the reduction of P . The coefficients in the falling-factorial expansion also reduce modulo p : reducing the

identity $P(x)^p = \sum_{j \geq 0} c_j(x)_j$ modulo p gives $\overline{P}^p = \sum_{j \geq 0} \overline{c_j}(x)_j$, so that the falling-factorial coefficients of $\overline{P}^p$ are the $\overline{c_j}$. That is, $\overline{c_j} = [(x)_j](\overline{P}^p)$ for each $j \geq 0$.

We first replace $\overline{P}$ by a polynomial of degree less than p that takes the same values on $\mathbb{F}_p$. Put $y = x^p - x$ and divide $\overline{P}$ by this monic polynomial y to obtain

$$\overline{P} = Q + yT, \quad \text{where } Q, T \in \mathbb{F}_p[x] \text{ with } \deg Q < p.$$

Fermat's little theorem says that $y(r) = 0$ for every $r \in \mathbb{F}_p$. Hence $Q(r) = \overline{P}(r)$ for every $r \in \mathbb{F}_p$, and thus our hypothesis (15) rewrites as

$$\sum_{r \in \mathbb{F}_p} Q(r) = 0. \quad (17)$$

This forces the coefficient of x^{p-1} in Q to vanish. The reason is the famous power-sum identity

$$\sum_{r \in \mathbb{F}_p} r^b = \begin{cases} 0, & \text{if } 0 \leq b < p-1, \\ -1, & \text{if } b = p-1 \end{cases} \quad \text{for every } 0 \leq b < p, \quad (18)$$

which we prove in Appendix A. Applying (18) term by term to Q shows that $\sum_{r \in \mathbb{F}_p} Q(r)$ is the negative of the x^{p-1} -coefficient of Q (since $\deg Q < p$). Thus, (17) tells us that the latter coefficient is 0. In other words,

$$\deg Q \leq p-2.$$

We can therefore apply Lemma 2.2 to $H = \overline{P}$ and $d = p-2$. It says that the falling-factorial coefficient $[(x)_{ap+b}](\overline{P}^p)$ at an index $ap+b$ vanishes whenever $0 \leq a, b < p$ and $a+b > p-2$. Our desired indices $k(p-1)$ (for $1 \leq k \leq p$) have exactly this property, since they can be written as

$$k(p-1) = ap+b \quad \text{for } a = k-1 \text{ and } b = p-k$$

(and these values a and b satisfy $0 \leq a < p$ and $0 \leq b < p$ and $a+b = (k-1) + (p-k) = p-1 > p-2$). Consequently, $[(x)_{k(p-1)}](\overline{P}^p)$ vanishes. This says exactly that $c_{k(p-1)} \in pR$, supplying the one factor of p still needed to complete the proof. $\square$

Remark 2.4. For example, when $p = 5$ and $\deg Q \leq 3$, the possible indices less than 25 in the falling-factorial expansion of $\overline{P}^5$ are

$$0, 1, 2, 3, \quad 5, 6, 7, \quad 10, 11, \quad 15.$$

None is one of 4, 8, 12, 16, 20, the indices needed for the first five averaging depths. The factorials at these latter indices have respective 5-adic valuations 0, 1, 2, 3, 4; the lemma supplies one further factor of 5.

3 From local interpolation to arbitrary functions

Recall that our interest lies in functions $f : \mathbb{Z} \rightarrow \mathbb{Q}$, not in polynomials. However, a value of the m -th finite difference $\Delta^m f$ depends on only $m+1$ values of f . Thus, even though our

function f need not itself be a polynomial, we can replace it by a polynomial separately on each interval of the required length. To be precise, expanding

$$\Delta^m = (E - \text{id})^m = \sum_{r=0}^m (-1)^{m-r} \binom{m}{r} E^r$$

gives

$$\Delta^m f(n) = \sum_{r=0}^m (-1)^{m-r} \binom{m}{r} f(n+r). \quad (19)$$

This formula makes precise how an interval of agreement shrinks when we take differences. Suppose two functions $f, g : \mathbb{Z} \rightarrow \mathbb{Q}$ agree at every integer in an interval $[a, b]$, where $a \leq b$ are integers. Then

$$\Delta^j f(t) = \Delta^j g(t) \quad \text{for all } 0 \leq j \leq b-a \text{ and } a \leq t \leq b-j. \quad (20)$$

Indeed, the formula for $\Delta^j f(t)$ uses only the values at $t, t+1, \dots, t+j$. All these arguments lie in $[a, b]$, so each value of f in the sum can be replaced by the corresponding value of g . Thus each further application of Δ loses the rightmost point of the interval of agreement, and j iterations of Δ lose its last j points.

In particular, suppose a polynomial $P \in \mathbb{Q}[x]$ satisfies $P(r) = f(n+r)$ for all $0 \leq r \leq m$ (where n and $m \geq 0$ are two fixed integers). Applying the same observation to the functions $r \mapsto P(r)$ and $r \mapsto f(n+r)$ gives

$$(\Delta^j P)(r) = (\Delta^j f)(n+r) \quad \text{for all } 0 \leq j \leq m \text{ and } 0 \leq r \leq m-j. \quad (21)$$

Here polynomial differences agree with differences of the evaluated function by (7). For $j = m$, only $r = 0$ remains: $\Delta^m P(0) = \Delta^m f(n)$. The same reasoning applies to P^p and f^p , since taking pointwise powers preserves the original equalities of values.

Newton interpolation constructs the polynomial we need directly from these differences. Recall that for each $j \geq 0$, the polynomial $\binom{x}{j} = (x)_j / j! \in \mathbb{Q}[x]$ is integer-valued². For any function $f : \mathbb{Z} \rightarrow \mathbb{Q}$, the polynomial

$$P(x) = \sum_{j=0}^m \Delta^j f(n) \binom{x}{j} \quad (22)$$

(known as the m -th *Newton interpolant* of f based at n) agrees with $f(n+x)$ at the integers $x = 0, \dots, m$. For example, to match three values $u = f(n)$, $v = f(n+1)$, and $w = f(n+2)$, the formula gives

$$P(x) = u + (v-u)x + (w-2v+u) \frac{x(x-1)}{2}.$$

Substitution of $x = 0, 1, 2$ gives u, v, w , respectively. In the general case, the formula

$$P(r) = f(n+r) \quad \text{for all } 0 \leq r \leq m \quad (23)$$

²Indeed, its values at nonnegative integers are the standard binomial coefficients, whereas its values at negative integers can be reduced to the nonnegative ones via the upper negation formula $\binom{-a}{j} = (-1)^j \binom{a+j-1}{j}$.

is verified by the same binomial theorem we used for the operators: since $E = \text{id} + \Delta$, for $0 \leq r \leq m$ we have

$$f(n+r) = (E^r f)(n) = ((\text{id} + \Delta)^r f)(n) = \sum_{j=0}^r \binom{r}{j} \Delta^j f(n) = P(r)$$

(here, all addends with $j > r$ in $P(r)$ vanish because $\binom{r}{j} = 0$). We apply this formula using the divisibility supplied by Lemma 1.4.

Proof of Theorem 1.1. The case $\ell = 0$ says only that f^p is integral, which is clear since f is integral. Suppose $1 \leq \ell \leq p$. Set $q = p - 1$. By Lemma 1.4 (equivalence of the first two conditions, applied to f^p instead of f), our goal is to show that $\Delta^{kq}(f^p) \in p^k \mathbb{Z}^{\mathbb{Z}}$ for all $0 \leq k \leq \ell$. In other words, we must prove that

$$p^k \mid \Delta^{kq}(f^p)(n) \quad \text{for all } n \in \mathbb{Z} \text{ and } 0 \leq k \leq \ell. \quad (24)$$

Fix such n and k , and put $m = kq$; thus, $m = kq \leq \ell q \leq pq = p(p-1) < p^2$. We assume that $k \geq 1$, since the case $k = 0$ is obvious. Define the polynomial

$$P(x) = \sum_{j=0}^m \Delta^j f(n) \binom{x}{j} = \sum_{j=0}^m \frac{\Delta^j f(n)}{j!} (x)_j. \quad (25)$$

Newton interpolation – i.e., the formula (23) – gives

$$P(r) = f(n+r) \quad \text{for all } 0 \leq r \leq m. \quad (26)$$

To apply Lemma 2.3, we must check its two hypotheses. First, the coefficients of P must belong to R . This is not automatic from the interpolation formula: division by $j!$ might put p in a denominator. However, the difference bounds prevent this. Indeed, Lemma 1.4 (this time applied to f itself) yields that each $j \leq m$ satisfies $\Delta^j f \in p^{\lfloor j/(p-1) \rfloor} \mathbb{Z}^{\mathbb{Z}}$ (since $j \leq m \leq \ell q$), so that $v_p(\Delta^j f(n)) \geq \lfloor j/(p-1) \rfloor$. Since $j \leq m < p^2$, the formula (6) yields

$$v_p(j!) = \left\lfloor \frac{j}{p} \right\rfloor \leq \left\lfloor \frac{j}{p-1} \right\rfloor.$$

The numerator $\Delta^j f(n)$ in (25) therefore contains at least as many factors of p as the denominator $j!$. After cancellation, the $(x)_j$ -coefficient therefore belongs to R . Since each falling factorial has integer coefficients, this proves $P \in R[x]$.

Second, the sum of P over $0, \dots, p-1$ must belong to pR . Indeed, since $k \geq 1$, we have $m = kq \geq q = p-1$, so that each $r \in \{0, 1, \dots, p-1\}$ satisfies $r \leq m$ and therefore $P(r) = f(n+r)$ by (26). Hence,

$$\sum_{r=0}^{p-1} P(r) = \sum_{r=0}^{p-1} f(n+r) = p(F_p f)(n) \in p\mathbb{Z}$$

(since $1 \leq \ell$ and the ℓ -average-integrality of f entail that $F_p f$ is integral, so $(F_p f)(n) \in \mathbb{Z}$). Thus, $\sum_{r=0}^{p-1} P(r) \in p\mathbb{Z} \subseteq pR$. Lemma 2.3 consequently yields $\Delta^{k(p-1)}(P^p)(0) \in p^k R$. In view

of $m = kq = k(p-1)$, this rewrites as $\Delta^m(P^p)(0) \in p^k R$. But (21) (applied to P^p , f^p , m and 0 instead of P , f , j and r) yields $\Delta^m(P^p)(0) = \Delta^m(f^p)(n)$. Hence,

$$\Delta^m(f^p)(n) = \Delta^m(P^p)(0) \in p^k R.$$

The left side is an integer, so it is therefore divisible by p^k . In other words, $p^k \mid \Delta^m(f^p)(n) = \Delta^{kq}(f^p)(n)$ (since $m = kq$). This proves (24). As n and k were arbitrary, Lemma 1.4 proves that f^p is ℓ -average-integral. Theorem 1.1 is thus proved. $\square$

4 Polynomials and repeated averaging

The finite-depth theorem is now proved. To study larger depths, it helps to answer a more specific question: which polynomials remain integer-valued after arbitrarily many averaging steps?

A monomial x^t need not have this property, but multiplying it by a suitable power of p will make it work. The exponent that appears is

$$d(t) = \left\lfloor \frac{t}{p-1} \right\rfloor - v_p(t!).$$

There is a reason for this expression. The difference criterion (Lemma 1.4) asks for $\lfloor t/(p-1) \rfloor$ factors of p in a t th difference, while the t th difference of x^t is $t!$. Thus $d(t)$ counts the additional factors needed at this highest difference order. The next lemma shows that these factors also suffice for **every** order. We then obtain the polynomial criterion from zeb's answer [3], and use it both for $p = 2, 3$ and for the counterexample when $p > 3$.

Lemma 4.1. *For $t \geq 0$, put*

$$d(t) = \left\lfloor \frac{t}{p-1} \right\rfloor - v_p(t!).$$

Then:

- (a) *We have $d(t) \geq 0$ for all $t \geq 0$.*
- (b) *We have $d(pt) = d(t)$ for all $t \geq 0$.*
- (c) *We have $d(t) \geq 1$ for all $t > 0$ satisfying $p-1 \mid t$.*
- (d) *For any $t \geq 0$, the function $n \mapsto p^{d(t)} n^t$ is average-integral.*

Proof. For $t > 0$, Legendre's formula (6) gives

$$v_p(t!) = \sum_{i \geq 1} \left\lfloor \frac{t}{p^i} \right\rfloor < \sum_{i \geq 1} \frac{t}{p^i} = \frac{t}{p-1}.$$

This implies part (c) (since $\lfloor \frac{t}{p-1} \rfloor = \frac{t}{p-1}$ when $p-1 \mid t$) and also part (a) (since $v_p(t!)$ is an integer, so that $v_p(t!) < \frac{t}{p-1}$ entails $v_p(t!) \leq \lfloor \frac{t}{p-1} \rfloor$). Also $d(0) = 0$. The same formula (6) yields

$$v_p((pt)!) = t + v_p(t!), \quad \left\lfloor \frac{pt}{p-1} \right\rfloor = \left\lfloor t + \frac{t}{p-1} \right\rfloor = t + \left\lfloor \frac{t}{p-1} \right\rfloor,$$

and hence $d(pt) = d(t)$ for all $t \geq 0$; thus, part (b) is proved.

It remains to prove part (d). First, we shall work with the polynomials $p^{d(t)}x^t$ for $t \geq 0$; then we will evaluate them at integers.

Define an R -submodule M_p of $R[x]$ by

$$M_p = \left\{ \sum_{t=0}^N c_t x^t : N \geq 0, c_t \in p^{d(t)}R \text{ for all } 0 \leq t \leq N \right\}. \quad (27)$$

In words, a polynomial belongs to M_p when its x^t -coefficient for each t contains at least $d(t)$ factors of p . Equivalently, the elements of M_p are the finite R -linear combinations of $p^{d(t)}x^t$. We will prove that averaging (i.e., the operator F_p on polynomials) preserves M_p . It suffices to check that $F_p(p^{d(t)}x^t) \in M_p$ for all $t \geq 0$, since averaging is linear.

For $t = 0$ the assertion is immediate: averaging fixes constants. For $t > 0$, start with the definition of F_p . In each summand, expand $(x + r)^t$ by the binomial theorem. This gives

$$\begin{aligned} F_p(p^{d(t)}x^t) &= \frac{p^{d(t)}}{p} \sum_{r=0}^{p-1} (x + r)^t \\ &= p^{d(t)-1} \sum_{r=0}^{p-1} \sum_{j=0}^t \binom{t}{j} x^j r^{t-j} \\ &= \sum_{j=0}^t p^{d(t)-1} \binom{t}{j} \left(\sum_{r=0}^{p-1} r^{t-j} \right) x^j. \end{aligned}$$

Write $S_u = \sum_{r=0}^{p-1} r^u$. The term with $j = t$ has $S_0 = p$ and $\binom{t}{t} = 1$, so it is $p^{d(t)}x^t$. Separating this term, we obtain

$$F_p(p^{d(t)}x^t) = p^{d(t)}x^t + \sum_{j=0}^{t-1} p^{d(t)-1} \binom{t}{j} S_{t-j} x^j.$$

The leading term of this polynomial is $p^{d(t)}x^t$, which is already in M_p . Fix a lower degree $j < t$ and write $u = t - j > 0$. Then, the coefficient of x^j in $F_p(p^{d(t)}x^t)$ is

$$C_j = p^{d(t)-1} \binom{t}{j} S_u. \quad (28)$$

We must show $C_j \in p^{d(j)}R$, that is, $v_p(C_j) \geq d(j)$. Two estimates give this.

First, the binomial coefficient accounts for the factorials in d . Since $t = j + u$, the factorial formula for $\binom{t}{j}$ gives

$$v_p \binom{t}{j} = v_p(t!) - v_p(j!) - v_p(u!).$$

Adding $d(t) = \lfloor t/(p-1) \rfloor - v_p(t!)$ cancels $v_p(t!)$, leaving the exact equality

$$d(t) + v_p \binom{t}{j} = \left\lfloor \frac{t}{p-1} \right\rfloor - v_p(j!) - v_p(u!).$$

Since $t = j + u$, we also have

$$\left\lfloor \frac{t}{p-1} \right\rfloor = \left\lfloor \frac{j+u}{p-1} \right\rfloor \geq \left\lfloor \frac{j}{p-1} \right\rfloor + \left\lfloor \frac{u}{p-1} \right\rfloor,$$

since all $x, y \in \mathbb{R}$ satisfy $\lfloor x + y \rfloor \geq \lfloor x \rfloor + \lfloor y \rfloor$. Subtracting $v_p(j!) + v_p(u!)$ from both sides yields

$$d(t) + v_p \binom{t}{j} \geq d(j) + d(u). \quad (29)$$

Second, we claim that

$$d(u) + v_p(S_u) \geq 1. \quad (30)$$

Indeed, if $p-1$ divides u , then we already proved $d(u) \geq 1$ in Lemma 4.1(c) (since $u > 0$), and so (30) follows. If $p-1$ does not divide u , then S_u is divisible by p (an easy consequence of (18)³), thus $v_p(S_u) \geq 1$, so that (30) follows again. In either case, (30) is proved.

Now, using (28), (29) and (30) in sequence, we find

$$v_p(C_j) = d(t) - 1 + v_p \binom{t}{j} + v_p(S_u) \geq d(j) + d(u) - 1 + v_p(S_u) \geq d(j),$$

so $C_j \in p^{d(j)}R$. That is, the x^j -coefficient of $F_p(p^{d(t)}x^t)$ is divisible by $p^{d(j)}$ in R . Since this holds for all j , we thus have $F_p(p^{d(t)}x^t) \in M_p$. Since this holds for all $t \geq 0$, we thus obtain $F_p M_p \subseteq M_p$.

Now, fix $t \geq 0$ and let $f : \mathbb{Z} \rightarrow \mathbb{Z}$ be the function given by $f(n) = p^{d(t)}n^t$. We must show that f is average-integral, i.e., that $F_p^k f$ is integral for every $k \geq 0$. Since $p^{d(t)}x^t \in M_p$ and $F_p M_p \subseteq M_p$, induction on k shows that

$$F_p^k(p^{d(t)}x^t) \in M_p \subseteq R[x] \quad \text{for every } k \geq 0.$$

Compatibility with evaluation, (7), therefore gives

$$(F_p^k f)(n) \in R \quad \text{for all } k \geq 0 \text{ and } n \in \mathbb{Z} \quad (31)$$

(since $f = \text{ev}(p^{d(t)}x^t)$ and thus $F_p^k f = \text{ev}(F_p^k(p^{d(t)}x^t))$). On the other hand, f is integral and each application of F_p adds p values and divides their sum by p . Inductively, this shows that $p^k(F_p^k f)(n)$ is an integer. Thus $(F_p^k f)(n)$ belongs to $\mathbb{Z}[1/p]$, the ring of rational numbers whose reduced denominators are powers of p . Combining this with (31), we see that $(F_p^k f)(n) \in R \cap \mathbb{Z}[1/p] = \mathbb{Z}$ (indeed, a rational number in $\mathbb{Z}[1/p]$ can have only p in its denominator, whereas a rational number in R cannot have p there; thus, a rational number in $R \cap \mathbb{Z}[1/p]$ must have denominator 1). That is, $(F_p^k f)(n)$ is an integer for every $k \geq 0$ and $n \in \mathbb{Z}$. This proves that f is average-integral. Thus, part (d) is proved. $\square$

³In more detail: Assume that $p-1$ does not divide u . Then, u has a nonzero remainder w upon division by $p-1$. Thus, $0 < w < p-1$ and $u \equiv w \pmod{p-1}$. Hence, Fermat's Little Theorem yields the congruence $r^u \equiv r^w \pmod{p}$ for each integer r not divisible by p ; but the same congruence holds for all integers r divisible by p because u and w are positive. Summing this congruence over all $r \in \{0, 1, \dots, p-1\}$, we find $S_u \equiv S_w \pmod{p}$. But (18) shows that the sum S_w is divisible by p , since $0 < w < p-1$; thus, $S_u \equiv S_w \equiv 0 \pmod{p}$.

Let us next record the analogue of Lemma 1.4 for R -valued functions. As before, p is our fixed prime and $q = p - 1$.

Lemma 4.2 (Difference criterion over R). *For a function $f: \mathbb{Z} \rightarrow \mathbb{Q}$ and a nonnegative integer ℓ , the following conditions are equivalent:*

1. *We have $F_p^k f \in R^{\mathbb{Z}}$ for all $0 \leq k \leq \ell$.*
2. *We have $\Delta^{kq} f \in p^k R^{\mathbb{Z}}$ for all $0 \leq k \leq \ell$.*
3. *We have $\Delta^j f \in p^{\lfloor j/q \rfloor} R^{\mathbb{Z}}$ for all $0 \leq j \leq \ell q$.*

Proof. This is an analogue of Lemma 1.4 with some $\mathbb{Z}$'s replaced by R 's; the exact same proof applies, mutatis mutandis. (The only thing we use about R is that R is a subring of $\mathbb{Q}$.) $\square$

We can now describe all the polynomials P for which every iterated average $F_p^k P$ takes values in R at all integers. The description has several useful forms: one bounds the ordinary monomial coefficients, and the other bounds the coefficients in the binomial (or, equivalently, falling-factorial) basis. We will use the R -submodule

$$M_p = \left\{ \sum_{t=0}^N c_t x^t : N \geq 0, c_t \in p^{d(t)} R \text{ for all } 0 \leq t \leq N \right\} \quad (32)$$

of $R[x]$ (which already appeared in the proof of Lemma 4.1).

We recall that each polynomial $P(x) \in \mathbb{Q}[x]$ can be written as a $\mathbb{Q}$ -linear combination $\sum_j b_j(x)_j$ of falling factorials. Since $(x)_j = j! \binom{x}{j}$, we can rewrite this as $\sum_j b_j j! \binom{x}{j}$; so every polynomial $P(x) \in \mathbb{Q}[x]$ can be written as a $\mathbb{Q}$ -linear combination $\sum_j a_j \binom{x}{j}$ of binomial coefficients. Again, the coefficients a_j here are unique.

Proposition 4.3 (Polynomial criterion). *For $P \in \mathbb{Q}[x]$, the following conditions are equivalent:*

1. *We have $F_p^k P(n) \in R$ for every $k \geq 0$ and $n \in \mathbb{Z}$.*
2. *Writing $P(x) = \sum_j c_j x^j$, we have $c_j \in p^{d(j)} R$ for every j ; in other words, $P \in M_p$.*
3. *Writing $P(x) = \sum_j b_j(x)_j$, we have $b_j \in p^{d(j)} R$ for every j .*
4. *Writing $P(x) = \sum_j a_j \binom{x}{j}$, we have $a_j \in p^{\lfloor j/(p-1) \rfloor} R$ for every j .*
5. *We have $\Delta^{r(p-1)} P(n) \in p^r R$ for each $n \in \mathbb{Z}$ and $r \geq 0$.*
6. *We have $\Delta^j P(n) \in p^{\lfloor j/(p-1) \rfloor} R$ for each $n \in \mathbb{Z}$ and $j \geq 0$.*

If P is integer-valued on $\mathbb{Z}$, these conditions are also equivalent to average-integrality of the function $n \mapsto P(n)$.

Proof. We first compare the conditions involving values and differences. Put $f(n) = P(n)$ for all $n \in \mathbb{Z}$. Apply Lemma 4.2 to f , using the compatibility with evaluation in (7). Allowing ℓ to range over all nonnegative integers proves $(1) \Leftrightarrow (5) \Leftrightarrow (6)$.

Next we prove $(6) \Rightarrow (4) \Rightarrow (5)$. Since $\binom{x}{j} = (x)_j / j!$, the expansion $P = \sum_j a_j \binom{x}{j}$ can also be written as

$$P = \sum_j \frac{a_j}{j!} (x)_j.$$

Thus (9) gives $\Delta^j P(0) = j!(a_j/j!) = a_j$. Applying condition (6) at $n = 0$, we thus obtain $a_j \in p^{\lfloor j/(p-1) \rfloor} R$, which is (4).

To prove $(4) \Rightarrow (5)$, divide (8) by $j!$ to obtain the normalized difference identity

$$\Delta \binom{x}{j} = \binom{x}{j-1} \quad \text{for all } j \geq 1, \quad \Delta \binom{x}{0} = 0.$$

So each iteration of Δ , when applied to a binomial-coefficient polynomial $\binom{x}{j}$, lowers the binomial index j by 1, until the term vanishes. So $\Delta^k \binom{x}{j} = \binom{x}{j-k}$ for all $j, k \geq 0$, where we interpret the right-hand side as 0 if $j - k < 0$. Thus, under condition (4), for every $r \geq 0$ we have

$$\Delta^{r(p-1)} P(x) = \sum_{j \geq r(p-1)} a_j \binom{x}{j - r(p-1)}. \quad (33)$$

For each index in this sum, $j \geq r(p-1)$, so $\lfloor j/(p-1) \rfloor \geq r$. Condition (4) therefore ensures $a_j \in p^r R$. Also, the binomial polynomials $\binom{x}{k}$ take integer values on all of $\mathbb{Z}$ (as we have already seen). Consequently, (33) shows that $\Delta^{r(p-1)} P$ takes values in $p^r R$ for every r , which is exactly condition (5). This completes the equivalence of conditions (1), (4), (5), and (6).

The equivalence of (3) and (4) is also easy, since the respective coefficients are related by $a_j = j!b_j$ (because $(x)_j = j! \binom{x}{j}$), and thus we have $a_j \in p^{\lfloor j/(p-1) \rfloor} R$ if and only if $b_j \in p^{d(j)} R$.

Next we prove $(2) \Rightarrow (1)$. Under (2), write

$$P(x) = \sum_{j=0}^N u_j p^{d(j)} x^j, \quad \text{with } u_j \in R.$$

For each j , put $f_j(n) = p^{d(j)} n^j$. Lemma 4.1(d) says that f_j is average-integral, so that $F_p^k f_j$ is integral for every $k \geq 0$. Since $\text{ev } P = \sum_{j=0}^N u_j f_j$ and F_p^k is $\mathbb{Q}$ -linear (and by (7)),

$$(F_p^k P)(n) = \sum_{j=0}^N \underbrace{u_j}_{\in R} \underbrace{(F_p^k f_j)(n)}_{\in \mathbb{Z}} \in R \quad \text{for all } k \geq 0 \text{ and } n \in \mathbb{Z}.$$

This is condition (1).

For the converse, we take a bird's-eye view. Do not fix P . As we recall, M_p is the set of all polynomials $P \in \mathbb{Q}[x]$ satisfying condition (2). Let M'_p be the set of all polynomials $P \in \mathbb{Q}[x]$ satisfying condition (1). Both M_p and M'_p are R -submodules of $\mathbb{Q}[x]$ (since each F_p^k is R -linear). Since we have proved $(2) \Rightarrow (1)$, we know that $M_p \subseteq M'_p$. We shall now show that $M'_p \subseteq M_p$.

Indeed, let $P \in M'_p$. We must show that $P \in M_p$. We shall show this by strong induction on $\deg P$; the base case $P = 0$ is obvious. Suppose $P \neq 0$ and let $c_t x^t$ be its leading term. The polynomial P satisfies condition (1) (since $P \in M'_p$), and thus also condition (6) (since (1) and (6) are equivalent), which at $j = t$ and $n = 0$ gives

$$\Delta^t P(0) \in p^{\lfloor t/(p-1) \rfloor} R.$$

Since $(x)_t$ is monic of degree t , the polynomial $P - c_t(x)_t$ has degree less than t , and thus is a linear combination of $(x)_0, (x)_1, \dots, (x)_{t-1}$. Thus the $(x)_t$ -coefficient of P is c_t , and so (9) gives $\Delta^t P(0) = t!c_t$. Therefore,

$$t!c_t = \Delta^t P(0) \in p^{\lfloor t/(p-1) \rfloor} R, \quad \text{hence} \quad c_t \in p^{d(t)} R.$$

Therefore, $c_t x^t \in M_p \subseteq M'_p$. Hence, $P - c_t x^t \in M'_p$ as well (because P and $c_t x^t$ belong to M'_p , and since M'_p is an R -module). By the induction hypothesis, this entails $P - c_t x^t \in M_p$, since $P - c_t x^t$ has smaller degree than P . Adding this to $c_t x^t \in M_p$, we find $P \in M_p$, since M_p is an R -module. So the induction is complete, and $M'_p \subseteq M_p$ is proved; that is, condition (1) implies condition (2).

Finally, suppose P is integer-valued, so f is integral. For every $k \geq 0$ and $n \in \mathbb{Z}$, the number $p^k (F_p^k f)(n)$ is an integer, so $(F_p^k f)(n)$ is a rational number whose denominator is a power of p . Under condition (1), the number $(F_p^k f)(n)$ also belongs to R , so the denominator argument at the end of the proof of Lemma 4.1(d) shows that it is an integer. Hence f is average-integral. Conversely, if f is average-integral, then $(F_p^k f)(n) \in \mathbb{Z} \subseteq R$ for every k, n , which is exactly condition (1). $\square$

For example, when $p = 3$ we have $d(0) = d(1) = 0$, $d(2) = 1$, and $d(3) = 0$. Thus a polynomial of degree at most 3 in $R[x]$ has all its averages valued in R exactly when its x^2 -coefficient is divisible by 3. In particular, x , $3x^2$, and x^3 are average-integral on $\mathbb{Z}$, but x^2 is not. Indeed,

$$F_3(x^2) = x^2 + 2x + \frac{5}{3}.$$

This example illustrates why the coefficient conditions are useful: we can decide about all averaging stages without calculating them.

The criterion also lets us return to arbitrary functions without losing control at large depths. Although we match the function only on a finite interval, the interpolating polynomial will have integral averages at *every* depth. This useful fact is the content of the following lemma.

Lemma 4.4 (Local lifting). *Let $f \in \mathbb{Q}^{\mathbb{Z}}$ be ℓ -average-integral, let $n \in \mathbb{Z}$, and put $m = \ell(p-1)$. There is an integer-valued polynomial $P \in M_p$ such that*

$$P(r) = f(n+r) \quad \text{for all } 0 \leq r \leq m.$$

In fact, the Newton interpolant

$$P(x) = \sum_{j=0}^m \Delta^j f(n) \binom{x}{j} \tag{34}$$

has these properties, and the function $r \mapsto P(r)$ is average-integral (at every depth).

Proof. Define P by (34). The matching values $P(r) = f(n+r)$ follow from (23). Also, each coefficient $\Delta^j f(n)$ is an integer (since f is integral), and the binomial polynomials $\binom{x}{j}$ take integer values on all of $\mathbb{Z}$, as noted above. Hence P is integer-valued on all of $\mathbb{Z}$, not just on the interval where it matches f .

For each $j \leq m$, Lemma 1.4 yields $\Delta^j f \in p^{\lfloor j/(p-1) \rfloor} \mathbb{Z}^\mathbb{Z}$, so that the integer $\Delta^j f(n)$ is divisible by $p^{\lfloor j/(p-1) \rfloor}$; in particular, $\Delta^j f(n) \in p^{\lfloor j/(p-1) \rfloor} R$. This accounts for all the addends in (34). Thus, the integer-valued polynomial $P \in \mathbb{Q}[x]$ meets condition (4) of Proposition 4.3 (with $a_j = \Delta^j f(n)$ for all $j \leq m$, and $a_j = 0$ for all $j > m$). Therefore, that proposition gives $P \in M_p$ and average-integrality of the function $r \mapsto P(r)$. $\square$

5 Every depth for $p = 2$ and $p = 3$

We now have the ingredients to prove Theorem 1.2. The key distinction between $p = 2$ and $p = 3$ will be visible in the coefficients of a product and a cube.

Proof of Theorem 1.2. We first work with polynomials. Proposition 4.3 reduces the question to coefficient divisibility: we must show that M_2 is closed under products and that M_3 is closed under cubes. Afterwards, local interpolation will transfer the result to arbitrary functions.

For $p = 2$, the definition of the function d (see Lemma 4.1) simplifies to $d(j) = j - v_2(j!)$; therefore

$$d(i+j) = d(i) + d(j) - v_2 \binom{i+j}{i} \leq d(i) + d(j) \quad \text{for all } i, j \geq 0.$$

The product of the two basic polynomials $2^{d(i)}x^i$ and $2^{d(j)}x^j$ has coefficient $2^{d(i)+d(j)}$ at degree $i+j$. The inequality we just proved says that this contains at least the required $d(i+j)$ factors of 2. The product therefore belongs to M_2 . More explicitly, if $P = \sum_i a_i 2^{d(i)}x^i$ and $Q = \sum_j b_j 2^{d(j)}x^j$, with $a_i, b_j \in R$, then

$$PQ = \sum_{i,j} a_i b_j \left(2^{d(i)}x^i \right) \left(2^{d(j)}x^j \right).$$

Each product in parentheses belongs to M_2 by the preceding calculation. Since M_2 is closed under finite R -linear combinations, $PQ \in M_2$. Closure under addition and additive inverses follows from the definition, and the constant polynomial 1 belongs to M_2 because $d(0) = 0$. Thus M_2 is a subring of $\mathbb{Q}[x]$.

For $p = 3$, products need not stay in M_3 . However, when we cube a sum, every mixed term has a coefficient divisible by 3. We will see that this extra factor is enough.

Here we have $d(j) = \lfloor j/2 \rfloor - v_3(j!)$. Thus, for all $i, j, k \geq 0$, we have

$$d(i+j+k) - d(i) - d(j) - d(k) = \left\lfloor \frac{i+j+k}{2} \right\rfloor - \left\lfloor \frac{i}{2} \right\rfloor - \left\lfloor \frac{j}{2} \right\rfloor - \left\lfloor \frac{k}{2} \right\rfloor - v_3 \binom{i+j+k}{i, j, k}.$$

Here the multinomial coefficient is an integer, and the difference of floors is at most 1. Indeed, write $i = 2a + \epsilon$, $j = 2b + \eta$, and $k = 2c + \theta$, where each of ϵ, η, θ is 0 or 1. The difference of

floors then becomes $\lfloor (\epsilon + \eta + \theta)/2 \rfloor \leq 1$. Also $v_3\binom{i+j+k}{i,j,k} \geq 0$, since this multinomial coefficient is an integer. Consequently,

$$d(i+j+k) \leq d(i) + d(j) + d(k) + 1 \quad \text{for all } i, j, k \geq 0. \quad (35)$$

Write $P = \sum_j c_j x^j \in M_3$, so that $v_3(c_j) \geq d(j)$ for every j . Multiplying three copies of P gives $P^3 = \sum_{i,j,k} c_i c_j c_k x^{i+j+k}$. Group the ordered triples of indices according to which indices coincide. There is one ordering when all three are equal, three when exactly two are equal, and six when all are distinct. Hence

$$P^3 = \sum_j c_j^3 x^{3j} + 3 \sum_{i \neq j} c_i^2 c_j x^{2i+j} + 6 \sum_{i < j < k} c_i c_j c_k x^{i+j+k}.$$

All sums are finite. First consider the pure terms $c_j^3 x^{3j}$. The 3-valuations $v_3(c_j^3) = 3v_3(c_j)$ of their coefficients are at least $3d(j)$ (since $v_3(c_j) \geq d(j)$), whereas membership in M_3 requires only $d(3j) = d(j)$, by Lemma 4.1(b). Since $d(j) \geq 0$, we have $3d(j) \geq d(j)$. These terms therefore belong to M_3 .

Each mixed term is $3c_i^2 c_j x^{2i+j}$ for some $i \neq j$, or $6c_i c_j c_k x^{i+j+k}$ for some $i < j < k$. In the first case the 3-valuation of its coefficient is at least $1 + 2d(i) + d(j) \geq d(2i+j)$, by (35) with indices i, i, j . In the second case it is at least $1 + d(i) + d(j) + d(k) \geq d(i+j+k)$, again by (35). Thus every mixed term also belongs to M_3 , and so does their sum with the pure terms. This proves $P^3 \in M_3$.

We now transfer these polynomial statements to functions. Fix a depth ℓ and an integer n . In the $p = 2$ case, let $f, g \in \mathbb{Q}^{\mathbb{Z}}$ be ℓ -average-integral. Lemma 4.4 (applied here with $m = \ell$, since $p - 1 = 1$) supplies integer-valued polynomials $P, Q \in M_2$ with

$$P(s) = f(n+s), \quad Q(s) = g(n+s) \quad \text{for all } 0 \leq s \leq \ell. \quad (36)$$

Their product PQ belongs to M_2 (since M_2 is a ring) and is integer-valued. Put $h(s) = P(s)Q(s)$ for all $s \in \mathbb{Z}$. Proposition 4.3 (specifically, its last sentence) gives $F_2^r h \in \mathbb{Z}^{\mathbb{Z}}$ for every $r \geq 0$, while (36) shows that

$$h(s) = (fg)(n+s) \quad \text{for all } 0 \leq s \leq \ell. \quad (37)$$

Consider an arbitrary function $w : \mathbb{Z} \rightarrow \mathbb{Q}$. Any value of $F_p^r w$ is determined by only $r(p-1) + 1$ consecutive values of w . This can be seen by writing each application of F_p as a sum of shifts. For $r \geq 0$, the equality (1) gives

$$F_p^r = \frac{1}{p^r} (\text{id} + E + \cdots + E^{p-1})^r = \frac{1}{p^r} \sum_{a_1, \dots, a_r \in \{0, 1, \dots, p-1\}} E^{a_1 + \cdots + a_r}$$

by distributivity (where for $r = 0$, the sum consists of a single addend equal to E^0 , corresponding to the empty 0-tuple). Here a_i records which shift was chosen from the i th factor. Applying this identity to a function w and evaluating at n gives

$$(F_p^r w)(n) = \frac{1}{p^r} \sum_{a_1, \dots, a_r \in \{0, 1, \dots, p-1\}} w(n + a_1 + \cdots + a_r) \quad (38)$$

for all $r \geq 0$. In the $p = 2$ case, all the inner sums $a_1 + \cdots + a_r$ lie between 0 and r . Thus, for $0 \leq r \leq \ell$, if we compute $(F_2^r(fg))(n)$ and $(F_2^r h)(0)$ using (38), then each summand in the formula for $(F_2^r(fg))(n)$ equals the corresponding summand for $(F_2^r h)(0)$, by (37). Consequently,

$$(F_2^r(fg))(n) = (F_2^r h)(0) \in \mathbb{Z} \quad \left(\text{since } F_2^r h \in \mathbb{Z}^{\mathbb{Z}} \right).$$

Since n was arbitrary, this shows that $F_2^r(fg) \in \mathbb{Z}^{\mathbb{Z}}$. Since $r \leq \ell$ was arbitrary, this means that fg is ℓ -average-integral. So the set of all ℓ -average-integral functions is closed under multiplication. Linearity of F_2 also gives closure under sums and additive inverses, and the constant function 1 is ℓ -average-integral because F_2 fixes it. Hence, the set of all ℓ -average-integral functions for $p = 2$ is a ring under pointwise addition and multiplication.

For $p = 3$, let us again fix an ℓ -average-integral function $f \in \mathbb{Q}^{\mathbb{Z}}$, and apply the same argument to an integer-valued polynomial $P \in M_3$ that satisfies

$$P(s) = f(n + s) \quad \text{for all } 0 \leq s \leq 2\ell. \quad (39)$$

(Such a P again exists by Lemma 4.4.) As we saw above, $P \in M_3$ entails $P^3 \in M_3$. Put $h(s) = P(s)^3$. Since P^3 is integer-valued and belongs to M_3 , Proposition 4.3 (again, its last sentence, applied to P^3) shows that the function h is average-integral, i.e., we have $F_3^r h \in \mathbb{Z}^{\mathbb{Z}}$ for all $r \geq 0$. For $0 \leq r \leq \ell$, the sum $a_1 + \cdots + a_r$ in (38) lies between 0 and $2r \leq 2\ell$. At all these arguments, we have $h(s) = P(s)^3 = f(n + s)^3$ by (39), and hence

$$(F_3^r(f^3))(n) = (F_3^r h)(0) \in \mathbb{Z}.$$

Since n was arbitrary, this shows that $F_3^r(f^3) \in \mathbb{Z}^{\mathbb{Z}}$, and therefore f^3 is ℓ -average-integral. $\square$

Proof of Corollary 1.3. Suppose f, g are average-integral for $p = 2$. For any $k \geq 0$, apply the $p = 2$ case of Theorem 1.2 with $\ell = k$ to see that fg is k -average-integral; thus, $F_2^k(fg) \in \mathbb{Z}^{\mathbb{Z}}$. This shows that fg is average-integral.

Likewise, if f is average-integral for $p = 3$, then Theorem 1.2 for $p = 3$ and $\ell = k$ gives $F_3^k(f^3) \in \mathbb{Z}^{\mathbb{Z}}$ for every $k \geq 0$. Thus f^3 is average-integral. $\square$

Remark 5.1. The polynomial strategy for $p = 2, 3$ was already sketched by Gjergji Zaimi in his June 16, 2011 answer to [2]. The proof above supplies the full coefficient estimates and the local lifting argument. The distinction between products and cubes is necessary: for $p = 3$, $f(n) = n$ is average-integral, since $F_3^r f(n) = n + r$, whereas

$$(F_3(f^2))(n) = n^2 + 2n + \frac{5}{3}$$

is never integral.

6 Sharpness for primes greater than three

We now verify zeb's counterexample from [3] and use Theorem 1.1 to identify its first failing depth.

Proposition 6.1 (zeb's example, with exact failing depth). *Let $p > 3$ be prime and put $f(n) = 1 + n^{p+1}$. Then f is average-integral, while f^p is p -average-integral but is not $(p+1)$ -average-integral. In particular, $F_p^{p+1}(f^p)$ is the first nonintegral iterate.*

For a first look at the obstruction, take $p = 5$. The relevant expansion is

$$(1 + x^6)^5 = 1 + 5x^6 + 10x^{12} + 10x^{18} + 5x^{24} + x^{30}.$$

The polynomial criterion requires two factors of 5 at degree 24, since $d(24) = 6 - 4 = 2$, but the coefficient is only 5. All other terms satisfy the criterion. This already shows that some iterate fails. The proof below also determines the first failing iterate, by using a falling factorial of degree 24.

Proof. Since $p > 3$, Legendre's formula gives $v_p((p+1)!) = 1$ and $\lfloor (p+1)/(p-1) \rfloor = 1$. Thus $d(p+1) = 0$, so Lemma 4.1(d) shows that $n \mapsto n^{p+1}$ is average-integral. Since F_p fixes the constant function 1 and is linear, $F_p^k f = 1 + F_p^k(n \mapsto n^{p+1})$ is integral for every $k \geq 0$. Thus f is average-integral. We now examine f^p . The idea is to separate its polynomial into terms that are already known to be average-integral and one term that fails the difference test. The binomial theorem first gives

$$(1 + x^{p+1})^p = \sum_{r=0}^p \binom{p}{r} (x^{p+1})^r = \sum_{r=0}^p \binom{p}{r} x^{r(p+1)}.$$

The terms with $r = 0$ and $r = p$ are 1 and x^{p^2+p} . The term with $r = p-1$ is $px^{(p-1)(p+1)}$. Set $N = p^2 - 1 = (p-1)(p+1)$ and separate these three terms to obtain

$$(1 + x^{p+1})^p = 1 + \sum_{r=1}^{p-2} \binom{p}{r} x^{r(p+1)} + px^N + x^{p^2+p}.$$

For $1 \leq r \leq p-2$, we have $r(p+1) < p^2$, so

$$v_p((r(p+1))!) = r, \quad d(r(p+1)) = \left\lfloor \frac{2r}{p-1} \right\rfloor \leq 1.$$

Also $v_p((p^2+p)!) = p+2 = \lfloor (p^2+p)/(p-1) \rfloor$, and hence $d(p^2+p) = 0$. For $1 \leq r \leq p-2$, the coefficient $\binom{p}{r}$ is divisible by p , which supplies the at most one factor demanded by $d(r(p+1)) \leq 1$. The last term needs no additional factor, since $d(p^2+p) = 0$. By Lemma 4.1(d), all the displayed summands except possibly px^N are therefore average-integral.

To test the remaining term px^N , we replace x^N by the falling factorial $(x)_N$, whose N th difference is especially simple. We must check that the lower-degree terms introduced by this replacement are harmless.

For $0 \leq j < N$, write $j = ap + b$ with $0 \leq b < p$. Then $0 \leq a \leq p-1$ and $a+b \leq 2p-3$: equality $a = b = p-1$ is excluded by $j < N$. Legendre's formula gives $v_p(j!) = a$, whence

$$d(j) = \left\lfloor \frac{ap+b}{p-1} \right\rfloor - a = \left\lfloor \frac{a+b}{p-1} \right\rfloor \leq 1.$$

Since $x^N - (x)_N$ is an integer polynomial of degree less than N , every coefficient of $p(x^N - (x)_N)$ is divisible by p . The inequalities $d(j) \leq 1$ show that each of its monomials satisfies the polynomial criterion. Thus $p(x^N - (x)_N)$ is average-integral. We conclude that

$$(1 + x^{p+1})^p = p(x)_N + A(x), \quad (40)$$

where A defines an average-integral function $a(n) = A(n)$. Put $h(n) = p(n)_N$. Evaluating the identity (40) at integers gives $f^p = h + a$, so linearity gives

$$F_p^k h = F_p^k(f^p) - F_p^k a \quad \text{for all } k \geq 0.$$

Since $F_p^k a$ is integral for every k , if f^p were $(p+1)$ -average-integral, then h would be $(p+1)$ -average-integral as well. Lemma 1.4, with $N = (p+1)(p-1)$, would force $\Delta^N h$ to belong to $p^{\lfloor N/(p-1) \rfloor} \mathbb{Z}^{\mathbb{Z}} = p^{p+1} \mathbb{Z}^{\mathbb{Z}}$, so that all values of the polynomial $\Delta^N(p(x)_N)$ at integers would be multiples of p^{p+1} . In other words, all values of the polynomial $\Delta^N(x)_N$ would be multiples of p^p . But this is false, since iterating (8) N times shows that $\Delta^N(x)_N$ is the constant polynomial $N!$, which is not a multiple of p^p (indeed, Legendre's formula shows that $v_p(N!) = p-1$). This contradiction shows that f^p cannot be $(p+1)$ -average-integral. Finally, Theorem 1.1 shows that f^p is p -average-integral, so failure first occurs at iteration $p+1$. $\square$

Remark 6.2. The proof of Lemma 2.3 also helps explain why the depth $p+1$ is different from the depths $1, \dots, p$. There we wrote $P(x)^p = \sum_j c_j(x)_j$ and used $\Delta^m(P^p)(0) = m!c_m$. For $m = k(p-1)$ with $1 \leq k \leq p$, the factorial $m!$ supplies exactly $k-1$ factors of p . The coefficient argument modulo p supplies the remaining factor by showing that $c_m \in pR$. At $k = p+1$, however, we have $m = p^2 - 1$ and still $v_p(m!) = p-1$. To obtain $\Delta^m(P^p)(0) \in p^{p+1}R$, we would now need $c_m \in p^2R$, rather than just $c_m \in pR$. Thus the argument modulo p no longer suffices. Proposition 6.1 shows that, for every $p > 3$, this stronger divisibility can actually fail.

7 Non-primes

The assumption that the averaging length p is prime is essential for Theorem 1.1. Here are some examples of what can go wrong when we average over composite-length windows.

For an integer $m \geq 2$, define an operator $F_m : \mathbb{Q}^{\mathbb{Z}} \rightarrow \mathbb{Q}^{\mathbb{Z}}$ by

$$F_m f(n) = \frac{f(n) + f(n+1) + \dots + f(n+m-1)}{m}$$

(generalizing the F_p we have studied above). An m -periodic function f exhibits very simple behavior under averaging: every block of m consecutive values has the same sum, so $F_m f$ is constant. If that sum is zero, then $F_m f = 0$, and hence $F_m^k f = 0$ for every $k \geq 1$.

Example 7.1. For $m = 4$, let f be the 4-periodic function whose values in one period are $1, -1, 0, 0$. Then $F_4 f = 0$, whereas

$$F_4(f^4) = \frac{1 + 1 + 0 + 0}{4} = \frac{1}{2}.$$

Thus, even when $F_4^k f$ is integral for every $k \geq 0$, its fourth power can fail at the very first averaging step. The same example, with $m - 2$ zeros in place of the two zeros, works for every even $m > 2$: it gives $F_m f = 0$ and $F_m(f^m) = 2/m$.

There is a similar obstruction whenever m is divisible by the square of a prime:

Example 7.2. Suppose that r is prime and $r^2 \mid m$. Take an m -periodic function f with r entries equal to 1, one entry equal to $-r$, and all remaining entries equal to zero in each period. There is room for these entries because $r + 1 \leq r^2 \leq m$. Their sum is zero, so again $F_m^k f = 0$ for every $k \geq 1$. On the other hand,

$$F_m(f^m) = \frac{r + (-r)^m}{m}.$$

The numerator is congruent to r modulo r^2 , and therefore is not divisible by m . Hence f^m fails to be integer-valued after one averaging step. In particular, this rules out an analogue of the power-preservation theorem for every composite prime power.

Odd squarefree lengths can fail as well:

Example 7.3. For $m = 15$, take the 15-periodic function with values 1, 1, -2 followed by twelve zeros in each period. Again the sum is zero, so $F_{15}^k f = 0$ for every $k \geq 1$. But

$$F_{15}(f^{15}) = \frac{1 + 1 + (-2)^{15}}{15} = -\frac{10922}{5} \notin \mathbb{Z}.$$

Thus the power-preservation assertion fails at the first step even for $m = 3 \cdot 5$.

More generally, suppose $m \geq 3$ is odd and some integer a satisfies $a^m \not\equiv a \pmod{m}$. We may choose $0 \leq a \leq m - 1$, since the congruence depends only on the residue of a modulo m . Take an m -periodic function with a entries equal to 1, one entry equal to $-a$, and $m - a - 1$ entries equal to zero in each period. Its sum is zero, whereas the sum of its m th powers is $a - a^m$, which is not divisible by m . This gives the same kind of counterexample.

Conversely, if

$$a^m \equiv a \pmod{m} \quad \text{for all } a \in \mathbb{Z}$$

(that is, m is either prime or a Carmichael number), then every integer-valued function f satisfies

$$\sum_{r=0}^{m-1} f(n+r)^m \equiv \sum_{r=0}^{m-1} f(n+r) \pmod{m} \quad \text{for all } n \in \mathbb{Z}.$$

Consequently, integrality of $F_m f$ implies integrality of $F_m(f^m)$. Thus, for odd m , the displayed congruence for all a is exactly the condition for power preservation at the first step. This does not imply preservation at all averaging steps. We can, however, obtain integrality at some larger depths, provided we increase the exponent along with the depth.

Proposition 7.4. *Let $m \geq 2$ be an integer such that*

$$a^m \equiv a \pmod{m} \quad \text{for all } a \in \mathbb{Z}.$$

Let $f: \mathbb{Z} \rightarrow \mathbb{Z}$ satisfy $F_m f \in \mathbb{Z}^{\mathbb{Z}}$. Then, for every integer $t \geq 0$, we have

$$F_m^k(f^{m^t}) \in \mathbb{Z}^{\mathbb{Z}} \quad \text{for all } 0 \leq k \leq t + 1.$$

Proof. Subtracting two consecutive averages gives

$$f(n+m) - f(n) = m((F_m f)(n+1) - (F_m f)(n)).$$

Thus

$$f(n+m) \equiv f(n) \pmod{m} \quad \text{for every } n \in \mathbb{Z} \quad (41)$$

(since $F_m f \in \mathbb{Z}^{\mathbb{Z}}$).

We first check how raising to the m th power improves this congruence. Suppose a and b are two integers such that $a \equiv b \pmod{m^r}$, where $r \geq 1$, and write $a = b + m^r u$ with $u \in \mathbb{Z}$. Expanding $a^m = (b + m^r u)^m$ using the binomial theorem gives

$$a^m - b^m = m^{r+1} u b^{m-1} + \sum_{i=2}^m \binom{m}{i} m^{ri} u^i b^{m-i}.$$

Every term is divisible by m^{r+1} , since $ri \geq r+1$ when $i \geq 2$. Thus, $a^m \equiv b^m \pmod{m^{r+1}}$. Applying this observation t times yields

$$a^{m^t} \equiv b^{m^t} \pmod{m^{r+t}}.$$

Thus, from (41), we obtain

$$f(n+m)^{m^t} \equiv f(n)^{m^t} \pmod{m^{t+1}} \quad \text{for all } n \in \mathbb{Z}.$$

Let $b : \mathbb{Z} \rightarrow \mathbb{Z}$ be the m -periodic function agreeing with f^{m^t} at $0, 1, \dots, m-1$. The last congruence shows that

$$f^{m^t} = b + m^{t+1} c$$

for an integer-valued function c on $\mathbb{Z}$. Repeatedly using the hypothesis $a^m \equiv a \pmod{m}$ gives $a^{m^t} \equiv a \pmod{m}$ for all $a \in \mathbb{Z}$. Consequently,

$$\sum_{s=0}^{m-1} b(s) = \sum_{s=0}^{m-1} f(s)^{m^t} \equiv \sum_{s=0}^{m-1} f(s) = m(F_m f)(0) \equiv 0 \pmod{m}.$$

Since b is m -periodic, $F_m b$ is therefore an integer constant, and every further average of b is the same constant. On the other hand, $m^k F_m^k c$ is integer-valued: expanding the k averaging operations expresses it as a sum of integer values of c . Hence $m^{t+1} F_m^k c$ is integer-valued for $0 \leq k \leq t+1$. Applying F_m^k to $f^{m^t} = b + m^{t+1} c$ proves the assertion. $\square$

Example 7.5 (The averaging length 561). The composite integer $m = 561 = 3 \cdot 11 \cdot 17$ is known to be a Carmichael number; thus it satisfies the hypothesis of Proposition 7.4. Hence, integrality of $F_{561} f$ guarantees (by Proposition 7.4) that the first two averages of f^{561} , the first three averages of f^{561^2} , and, more generally, the first $t+1$ averages of f^{561^t} are integral.

Keeping the exponent equal to 561, however, does not give preservation at every depth. Take $f(n) = n$. Then

$$(F_{561}^k f)(n) = n + 280k \quad \text{for all } k \geq 0 \text{ and } n \in \mathbb{Z},$$

so every average of f is integer-valued. Nevertheless, a mechanical computation (see Appendix B) shows that

$$\sum_{a_1, \dots, a_5=0}^{560} (a_1 + \dots + a_5)^{561} \equiv 162 \pmod{243}. \quad (42)$$

The left-hand side is $561^5 (F_{561}^5(f^{561}))(0)$ (by (38)). Since $243 = 3^5$ divides 561^5 , the congruence shows that $(F_{561}^5(f^{561}))(0)$ is not an integer.

8 Connections with the literature

The averaging problem also belongs to a longer history of divisibility conditions on finite differences. We describe some connections here, including two results from which parts of our conclusions can be deduced. This section is independent of the proofs above. The paragraphs on p -adic functions provide additional context and may be skipped without affecting the rest of the note.

8.1 An exact antecedent for the binary product theorem

Sun [4, Definition 2.1] calls a function $u: \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{(p)}$ *p-regular* if

$$\sum_{r=0}^j (-1)^r \binom{j}{r} u(r) \in p^j \mathbb{Z}_{(p)} \quad \text{for all } j \geq 0.$$

These sums equal $(-1)^j \Delta^j u(0)$. His Product Theorem, [4, Theorem 2.3, p. 200], states that products of p -regular functions are p -regular. For $p = 2$, it gives our infinite-depth product assertion. To see the connection, let f, g be average-integral functions and fix $n \in \mathbb{Z}$. The functions $u(r) = f(n+r)$ and $v(r) = g(n+r)$ on the nonnegative integers are 2-regular by Lemma 1.4, since $q = 1$. Sun's theorem gives $2^j \mid \Delta^j(fg)(n)$ in $\mathbb{Z}_{(2)}$ for every j . These differences are ordinary integers, so the same divisibility holds in $\mathbb{Z}$. Since n was arbitrary, the difference criterion proves that fg is average-integral. For $p = 3$ our bound is $\lfloor j/2 \rfloor$, not j , so Sun's theorem does not cover our class.

There is a related formulation using p -adic integers. Recall that $\mathbb{Z}_p$ records compatible residues modulo $p, p^2, p^3, \dots$; in its notion of convergence, $a_j \rightarrow 0$ means that for each k , all sufficiently late terms are divisible by p^k . This ring is different from $R = \mathbb{Z}_{(p)}$, whose elements are rational numbers.

Kellner [5, Definition 3.1 and Theorem 6.1] studies the p -adic Kummer space $\mathcal{K}_{p,2}$ defined by $\Delta^j u(x) \in p^j \mathbb{Z}_p$ and proves that it is a ring. At $p = 2$, this is exactly the continuous p -adic version of the present class.

8.2 Mahler expansions and the exact coefficient condition

Newton interpolation has an infinite-series analogue over $\mathbb{Z}_p$. Mahler's interpolation theorem [6] identifies continuous functions $u: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ with uniformly convergent series

$$u(x) = \sum_{j \geq 0} a_j \binom{x}{j}, \quad a_j = \Delta^j u(0) \rightarrow 0.$$

This extension procedure also allows us to recover Sun's earlier product theorem from Kellner's ring theorem, for every prime p . Indeed, let $u: \mathbb{Z}_{\geq 0} \rightarrow R$ be p -regular and put $a_j = \Delta^j u(0)$. Embed R in $\mathbb{Z}_p$. Since $a_j \in p^j R$, the series

$$U(x) = \sum_{j \geq 0} a_j \binom{x}{j}$$

converges uniformly on $\mathbb{Z}_p$. Newton interpolation gives $U(n) = u(n)$ for every nonnegative integer n . The identity $\Delta \binom{x}{j} = \binom{x}{j-1}$ for $j \geq 1$ then gives

$$\Delta^k U(x) = \sum_{j \geq k} a_j \binom{x}{j-k} \in p^k \mathbb{Z}_p \quad \text{for all } k \geq 0 \text{ and } x \in \mathbb{Z}_p.$$

Here each binomial polynomial takes values in $\mathbb{Z}_p$, and the series still converges after division by p^k . Thus $U \in \mathcal{K}_{p,2}$. If v is another p -regular function, its extension V likewise belongs to $\mathcal{K}_{p,2}$. Kellner's ring theorem gives $UV \in \mathcal{K}_{p,2}$, so $\Delta^k(uv)(0) \in p^k \mathbb{Z}_p$. This finite difference also belongs to R . Since $R \cap p^k \mathbb{Z}_p = p^k R$, it follows that uv is p -regular, as asserted by Sun's theorem.

In this language the continuous p -adic version of average-integrality is precisely

$$a_j \in p^{\lfloor j/(p-1) \rfloor} \mathbb{Z}_p \quad \text{for all } j \geq 0. \quad (43)$$

For necessity, apply the difference criterion and evaluate the j th difference at zero. For sufficiency, note that $\Delta \binom{x}{j} = \binom{x}{j-1}$ for $j \geq 1$, by (8) after division by $j!$. Taking $r(p-1)$ differences therefore shifts the binomial indices by $r(p-1)$, just as in (33). Every remaining coefficient is divisible by p^r , and division by p^r still leaves a convergent series with values in $\mathbb{Z}_p$. The difference criterion now gives the desired integrality of the averages.

Every average-integral function has such a continuous extension, including agreement at negative integers. Here is a direct justification of the latter point. The bounds on $a_j = \Delta^j f(0)$ make the Mahler series converge uniformly and agree with f at nonnegative integers. For a fixed negative integer n , put

$$A_j = \Delta^j f(n), \quad P_N(x) = \sum_{j=0}^N A_j \binom{x-n}{j}.$$

The same bounds give a uniform limit u_n on $\mathbb{Z}_p$, and Newton interpolation gives $u_n(x) = f(x)$ for every integer $x \geq n$. Thus u_n and the extension constructed at zero agree on the dense set $\mathbb{Z}_{\geq 0}$, hence everywhere; in particular they agree at n . This proves the asserted extension. Conversely, the restriction of a function satisfying (43) is average-integral whenever its values on $\mathbb{Z}$ are ordinary integers.

8.3 A second proof using Stirling numbers

The Stirling number $S(N, m)$ of the second kind counts the partitions of an N -element set into m nonempty blocks. These integers also express ordinary powers in the falling-factorial basis:

$$x^N = \sum_{m=0}^N S(N, m)(x)_m.$$

For a nonnegative integer x , this identity counts maps from an N -element set to an x -element set: first partition the domain into its nonempty fibers, then choose distinct images for the blocks. It holds as a polynomial identity because it holds at every nonnegative integer x . Thus a divisibility theorem about $S(N, m)$ can be used in place of our elementary coefficient argument.

In [7, Lemma 2.1], Adelberg proves a lower bound on $v_p(S(N, m))$. In terms of factorial valuations, his bound is

$$v_p(S(N, m)) \geq v_p(N!) - v_p(m!) - \left\lfloor \frac{N - m}{p - 1} \right\rfloor \quad (44)$$

for all $N \geq m \geq 1$. This bound yields another proof of Theorem 1.1, as follows.

Let f be ℓ -average-integral, where $1 \leq \ell \leq p$, and fix $n \in \mathbb{Z}$. Lemma 4.4 supplies an integer-valued polynomial $P \in M_p$ with $P(r) = f(n + r)$ for $0 \leq r \leq \ell(p - 1)$. Write $P(x) = \sum_j c_j x^j$. By definition of M_p , $c_j \in p^{d(j)}R$. Since $d(j) \geq 0$, reducing modulo p therefore gives

$$\overline{P}(x) = \sum_{j: d(j)=0} \overline{c_j} x^j.$$

In characteristic p , the intermediate binomial coefficients $\binom{p}{a}$, for $1 \leq a < p$, vanish. Hence $(U + V)^p = U^p + V^p$; applying this identity repeatedly to the finite sum above gives

$$\overline{P}(x)^p = \sum_{j: d(j)=0} \overline{c_j}^p x^{pj} = \sum_{j: d(j)=0} \overline{c_j} x^{pj}.$$

The last equality uses $c^p = c$ in $\mathbb{F}_p$. Every exponent $N = pj$ in this sum satisfies $d(N) = d(j) = 0$, by Lemma 4.1(b).

As in the elementary proof, we want the falling-factorial coefficient at $m = k(p - 1)$ to be divisible by p . Stirling numbers describe exactly this change of basis. We will check that every monomial surviving modulo p has the required coefficient divisible by p . For $m = k(p - 1)$ with $1 \leq k \leq p$, we have $v_p(m!) = k - 1$. Therefore the right side of (44) simplifies to

$$v_p(N!) - (k - 1) - \left(\left\lfloor \frac{N}{p - 1} \right\rfloor - k \right) = 1 - d(N).$$

Taking $N = pj$ gives $p \mid S(pj, m)$ whenever $pj \geq m$ and $d(j) = 0$; for $pj < m$ the Stirling number is zero. It follows that

$$[(x)_m](\overline{P}^p) = \sum_{j: d(j)=0} \overline{c_j} \overline{S(pj, m)} = 0.$$

Thus, if b_m is the $(x)_m$ -coefficient of P^p , then $b_m \in pR$. By (9),

$$\Delta^m(P^p)(0) = m!b_m \in p^k R,$$

since $v_p(m!) = k - 1$. For $1 \leq k \leq \ell$, we have $m = k(p - 1) \leq \ell(p - 1)$. The polynomials P^p and the shifted function $r \mapsto f(n + r)^p$ therefore agree at $r = 0, \dots, m$. Applying (21) gives

$$\Delta^{k(p-1)}(f^p)(n) = \Delta^m(P^p)(0) \in p^k R.$$

The left side is an integer, so it is divisible by p^k . Since n and k were arbitrary, Lemma 1.4 gives $F_p^k(f^p) \in \mathbb{Z}^{\mathbb{Z}}$ for $0 \leq k \leq \ell$. The case $\ell = 0$ follows directly from integrality of f .

Thus the finite-depth conclusion can be deduced from this published valuation bound and the polynomial criterion already present in the 2011 MathOverflow discussion. The proof in Section 2 gives a direct finite-field argument for the particular vanishing needed here.

A Power sums over a prime field

We prove (18), working throughout in $\mathbb{F}_p$.

When $b = 0$, our convention $r^0 = 1$ makes the sum a sum of p ones, which is zero in $\mathbb{F}_p$. When $b = p - 1$, Fermat's little theorem gives $r^{p-1} = 1$ for every nonzero r , while $0^{p-1} = 0$. Hence the sum is $p - 1 = -1$ in $\mathbb{F}_p$.

It remains to consider $1 \leq b < p - 1$. Choose a nonzero residue a with $a^b \neq 1$. Such a residue exists because the polynomial $X^b - 1$ has at most b roots, fewer than the $p - 1$ nonzero residues. Write $S = \sum_{r \in \mathbb{F}_p} r^b$. Multiplication by a permutes $\mathbb{F}_p$, so

$$S = \sum_{r \in \mathbb{F}_p} (ar)^b = a^b \sum_{r \in \mathbb{F}_p} r^b = a^b S.$$

Since $a^b - 1 \neq 0$ in the field $\mathbb{F}_p$, the equality forces $S = 0$.

B Verification of Example 7.5

The following exact computation in Python verifies (42). After r passes through the loop, `counts[s]` is, modulo 243, the number of r -tuples with entries in $\{0, 1, \dots, 560\}$ whose sum is congruent to s modulo 243. Appending an entry a gives the recurrence used in the loop. The final sum weights each residue s by s^{561} modulo 243.

```
modulus = 243
counts = [1] + [0] * (modulus - 1)
for r in range(5):
    counts = [
        sum(counts[(s - a) % modulus] for a in range(561)) % modulus
        for s in range(modulus)
    ]
remainder = sum(counts[s] * pow(s, 561, modulus)
    for s in range(modulus)) % modulus
assert remainder == 162
```

This establishes failure of average-integrality by the fifth step; it does not assert that the fifth step is the first failure.

References

- [1] Darij Grinberg, *Generalization of Tamarkin's ARO 1993, final round, problem 10/8: part II*, MathOverflow question 68595, June 23, 2011. <https://mathoverflow.net/q/68595>.

- [2] Darij Grinberg, *Generalization of Tamarkin's ARO 1993, final round, problem 10/8: still a conjecture?*, MathOverflow question 66959, June 5, 2011. <https://mathoverflow.net/q/66959>.
- [3] zeb, Answer to MathOverflow question 66959, June 16, 2011 (counterexample and the weighted-monomial criterion). <https://mathoverflow.net/a/68000>.
- [4] Zhi-Hong Sun, *Congruences concerning Bernoulli numbers and Bernoulli polynomials*, Discrete Applied Mathematics **105** (2000), 193–223. doi:10.1016/S0166-218X(00)00184-0.
- [5] Bernd C. Kellner, *Classification of p -adic functions satisfying Kummer type congruences*, preprint, 2009, arXiv:0909.0743v2. <https://arxiv.org/abs/0909.0743>.
- [6] Kurt Mahler, *An interpolation series for continuous functions of a p -adic variable*, Journal für die reine und angewandte Mathematik **199** (1958), 23–34. doi:10.1515/crll.1958.199.23.
- [7] Arnold Adelberg, *The p -adic analysis of Stirling numbers via higher order Bernoulli numbers*, International Journal of Number Theory **14** (2018), no. 10, 2767–2779. Preprint: <https://arxiv.org/abs/1805.00995>; Lemma 2.1.
- [8] Ira M. Gessel, *Applications of the classical umbral calculus*, Algebra Universalis **49** (2003), no. 4, 397–434. doi:10.1007/s00012-003-1813-5. Preprint: <https://arxiv.org/abs/math/0108121>.
- [9] D. Tamarkin, Problem 8 for grade 10, final round of the 1993 All-Russian Mathematical Olympiad. Archived statement and solution, Problems.ru, problem 109520. https://problems.ru/view_problem_details_new.php?id=109520.