



Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt



General Section

Functions with integer-valued divided differences [☆]

Andrew O'Desky

Department of Mathematics, Princeton University, Princeton, NJ 08544-1000, USA

ARTICLE INFO

Article history:

Received 27 May 2020

Received in revised form 4 June 2021

Accepted 9 June 2021

Available online 22 July 2021

Communicated by F. Pellarin

MSC:

primary 11B99, 26E30

secondary 11C08, 13F20

Keywords:

Divided differences

Harmonic numbers

Integer-valued polynomials

Nonarchimedean analysis

ABSTRACT

Let s_0, s_1, s_2, \dots be a sequence of rational numbers whose m th divided difference $\delta_m s$ is integer-valued. We prove that s_n is a polynomial function in n if $|s_n| \ll \theta^n$ for some positive number θ satisfying $\theta < e^{1+\frac{1}{2}+\dots+\frac{1}{m}} - 1$.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

In 1971, Hall and Ruzsa independently discovered an elegant characterization of polynomial functions among congruence-preserving functions.

Theorem (Hall–Ruzsa). Let $s: \mathbb{N} \rightarrow \mathbb{Q}$ be a function. Suppose that¹

[☆] This work was supported in part by National Science Foundation RTG grant DMS-1840234.

E-mail address: andy.odesky@gmail.com.

URL: <http://www.andrewodesky.com/>.

¹ We write $r_n \ll q_n$ to mean there is a constant $C > 0$ such that $|r_n| \leq C|q_n|$ for all n .

- (i) s is integer-valued,
- (ii) $s(n+k) \equiv s(n) \pmod{k}$ for all $n, k \in \mathbb{N}$, and
- (iii) $|s(n)| \ll \theta^n$ for some positive number θ satisfying $\theta < e - 1$.

Then $s(n)$ is a polynomial in n .

In this article we generalize this theorem by using the integrality of higher divided differences. The first two hypotheses of the Hall–Ruzsa theorem are equivalent to requiring that $\delta_0 s := s$ and $\delta_1 s$ are integer-valued, where

$$\delta_1 s(n, m) := \frac{s(n) - s(m)}{n - m} \quad (n \neq m)$$

is the first divided difference of s .

Our main result is the following theorem.

Theorem 1. *Let $s: \mathbb{N} \rightarrow \mathbb{Q}$ be a function. Suppose that*

- (i) *the m th divided difference $\delta_m s$ of s is integer-valued, and*
- (ii) *$|s(n)| \ll \theta^n$ for some positive number θ satisfying*

$$\theta < e^{1 + \frac{1}{2} + \cdots + \frac{1}{m}} - 1.$$

Then $s(n)$ is a polynomial in n .

Recall the m th divided difference of s is the function $\delta_m s$ defined by

$$\delta_m s(n_0, \dots, n_m) := \sum_{i=0}^m \left\{ \prod_{j \neq i} (n_i - n_j)^{-1} \right\} s(n_i) \quad (n_i \text{ distinct}).$$

Divided differences are a classical construction of basic importance in interpolation theory and nonarchimedean analysis (cf. e.g. [7], [9], [2]).

1.1. Outline of the paper

The first part of this paper (§3) consists of a local analysis of $\delta_m s$ at nonarchimedean places. Our first new result is a formula for the p -adic supremum $\|\delta_m s\|_p$ of $\delta_m s$. To state our formula, we make use of a (possibly new) elementary arithmetic function. Let $\tau_{p,m}(n)$ denote the largest possible p -adic valuation of a product of m distinct positive integers which are all less than or equal to n . (We will compute an explicit formula for $\tau_{p,m}$ in Proposition 1.) Let $s: \mathbb{N} \rightarrow \mathbb{Q}_p$ be a function and set $c(n) := \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} s(k)$. We show that

$$\|\delta_m s\|_p = \sup_{n \geq m} |c(n)|_p p^{\tau_{p,m}(n)}. \quad (1)$$

This formula gives a Mahler-type criterion for the p -integrality of $\delta_m s$. Along the way, we prove a generalization of Mahler's theorem to bounded, not-necessarily-continuous functions (Theorem 2).

The second part of this paper (§4) consists of a global calculation involving some analytic estimates of certain sums over primes. Our third new result is that

$$\sum_{p \leq n} \tau_{p,m}(n) \log p = nH_m + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n) \quad (2)$$

for some positive constant α .

In the third part of this paper (§5) we combine the local and global results — (1) and (2) — to prove Theorem 1. In fact we will prove a more general version which simultaneously generalizes the main result of [4].

The p -integrality of divided differences is a natural condition on functions $s: \mathbb{N} \rightarrow \mathbb{Q}_p$ which does not appear to have been given a thorough treatment in the present literature on interpolation and p -adic analysis. Therefore it seems worthwhile to us to explore this condition somewhat. In the last section (§6), which is independent from the rest of the paper, we give two interpretations of this condition.

Notation. A place v of a number field K is an equivalence class of isometric embeddings $\sigma: K \rightarrow \mathbb{C}_p$ with $p \in M_{\mathbb{Q}} = \{2, 3, 5, \dots, \infty\}$ where \mathbb{C}_p is the metric completion of an algebraic closure of the p -adic field \mathbb{Q}_p (where $\mathbb{Q}_{\infty} := \mathbb{R}$). The set of all places of K is denoted by M_K . The local degree at v is denoted by d_v . We write $|\cdot|_p$ for the unique norm on \mathbb{C}_p extending the norm on \mathbb{Q}_p , normalized by $|p|_p = 1/p$. The norm associated to a place v of K extending $p \in M_{\mathbb{Q}}$ is defined by $|x|_v := |\sigma(x)|_p$ for a representative embedding σ of v . For nonarchimedean places, the additive v -adic valuation is defined by $\nu_v := \log_p |\cdot|_v^{-1}$ where \log_p is the base- p logarithm. We write $[\cdot]$ for the floor function and $H_m = 1 + \frac{1}{2} + \dots + \frac{1}{m}$ for the m th harmonic number. We write $\vartheta(n) = \sum_{p \leq n} \log p$ for the Chebyshev function and $\pi(n)$ for the number of rational primes less than or equal to n . We define $\Delta_m := \{(n_0, \dots, n_m) \in \mathbb{N}^{m+1} : n_i \text{ all distinct}\}$.

2. Background from difference calculus and p -adic analysis

Let m be a non-negative integer. Define

$$\Delta_m := \{(n_0, \dots, n_m) \in \mathbb{N}^{m+1} : n_i \text{ all distinct}\}.$$

Let $s: \mathbb{N} \rightarrow E$ be a function valued in a \mathbb{Q} -algebra E . Recall the m th divided difference of s is the function $\delta_m s: \Delta_m \rightarrow E$ given by

$$\delta_m s(n_0, \dots, n_m) := \sum_{i=0}^m \left\{ \prod_{j \neq i} (n_i - n_j)^{-1} \right\} s(n_i).$$

(It is well-known that s can be reconstructed using the values of its divided differences by means of Newton's interpolation formula (cf. (22) or [7], §1).) Recall the n th finite difference of s is defined by²

$$c(n) := \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} s(k). \quad (3)$$

Lemma 1. *Let $s: \mathbb{N} \rightarrow E$ be a function and let $c: \mathbb{N} \rightarrow E$ be its finite differences. Then s is polynomial if and only if c is eventually zero.*

Proof. Let S be the forward shift operator on sequences defined by $(Ss)(n) := s(n+1)$ for all non-negative integers n . Then for any non-negative integer ℓ ,

$$\{(S - \text{id})^n s\}(\ell) = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} s(\ell + k),$$

and in particular, $\{(S - \text{id})^n s\}(0) = c(n)$. We have that $(S - \text{id})(n^d) = dn^{d-1} + O(n^{d-2})$, and so the restriction of $S - \text{id}$ to the space of polynomial sequences is nilpotent. This shows that c is eventually zero if s is polynomial.

Conversely, assume that c is eventually zero. It is easy to verify that the inverse relation of (3) is given by

$$s(n) = \sum_{k=0}^n \binom{n}{k} c(k), \quad (4)$$

and this shows that s is given by a polynomial of degree $\leq N$ if $c(n) = 0$ for $n > N$. \square

We recall a classical theorem of Mahler. Let $\binom{x}{n} := \frac{x(x-1)\cdots(x-n+1)}{n!} \in \mathbb{Q}[x]$.

Theorem (Mahler [6]). *Let $s: \mathbb{N} \rightarrow \mathbb{C}_p$ be a function and let $c: \mathbb{N} \rightarrow \mathbb{C}_p$ denote its finite differences. Then s extends to a continuous function $f: \mathbb{Z}_p \rightarrow \mathbb{C}_p$ if and only if $\lim_{n \rightarrow \infty} |c(n)|_p = 0$. Furthermore, if s extends to a continuous function $f: \mathbb{Z}_p \rightarrow \mathbb{C}_p$, then the ‘Mahler series’*

$$\sum_{n=0}^{\infty} c(n) \binom{x}{n}$$

converges uniformly to f in $C(\mathbb{Z}_p, \mathbb{C}_p)$ and

² Strictly speaking, this is the sequence obtained by evaluating the finite differences of s at zero.

$$\sup_{x \in \mathbb{Z}_p} |f(x)|_p = \sup_{n \geq 0} |c(n)|_p.$$

3. The p -adic size of divided differences

Let p be a fixed prime integer. Let E be a finite extension of \mathbb{Q}_p and let $|\cdot|$ denote the norm on E which extends the usual p -adic norm on \mathbb{Q}_p . Let $s: \mathbb{N} \rightarrow E$ be a function and set $c(n) = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} s(k)$. The goal of this section is to prove the formula

$$\sup_{0 \leq n_0 < \dots < n_m} |\delta_m s(n_0, \dots, n_m)| = \sup_{n \geq m} |c(n)| p^{\tau_{p,m}(n)}$$

where $\tau_{p,m}(n)$ is the maximal p -adic valuation of a product of m distinct positive integers $\leq n$, i.e.

$$\tau_{p,m}(n) := \max_{\substack{S \subset \{1, \dots, n\}, \\ \#S=m}} \nu_p \left(\prod_{k \in S} k \right).$$

3.1. The Mahler series of a bounded function

Let $\ell^\infty(\mathbb{N}^{m+1})$ denote the E -Banach space of bounded functions $F: \mathbb{N}^{m+1} \rightarrow E$ under pointwise addition and scalar multiplication equipped with the norm

$$\|F\| := \sup_{\underline{n} \in \mathbb{N}^{m+1}} |F(\underline{n})|.$$

This p -adic Banach space contains two important closed subspaces:

$$C(\mathbb{N}^{m+1}) := \{F: \mathbb{N}^{m+1} \rightarrow E \mid F \text{ extends to a continuous function } \mathbb{Z}_p^{m+1} \rightarrow E\}$$

and³

$$c_0(\mathbb{N}^{m+1}) := \{C: \mathbb{N}^{m+1} \rightarrow E \mid C \text{ vanishes at infinity}\}$$

Let $\underline{x} = (x_0, \dots, x_m)$ be indeterminates, $\underline{j} = (j_0, \dots, j_m)$ be non-negative integers, and define $\binom{\underline{x}}{\underline{j}} := \binom{x_0}{j_0} \binom{x_1}{j_1} \dots \binom{x_m}{j_m} \in \mathbb{Q}[x_0, \dots, x_m]$. Mahler's theorem easily generalizes to multivariate functions (cf. [1], Corollaire 1, §2.7) and amounts to the assertion that there is a unique isometry

$$\begin{aligned} \mathcal{M}_0: C(\mathbb{N}^{m+1}) &\xrightarrow{\sim} c_0(\mathbb{N}^{m+1}) \\ F &\mapsto C \end{aligned}$$

³ A function $C: \mathbb{N}^{m+1} \rightarrow E$ vanishes at infinity if $\lim_{N \rightarrow \infty} \sup_{j_0 + \dots + j_m > N} |C(\underline{j})| = 0$.

with the property that for all $F \in C(\mathbb{N}^{m+1})$ the infinite series

$$\sum_{\underline{j} \in \mathbb{N}^{m+1}} C(\underline{j}) \binom{\underline{x}}{\underline{j}}$$

converges in $C(\mathbb{N}^{m+1})$ to F .

We are interested in computing the norm of $\delta_m s$ (strictly speaking, $\delta_m s$ is not defined on all of \mathbb{N}^{m+1} , but this technical issue will be easily resolved later). Mahler's theorem is therefore insufficient for our purposes, since $\delta_m s$ is only assumed to be p -integral and inhabits neither $C(\mathbb{N}^{m+1})$ nor $c_0(\mathbb{N}^{m+1})$ in general.

To resolve this problem, we will show that the isometry \mathcal{M}_0 extends to a self-isometry of $\ell^\infty(\mathbb{N}^{m+1})$. The following notion will be helpful: if $F, C: \mathbb{N}^{m+1} \rightarrow E$ are functions, not necessarily bounded or continuous, we say that

$$\sum_{\underline{j} \in \mathbb{N}^{m+1}} C(\underline{j}) \binom{\underline{x}}{\underline{j}}$$

is a *generalized Mahler series* for F if

$$F(\underline{n}) = \sum_{\underline{j} \in \mathbb{N}^{m+1}} C(\underline{j}) \binom{\underline{n}}{\underline{j}} \quad (\text{finite sum}) \quad (5)$$

for all $\underline{n} \in \mathbb{N}^{m+1}$. The values of C are called the *Mahler coefficients* of F . The following theorem generalizes Mahler's theorem to bounded functions which are not necessarily continuous.

Theorem 2. *Every function $F: \mathbb{N}^{m+1} \rightarrow E$ admits a generalized Mahler series with unique Mahler coefficients. Furthermore, the mapping $F \mapsto C$ which sends a bounded function F to its Mahler coefficients C extends $\mathcal{M}_0: C(\mathbb{N}^{m+1}) \xrightarrow{\sim} c_0(\mathbb{N}^{m+1})$ to a self-isometry \mathcal{M} of $\ell^\infty(\mathbb{N}^{m+1})$.*

Note however that the proposition does not imply that $\{\binom{\underline{x}}{\underline{j}}\}_{\underline{j} \in \mathbb{N}^{m+1}}$ is an orthonormal Banach basis for $\ell^\infty(\mathbb{N}^{m+1})$, as generalized Mahler series are typically divergent in $\ell^\infty(\mathbb{N}^{m+1})$.

Proof. We proceed by induction on m . If $m = 0$ then we take $C(n)$ to be the n th finite difference of F given by $C(n) := \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} F(k)$. The inverse relation is given by (4) and uniqueness follows. If F is bounded then C is bounded by the ultrametric inequality and vice versa. This mapping restricts to an isometry $C(\mathbb{N}) \xrightarrow{\sim} c_0(\mathbb{N})$ by Mahler's theorem (cf. §2). To see that $\mathcal{M}: F \mapsto C$ is an isometry when F is bounded it suffices to observe that the relation (3) and its inverse (4) are both defined over \mathbb{Z} and to apply the ultrametric inequality.

Now suppose m is a positive integer. Fix a non-negative integer a and define the function G_a by

$$G_a(n_0, \dots, n_{m-1}) := F(n_0, \dots, n_{m-1}, a).$$

By induction, G_a has a generalized Mahler series

$$\sum_{\underline{j} \in \mathbb{N}^m} D_a(\underline{j}) \binom{\underline{x}}{\underline{j}}$$

for uniquely determined $D_a: \mathbb{N}^m \rightarrow E$. For fixed $\underline{j} \in \mathbb{N}^m$, the function $a \mapsto D_a(\underline{j})$ also has a generalized Mahler series $\sum_{k \geq 0} d_{\underline{j}}(k) \binom{y}{k}$. We put these together to get

$$\sum_{\underline{j} \in \mathbb{N}^m} \sum_{k \geq 0} d_{\underline{j}}(k) \binom{\underline{x}}{\underline{j}} \binom{y}{k}$$

which is therefore the generalized Mahler series of F , with Mahler coefficients given by $C(j_0, \dots, j_m) := d_{(j_0, \dots, j_{m-1})}(j_m)$. Uniqueness of C follows from uniqueness of $d_{\underline{j}}$ and D_a .

If F is bounded then

$$\|F\| = \sup_{\underline{n} \in \mathbb{N}^{m+1}} |F(\underline{n})| = \max_{\underline{n} \in \mathbb{N}^{m+1}} |F(\underline{n})| = \max_{a \in \mathbb{N}} \|G_a\|$$

since $|E^\times| \subset \mathbb{R}^{>0}$ is discrete. By induction, and discreteness of $|E^\times|$ again,

$$\|F\| = \max_{a \in \mathbb{N}} \|G_a\| = \max_{a \in \mathbb{N}} \|D_a\| = \max_{\underline{j} \in \mathbb{N}^m} \max_{a \in \mathbb{N}} |D_a(\underline{j})| = \max_{\underline{j} \in \mathbb{N}^m} \max_{a \in \mathbb{N}} |d_{\underline{j}}(a)| = \|C\|.$$

This proves that C is bounded if F is bounded; conversely, if C is bounded then F is bounded by applying the ultrametric inequality to (5). When F and C are both bounded, we see that $\mathcal{M}: F \mapsto C$ is a self-isometry of $\ell^\infty(\mathbb{N}^{m+1})$. The restriction of \mathcal{M} to $C(\mathbb{N}^{m+1})$ is equal to \mathcal{M}_0 by uniqueness of Mahler coefficients. \square

3.2. Proof of supremum formula

Let $s: \mathbb{N} \rightarrow E$ be a function and let m be a non-negative integer. As before, let $\tau_{p,m}(n) = \max_{1 \leq i_1 < \dots < i_m \leq n} (\nu_p(i_1 \cdots i_m))$ and $c(n) := \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} s(k)$.

Theorem 3. *We have the equality*

$$\|\delta_m s\| = \sup_{n \geq m} |c(n)| p^{\tau_{p,m}(n)}. \quad (6)$$

We will show that (6) holds even if both sides are infinite.

Proof. As $\delta_0 s = s$ and $\tau_{p,0}(n) = 0$, the $m = 0$ case follows from Mahler's theorem (cf. §2). We suppose $m \geq 1$. Theorem 2 does not apply directly to $\delta_m s$, as the domain of $\delta_m s$ is Δ_m and not all of \mathbb{N}^{m+1} . We can resolve this with a change of variables as follows. Let $(a_0, \dots, a_m) \in \mathbb{N}^{m+1}$ and define

$$\begin{aligned} n_m &:= a_m, \\ n_{m-1} &:= a_m + (a_{m-1} + 1), \\ &\vdots \\ n_0 &:= a_m + (a_{m-1} + 1) + \dots + (a_0 + 1). \end{aligned}$$

This defines a bijection $(a_0, \dots, a_m) \leftrightarrow (n_0, \dots, n_m)$ from \mathbb{N}^{m+1} to the subset of Δ_m whose entries are strictly decreasing. Define

$$\begin{aligned} F: \mathbb{N}^{m+1} &\rightarrow E \\ (a_0, \dots, a_m) &\mapsto \delta_m s(n_0, \dots, n_m). \end{aligned}$$

Since $\delta_m s$ is a symmetric function, all of its values are realized on the subset of Δ_m whose entries are strictly decreasing, so F takes the same values as $\delta_m s$. In particular, $\|F\| = \|\delta_m s\|$.

We continue letting $\underline{a} \leftrightarrow \underline{n}$ using the bijection above. To compute the Mahler coefficients of F we make use of the following formula of Schikhof (cf. [9], Theorem 54.1),

$$\begin{aligned} F(\underline{a}) &= \sum_{\underline{j} \in \mathbb{N}^{m+1}} C(\underline{j}) \binom{\underline{a}}{\underline{j}} \quad \text{for all } \underline{a} \in \mathbb{N}^{m+1} \\ \text{where } C(\underline{j}) &= \frac{c(j_0 + \dots + j_m + m)}{(j_{m-1} + 1)(j_{m-1} + j_{m-2} + 2) \cdots (j_{m-1} + \dots + j_0 + m)}. \end{aligned}$$

By uniqueness of Mahler coefficients, this is the generalized Mahler series for F (Theorem 2).

Set $i_k := j_{m-1} + \dots + j_{m-k} + k$. Then

$$\|C\| = \sup_{1 \leq i_1 < \dots < i_m \leq \ell} \left| \frac{c(\ell)}{i_1 i_2 \cdots i_m} \right| = \sup_{m \leq \ell} |c(\ell)| p^{\tau_{p,m}(\ell)}$$

where $\tau_{p,m}(\ell) := \max_{1 \leq i_1 < \dots < i_m \leq \ell} (\nu_p(i_1) + \dots + \nu_p(i_m))$. Theorem 2 asserts that

$$\|F\| = \|C\|.$$

Since $\|F\| = \|\delta_m s\|$ this concludes the proof. \square

Corollary 1. *Let $s: \mathbb{N} \rightarrow \mathbb{Q}$ be a function. If $\delta_{m+1}s$ is \mathbb{Z} -valued then there is an explicit positive integer C such that $C\delta_k s$ is \mathbb{Z} -valued for all $k \in \{0, \dots, m\}$.*

Proof. The proposition implies that for any $k \in \{0, \dots, m\}$,

$$\begin{aligned} \|\delta_k s\| &= \sup_{n \geq k} |c(n)| p^{\tau_{p,k}(n)} = \sup \left\{ |c(k)| p^{\tau_{p,k}(k)}, \sup_{n \geq k+1} |c(n)| p^{\tau_{p,k}(n)} \right\} \\ &\leq \sup \left\{ |c(k)| |k!|^{-1}, \|\delta_{k+1} s\| \right\}. \end{aligned}$$

Thus

$$\sup \left\{ \|\delta_0 s\|, \|\delta_1 s\|, \dots, \|\delta_m s\| \right\} \leq \sup \left\{ \left| \frac{c(0)}{0!} \right|, \left| \frac{c(1)}{1!} \right|, \dots, \left| \frac{c(m)}{m!} \right|, 1 \right\},$$

and we can take

$$C = \prod_{p \text{ prime}} \sup \left\{ \left| \frac{c(0)}{0!} \right|_p, \left| \frac{c(1)}{1!} \right|_p, \dots, \left| \frac{c(m)}{m!} \right|_p, 1 \right\}.$$

Note that only finitely many terms in the product differ from unity. \square

4. Asymptotic formulas for certain sums over primes

Let K denote a finite extension of \mathbb{Q} . In the previous section we established a local estimate for the finite differences c of a function $s: \mathbb{N} \rightarrow K$ with integral m th divided difference. Namely, for any nonarchimedean place v of K lying over p_v , Theorem 3 shows that

$$\delta_m s \text{ } v\text{-integral} \implies |c(n)|_v \leq p_v^{-\tau_{p_v,m}(n)}.$$

Now suppose that $\delta_m s$ is v -integral for *all* nonarchimedean places v of K . Then

$$\prod_{v \text{ finite}} |c(n)|_v \leq \prod_{v \text{ finite}} p_v^{-\tau_{p_v,m}(n)} = \exp \left\{ - \sum_{v \text{ finite}} \tau_{p_v,m}(n) \log p_v \right\}.$$

Our goal is to achieve a precise upper bound for $\prod_{v \text{ finite}} |c(n)|_v$, which will be used later to obtain a lower bound for $\prod_{v \text{ infinite}} |c(n)|_v$ by the product formula. Therefore we presently turn to the asymptotic growth of $\sum_{p \leq n} \tau_{p,m}(n) \log p$ in the limit of large n . The goal of this section is to prove that

$$\sum_{p \leq n} \tau_{p,m}(n) \log p = n \left(1 + \frac{1}{2} + \dots + \frac{1}{m} \right) + o(n). \quad (7)$$

The standard bound for the Chebyshev function $\vartheta(x) = x + O\left(\frac{x}{\log x}\right)$ is not strong enough to establish the above asymptotic formula, so we employ the following useful estimate due to Rosser–Schoenfeld (cf. [8], (2.29)):

$$\vartheta(x) = x + O(x \exp\{-\alpha(\log x)^{1/2}\}) \quad (8)$$

for some positive constant α . In fact, this will lead to the improved error term $O(n \exp\{-\alpha(\log n)^{1/2}\} \log n)$ in place of $o(n)$ in (7).

4.1. A formula for $\tau_{p,m}$

We begin by deriving an explicit formula for $\tau_{p,m}(n)$ which will be needed later. Let \log_p denote the base- p logarithm and let $[\cdot]$ denote the floor function. Let $a_p(n) := [np^{-[\log_p n]}]$ (the top nonzero p -adic digit in the usual base- p expansion of n).

Proposition 1. *Let $m > 0$, $n \geq m$ be integers and let p be a prime $\geq m$. Then*

$$\tau_{p,m}(n) = \begin{cases} m[\log_p n] & \text{if } a_p(n) \geq m, \\ m[\log_p n] + a_p(n) - m & \text{if } a_p(n) < m. \end{cases} \quad (9)$$

The formula can fail if $p < m$, e.g. $\tau_{p+1,p}(p^2) = p + 1$ whereas (9) gives $p + 2$.

Proof. Write $t := [\log_p n]$. We will construct integers $1 \leq i_1 < \dots < i_m \leq n$ whose product realizes the maximum p -adic valuation. If $a_p(n) \geq m$ then set

$$i_k = (a_p(n) - m + k)p^t, \quad k \in \{1, \dots, m\}.$$

The p -adic valuation of $i_1 \dots i_m$ is mt and this is maximal since each integer has the maximum possible p -adic valuation of any positive integer $\leq n$. If $a_p(n) < m$ then set

$$i_k = \begin{cases} (a_p(n) - m + k)p^t & \text{if } k \in \{m - a_p(n) + 1, \dots, m\}, \\ p^t - (m - a_p(n) + 1 - k)p^{t-1} & \text{if } k \in \{1, \dots, m - a_p(n)\}. \end{cases}$$

As $p \geq m$ by hypothesis, $m - a_p(n) \leq p - a_p(n) \leq p - 1$ and so $m - a_p(n) + 1 - k \in \{1, \dots, p - 1\}$ for any $k \in \{1, \dots, m - a_p(n)\}$. Therefore the p -adic valuation of $p^t - (m - a_p(n) + 1 - k)p^{t-1}$ is $t - 1$ for any $k \in \{1, \dots, m - a_p(n)\}$. Putting these together obtains

$$\tau_{p,m}(n) = ta_p(n) + (t - 1)(m - a_p(n)) = mt + a_p(n) - m. \quad \square$$

4.2. Asymptotic formulas for certain sums over primes

First we prove a simple lemma. Let $[\cdot]: \mathbb{R} \rightarrow \mathbb{Z}$ denote the floor function.

Lemma 2.

$$\sum_{p \leq n} [\log_p n] \log p = n + O(n \exp\{-\alpha(\log n)^{1/2}\})$$

for some positive constant α .

Proof. For any prime p in the sum, $r := [\log_p n]$ must be positive. Then

$$[\log_p n] = r \iff \frac{\log n}{r+1} < \log p \leq \frac{\log n}{r} \iff n^{\frac{1}{r+1}} < p \leq n^{\frac{1}{r}}$$

and

$$\begin{aligned} 0 \leq \sum_{p \leq n} [\log_p n] \log p &= \sum_{r=1}^{\infty} \sum_{n^{\frac{1}{r+1}} < p \leq n^{\frac{1}{r}}} r \log p \\ &\leq \sum_{\sqrt{n} < p \leq n} \log p \\ &= \vartheta(n) - \vartheta(\sqrt{n}). \end{aligned}$$

The last term is $n + O(n \exp\{-\alpha(\log n)^{1/2}\})$ by (8). \square

Let $H_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m}$ be the m th harmonic number.

Theorem 4.

$$\sum_{p \leq n} \tau_{p,m}(n) \log p = nH_m + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n)$$

for some positive constant α .

Proof. This is clear if m is zero so we suppose m is positive. The explicit formula in Proposition 1 suggests that we separate $\tau_{p,m}(n)$ into a logarithmic term and a ‘remainder term’. The asymptotic contribution to $\sum_{p \leq n} \tau_{p,m}(n) \log p$ from the logarithmic term is established by Lemma 2:

$$\sum_{p \leq n} m[\log_p n] \log p = mn + O(n \exp\{-\alpha(\log n)^{1/2}\}). \quad (10)$$

The asymptotic contribution to the sum from the remainder $m[\log_p n] - \tau_{p,m}(n)$ is significantly more difficult to establish. We will show that for some positive constant α ,

$$\sum_{p \leq n} (m[\log_p n] - \tau_{p,m}(n)) \log p = (m - H_m)n + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n). \quad (11)$$

Taking the difference of (10) and (11) immediately proves the result so we now establish (11). Let $a \in \{1, 2, \dots, m-1\}$ and $n \in \{m, m+1, \dots\}$. For a prime p , let $a_p(n) = [np^{-[\log_p n]}]$ (the top nonzero p -adic digit in the base- p expansion of n). We separate primes according to the value of $a_p(n)$. Define the set

$$P_{a,n} := \{p \text{ prime} : a = a_p(n), \ m \leq p \leq n\}.$$

Consider the sum

$$\gamma(n) := \sum_{\substack{p: m \leq p \leq n, \\ a_p(n) < m}} (m - a_p(n)) \log p = \sum_{a=1}^{m-1} \sum_{p \in P_{a,n}} (m - a) \log p.$$

Note that

$$\tau_{p,m}(n) \leq \max_{1 \leq i_1 \leq \dots \leq i_m \leq n} (\nu_p(i_1) + \dots + \nu_p(i_m)) = m \max_{1 \leq i \leq n} \nu_p(i) = m[\log_p n],$$

and so $\sum_{p < m} (m[\log_p n] - \tau_{p,m}(n)) \log p = O(\log n)$. Then by Proposition 1,

$$\begin{aligned} \sum_{p \leq n} (m[\log_p n] - \tau_{p,m}(n)) \log p \\ = O(\log n) + \sum_{\substack{m \leq p \leq n \\ a_p(n) < m}} (m - a_p(n)) \log p = \gamma(n) + O(\log n). \end{aligned} \quad (12)$$

We will show that

$$\gamma(n) = (m - H_m)n + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n)$$

which will prove (11) in view of (12).

We claim that a prime p is in $P_{a,n}$ if and only if $p \geq m$ and $a = [np^{-t}]$ for some $t \in \{1, 2, \dots, [\log_2 n]\}$. Indeed, if $p \in P_{a,n}$ then $a = [np^{-t}]$ where $t = [\log_p n] \in \{1, 2, \dots, [\log_2 n]\}$. Conversely, if $a = [np^{-t}]$ for some $t \in \{1, 2, \dots, [\log_2 n]\}$, then $a \geq 1$ implies that $p^t \leq n$ and $p \geq m > a$ implies that $n < p^{t+1}$. Therefore $t = [\log_p n]$ and $a = a_p(n)$. This proves the claim. It follows that

$$\gamma(n) = \sum_{a=1}^{m-1} (m - a) \sum_{t=1}^{[\log_2 n]} \sum_{\substack{p: p \geq m, \\ a = [np^{-t}]}} \log p.$$

Now observe that

$$\begin{aligned} a = [np^{-t}] &\iff a \leq np^{-t} < a + 1 \\ &\iff an^{-1} \leq p^{-t} < (a + 1)n^{-1} \\ &\iff (n(a + 1)^{-1})^{\frac{1}{t}} < p \leq (na^{-1})^{\frac{1}{t}} \end{aligned}$$

and thus

$$\begin{aligned}
\gamma(n) &= \sum_{a=1}^{m-1} (m-a) \sum_{t=1}^{\lfloor \log_2 n \rfloor} \sum_{\substack{p: p \geq m, \\ (n(a+1)^{-1})^{1/t} < p \leq (na^{-1})^{1/t}}} \log p \\
&= O(1) + \sum_{t=1}^{\lfloor \log_2 n \rfloor} \sum_{a=1}^{m-1} (m-a) (\vartheta(\{na^{-1}\}^{\frac{1}{t}}) - \vartheta(\{n(a+1)^{-1}\}^{\frac{1}{t}}))
\end{aligned}$$

where $\vartheta(x) = \sum_{p \leq x} \log p$. Let $\gamma_t(n)$ denote the inner sum of the last expression for $1 \leq t \leq \lfloor \log_2 n \rfloor$ so that $\gamma(n) = O(1) + \sum_{t=1}^{\lfloor \log_2 n \rfloor} \gamma_t(n)$. Then

$$\gamma_t(n) = m\vartheta(\{\frac{n}{1}\}^{1/t}) - \vartheta(\{\frac{n}{1}\}^{1/t}) - \vartheta(\{\frac{n}{2}\}^{1/t}) - \dots - \vartheta(\{\frac{n}{m}\}^{1/t}).$$

By (8) there is a positive constant α such that

$$\vartheta(\{\frac{n}{a}\}^{1/t}) = (\frac{n}{a})^{1/t} + O(n^{1/t} \exp\{-\alpha(\log n)^{1/2}\}).$$

Therefore

$$\gamma_1(n) = (m - H_m)n + O(n \exp\{-\alpha(\log n)^{1/2}\}).$$

Since $n^{1/t} = O(n \exp\{-\alpha(\log n)^{1/2}\})$ for $t \geq 2$, we get that

$$\begin{aligned}
\gamma(n) &= \gamma_1(n) + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n) \\
&= (m - H_m)n + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n). \quad \square
\end{aligned}$$

Remark. The proof shows that any improved bound $\vartheta(x) - x = O(f(x))$ would lead to an improved bound of the form $nH_m - \sum_{p \leq n} \tau_{p,m}(n) \log p = O(f(n) \log n)$. In particular, if the Riemann hypothesis is true then for any $\varepsilon > 0$,

$$\sum_{p \leq n} \tau_{p,m}(n) \log p = nH_m + O(n^{1/2+\varepsilon} \log n).$$

5. Proof of Theorem 1

In this section we combine the local and global calculations from §3 and §4 to prove the main theorem of the paper. It is easy to work with a general number field in place of \mathbb{Q} , and we choose to do so for the sake of generality. Let K be a finite extension of \mathbb{Q} of degree d . Let M_K denote the set of places of K . For any place $v \in M_K$ let d_v denote the local degree at v .

Lemma 3. *Let $s: \mathbb{N} \rightarrow K$ be a function and let $S \subset M_K$ be a finite set containing the archimedean places. Suppose that*

- (i) $\delta_m s$ is S -integer-valued,⁴ and
(ii) for each v in S there is a positive number ρ_v such that $|c(n)|_v \ll \rho_v^n$.

If

$$\prod_{v \in S} \rho_v^{d_v} < e^{d(1+\frac{1}{2}+\dots+\frac{1}{m})} \quad (13)$$

then $s(n)$ is a polynomial in n .

Proof. By Lemma 1 the conclusion is equivalent to c being eventually zero, so for the sake of contradiction suppose that c has infinitely many nonzero terms.

Let v be a place of K not in S , p_v the rational prime v lies over, σ_v a representative embedding for v , and d_v the local degree at v . By Theorem 3, $|c(n)|_v = |\sigma_v c(n)|_{p_v} \leq p_v^{-\tau_{p_v,m}(n)}$ for $n \geq m$, and so

$$\prod_{v \notin S} |c(n)|_v^{d_v} \leq \prod_{v \notin S} p_v^{-d_v \tau_{p_v,m}(n)}. \quad (14)$$

Note that both sides amount to finite products ($\tau_{p_v,m}(n) = 0$ if $p_v > n$).

Recall that the sum of local degrees at all places lying over a given prime is the global degree, and also that $\tau_{p,m}(n) = O(\log n)$. Then

$$\begin{aligned} \sum_{v \notin S} d_v \tau_{p_v,m}(n) \log p_v &= \sum_{p_v \leq n} d_v \tau_{p_v,m}(n) \log p_v - \sum_{v \in S} d_v \tau_{p_v,m}(n) \log p_v \\ &= d \sum_{p \leq n} \tau_{p,m}(n) \log p + O(\log n). \end{aligned}$$

Then with the help of Theorem 4 we see that

$$\sum_{v \notin S} d_v \tau_{p_v,m}(n) \log p_v = dnH_m + o(n).$$

Putting this together with (14) obtains

$$\limsup_{n \rightarrow \infty} \prod_{v \notin S} |c(n)|_v^{d_v/n} \leq e^{-dH_m}.$$

Let n_i be chosen so that $c(n_i) \neq 0$ for all non-negative integers i . With the help of the product formula, $\prod_{v \in M_K} |c(n_i)|_v^{d_v} = 1$, we obtain

$$\prod_{v \in S} \rho_v^{-d_v} \leq \liminf_{i \rightarrow \infty} \prod_{v \in S} |c(n_i)|_v^{-d_v/n_i} = \limsup_{i \rightarrow \infty} \prod_{v \notin S} |c(n_i)|_v^{d_v/n_i} \leq e^{-dH_m}.$$

⁴ i.e., $\delta_m s(\underline{n})$ is v -integral for all $\underline{n} \in \Delta_m$ and $v \notin S$.

This contradicts (13) and concludes the proof. \square

Now we will prove a generalization of Theorem 1. In addition to the integrality of higher divided differences we will simultaneously consider *p-adic analytic interpolation*, i.e. the existence of a power series $f(x) \in \mathbb{C}_p[[x]]$ which converges for all $x \in \mathbb{D}_p^{<R} := \{x \in \mathbb{C}_p : |x|_p < R\}$ such that $R > 1$ and $f(n) = \sigma_v s(n)$ for all $n \geq 0$, where σ_v is a representative embedding for v .

Theorem 5. *Let $s: \mathbb{N} \rightarrow K$ be a function, let $S \subset M_K$ be a finite set containing the archimedean places M_K^∞ , and let $T \subset M_K$ be a finite set disjoint from S . Suppose that*

- (i) $\delta_m s$ is $(S \cup T)$ -integer-valued,
- (ii) for each v in S there is a positive number θ_v such that $|s(n)|_v \ll \theta_v^n$, and
- (iii) for each v in T there is a positive number $R_v > 1$ such that $\sigma_v s$ extends to a *p*-adic analytic function $\mathbb{D}_{p_v}^{<R_v} \rightarrow \mathbb{C}_{p_v}$.

If

$$\prod_{v \in M_K^\infty} (1 + \theta_v)^{d_v} \prod_{v \in S \setminus M_K^\infty} \max\{1, \theta_v\}^{d_v} \prod_{v \in T} (p_v^{\frac{1}{p_v-1}} R_v)^{-d_v} < e^{d(1+\frac{1}{2}+\dots+\frac{1}{m})} \quad (15)$$

then $s(n)$ is a polynomial in n .

Proof of Theorem 5. By Lemma 3 we see that s is polynomial if for each $v \in S \cup T$ there are positive constants D_v and ρ_v such that

$$|c(n)|_v \leq D_v \rho_v^n \quad \text{for all } n \geq 0 \quad (16)$$

and

$$\prod_{v \in S \cup T} \rho_v^{d_v} < e^{dH_m}. \quad (17)$$

We can construct the D_v and ρ_v as follows. First suppose that v is a nonarchimedean place in S . As $|s(n)|_v \ll \theta_v^n$ by hypothesis, there is a positive constant D_v such that

$$|c(n)|_v \leq \max_{0 \leq k \leq n} |s(k)|_v \leq \max_{0 \leq k \leq n} D_v \theta_v^k = \begin{cases} D_v & \text{if } \theta_v < 1, \\ D_v \theta_v^n & \text{if } \theta_v \geq 1. \end{cases}$$

In either case we can take $\rho_v = \max\{1, \theta_v\}$. Now suppose v is a archimedean place in S . We have that

$$|c(n)|_v \leq \sum_{0 \leq k \leq n} \binom{n}{k} |s(k)|_v.$$

Therefore for some positive constant D_v we have that $|c(n)|_v \leq D_v(1 + \theta_v)^n$, so we can take $\rho_v = 1 + \theta_v$. Then (16) is satisfied for every v in S .

Now we consider the places v in T . By hypothesis, for any place v in T there exists a power series $f_v(x) \in \mathbb{C}_p[[x]]$ converging on $\mathbb{D}_{p_v}^{<R_v}$ such that $\sigma_v s(n) = f_v(n)$ for all $n \geq 0$. Without loss of generality, we may assume that f_v converges on a disk of radius strictly larger than R_v for all $v \in T$ since the inequality (15) remains valid even if R_v is replaced with a sufficiently close but smaller quantity; thus, there is an $\varepsilon > 0$ such that for all v in T , f_v converges on an open disk of radius $R_v + \varepsilon$. With the help of a theorem of Iwasawa (cf. [5], Theorem 3), we see that

$$\lim_{n \rightarrow \infty} |c(n)|_v r^{-n} = 0 \quad (18)$$

for any real number r such that

$$p_v^{\frac{-1}{p_v-1}} (R_v + \varepsilon)^{-1} < r. \quad (19)$$

Then (18) implies that $|c(n)|_v^{1/n} > r$ for only finitely many n , and thus

$$\limsup_{n \rightarrow \infty} |c(n)|_v^{1/n} \leq r.$$

This shows that for every $v \in T$ there are positive constants D_v, ρ_v satisfying (16), and additionally that $\rho_v \leq r$. As r was arbitrary subject to (19), we may suppose $r \leq p_v^{\frac{-1}{p_v-1}} R_v^{-1}$. Therefore the constants D_v, ρ_v may be chosen to satisfy (16) as well as the inequality

$$\prod_{v \in T} \rho_v^{d_v} \leq \prod_{v \in T} (p_v^{\frac{1}{p_v-1}} R_v)^{-d_v}. \quad (20)$$

Putting (20) together with the constructed ρ_v for v in S shows that

$$\prod_{v \in S \cup T} \rho_v^{d_v} \leq \prod_{v \in M_K^\infty} (1 + \theta_v)^{d_v} \prod_{v \in S \setminus M_K^\infty} \max\{1, \theta_v\}^{d_v} \prod_{v \in T} (p_v^{\frac{1}{p_v-1}} R_v)^{-d_v}.$$

This inequality shows that (15) implies (17) and concludes the proof. \square

Corollary 2. *Let $s: \mathbb{N} \rightarrow K$ be a function. Suppose that*

- (i) $\delta_m s$ takes values in the ring of integers of K , and
- (ii) for each $v \in M_K^\infty$ there is a positive number θ_v such that $|s(n)|_v \ll \theta_v^n$.

If

$$\prod_{v \in M_K^\infty} (1 + \theta_v)^{d_v} < e^{d(1 + \frac{1}{2} + \dots + \frac{1}{m})}$$

then $s(n)$ is a polynomial in n .

Proof of Corollary 2. Take $S = M_K^\infty$ and $T = \emptyset$ in Theorem 5. \square

Proof of Theorem 1. Take $K = \mathbb{Q}$ in Corollary 2. \square

Remark. It is well-known that a power series $a_0 + a_1x + a_2x^2 + \cdots$ is the expansion of a rational function if and only if there exists a positive integer N such that

$$c(n) := \det(a_{n+i+j})_{i,j=0}^N \quad (21)$$

is zero for all sufficiently large n . Theorem 5 may be applied to the sequence s whose finite differences are given by (21) to obtain a generalization of Theorems 2 and 3 of [3]. We have not emphasized this application however as hypotheses (i) and (iii) of Theorem 5 do not appear to be natural conditions on power series.

6. Integrality of divided differences

In this section we give two interpretations of the integrality of higher divided differences. Our first interpretation generalizes the observation that the first divided difference $\delta_1 s$ of a function $s: \mathbb{N} \rightarrow \mathbb{Q}$ is integer-valued if and only if s preserves all congruences. Roughly speaking, our first interpretation says that a function whose m th divided difference is integral is locally approximated to m th order by polynomials. Here ‘locally’ refers to the topology on \mathbb{N} inherited from its inclusion into the ring of adeles. In this topology, the open neighborhoods are infinite arithmetic progressions and small neighborhoods are infinite arithmetic progressions with highly divisible periods.

If $U \subset \mathbb{N}$ is a subset, $g: U \rightarrow \mathbb{Q}$ is any function, and ε is an integer, we write

$$g(n) = O(\varepsilon)$$

to mean $g(n) \in \varepsilon\mathbb{Z}$ for all $n \in U$. For example, $g = O(1)$ if and only if g is integer-valued.

Proposition 2. Suppose $s: \mathbb{N} \rightarrow \mathbb{Q}$ is a function whose m th divided difference is integer-valued. Let $U = n_0 + \mathbb{N}\varepsilon \subset \mathbb{N}$ be any infinite arithmetic progression. Then there is a polynomial $f(x) \in \mathbb{Q}[x]$ of degree less than m such that

$$s|_U(n) = f|_U(n) + O(\varepsilon^m).$$

Furthermore, the denominators of the coefficients of $f(x)$ are divisors of

$$C = \prod_{p \text{ prime}} \sup \left\{ \left| \frac{c(0)}{0!} \right|_p, \left| \frac{c(1)}{1!} \right|_p, \dots, \left| \frac{c(m-1)}{(m-1)!} \right|_p, 1 \right\}$$

where $c(n) := \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} s(k)$, and

$$\gcd(f(n) : n \in U) = \gcd(\varepsilon^k k! \cdot \delta_k s(n_0, n_1, \dots, n_k) : 0 \leq k < m).$$

Proof. We make use of a classical interpolation formula due to Newton (cf. e.g. [7], §1). Let $\{n_0, n_1, \dots\} \subset \mathbb{Q}$ be an infinite subset and let $s : \{n_0, n_1, \dots\} \rightarrow \mathbb{Q}$ be a function. Then for all $n \in \{n_0, n_1, \dots\}$,

$$s(n) = s(n_0) + \sum_{k=1}^{\infty} \delta_k s(n_0, n_1, \dots, n_k) \prod_{j=0}^{k-1} (n - n_j) \quad (\text{finite sum}). \quad (22)$$

Set $n_k := n_0 + k\varepsilon$ for $k \in \mathbb{N}$. From (22) we obtain that

$$s(n) = f(n) + \sum_{k=m}^{\infty} \delta_k s(n_0, n_1, \dots, n_k) \prod_{j=0}^{k-1} (n - n_j) \quad \text{for all } n \in U \quad (23)$$

where $f(x) := s(n_0) + \sum_{k=1}^{m-1} \delta_k s(n_0, n_1, \dots, n_k) \prod_{j=0}^{k-1} (x - n_j) \in \mathbb{Q}[x]$.

We claim that for any integer $k \geq m$ and $a, a_0, a_1, \dots \in U$,

$$\delta_k s(a_0, a_1, \dots, a_k) \prod_{j=0}^{k-1} (a - a_j)$$

is integral and divisible by ε^m . We proceed by induction on k . The case $k = m$ is true by assumption. Suppose $k > m$. By the recursive formula for divided differences (cf. [7], §1),

$$\delta_k s(a_0, \dots, a_k) = (a_1 - a_0)^{-1} (\delta_{k-1} s(\widehat{a_0}, a_1, \dots, a_k) - \delta_{k-1} s(a_0, \widehat{a_1}, \dots, a_k))$$

(the $\widehat{}$ indicates an omission). Then

$$\begin{aligned} \delta_k s(a_0, a_1, \dots, a_k) \prod_{j=0}^{k-1} (a - a_j) = \\ \left(\frac{a - a_0}{a_1 - a_0} \right) \left\{ \delta_{k-1} s(\widehat{a_0}, a_1, \dots, a_k) \prod_{j \in \{1, \dots, k-1\}} (a - a_j) \right\} \\ - \left(\frac{a - a_1}{a_1 - a_0} \right) \left\{ \delta_{k-1} s(a_0, \widehat{a_1}, \dots, a_k) \prod_{j \in \{0, 2, \dots, k-1\}} (a - a_j) \right\}. \end{aligned}$$

By the inductive hypothesis, and the fact that $\delta_{k-1} s$ is a symmetric function, each of the two terms in $\{\cdot\}$ -brackets is integral and divisible by ε^m . The remaining two factors are integers, and we have proven the claim. It follows from (23) that $s(n) - f(n)$ is integral and divisible by ε^m for any $n \in U$.

The explicit bound on the denominators of f follows from Corollary 1. For the last claim, it is well-known that the binomial polynomials $\{\binom{y}{j}\}_{j \geq 0}$ form an orthonormal Banach basis of $C(\mathbb{N}, \mathbb{Q}_p)$ for all p . Observe that

$$f(x) = \tilde{f}(y) = s(n_0) + \sum_{k=1}^{m-1} \varepsilon^k k! \cdot \delta_k s(n_0, n_1, \dots, n_k) \binom{y}{j} \quad \text{where } y := \frac{x - n_0}{\varepsilon}.$$

Therefore

$$\begin{aligned} \gcd(f(n) : n \in U) &= \gcd(\tilde{f}(n) : n \in \mathbb{N}) \\ &= \gcd(\varepsilon^k k! \cdot \delta_k s(n_0, n_1, \dots, n_k) : 0 \leq k < m). \quad \square \end{aligned}$$

Integrality of $\delta_1 s$ implies that for all primes p and integers $m, n \in \mathbb{N}$,

$$|s(m) - s(n)|_p \leq |m - n|_p.$$

In other words, $\delta_1 s$ is \mathbb{Z} -valued if and only if s is Lipschitz continuous with Lipschitz constant 1 for every p -adic metric on \mathbb{N} . Our second interpretation of the integrality of higher divided differences generalizes this observation.

Proposition 3. *Let $s : \mathbb{N} \rightarrow \mathbb{Q}_p$ be a function and let m be a positive integer. Suppose that $\|\delta_m s\|_p \leq 1$. Then s extends to an element f of $C^{m-1}(\mathbb{Z}_p, \mathbb{Q}_p)$ and $f^{(m-1)}$ is Lipschitz continuous with Lipschitz constant $|(m-1)!|_p$.*

Here we mean *strict* differentiability (cf. [9], §29).

Proof. For $m = 1$ this was already explained just before the proposition. Assume $m \geq 2$. By the recursive formula for divided differences (cf. [7], §1),

$$|\delta_{m-1} s(a_0, \dots, a_{m-1}) - \delta_{m-1} s(a_1, \dots, a_m)| \leq |a_0 - a_m|$$

for all $(a_0, \dots, a_m) \in \Delta_m$. Thus,

$$|\delta_{m-1} s(a_i; a_{ij}) - \delta_{m-1} s(a_j; a_{ij})| \leq |a_i - a_j| \quad (24)$$

where $0 \leq i < j \leq m$ and $a_{ij} := (a_0, \dots, \widehat{a_i}, \dots, \widehat{a_j}, \dots, a_m) \in \Delta_{m-2}$ (the $\widehat{}$ indicates an omission).

We equip \mathbb{Z}_p^m with the metric given by

$$d(x, y) := \max_{1 \leq i \leq m} |x_i - y_i|_p.$$

We will show that $\delta_{m-1} s$ is Lipschitz continuous for this metric on the dense subset $\Delta_{m-1} \subset \mathbb{Z}_p^m$. By a limiting argument, it suffices to show the Lipschitz condition for

$\delta_{m-1}s$ on elements $x = (x_1, \dots, x_m)$ and $y = (y_1, \dots, y_m)$ in Δ_{m-1} such that $x \cap y = \emptyset$. Define $z^0 := x$, $z^m := y$, and

$$z^i := (x_1, x_2, \dots, x_{m-i}, y_{m-i+1}, \dots, y_m) \in \mathbb{Z}_p^m \quad \text{for } 0 < i < m.$$

Then $d(z^i, z^{i+1}) = |x_{m-i} - y_{m-i}|_p$ for $0 \leq i < m$. Since z^i and z^{i+1} differ by only one coordinate and $x \cap y = \emptyset$, we may apply (24). We conclude that

$$|\delta_{m-1}s(z^i) - \delta_{m-1}s(z^{i+1})|_p \leq |x_{m-i} - y_{m-i}|_p = d(z^i, z^{i+1})$$

for $0 \leq i < m$. Using $\delta_{m-1}s(z^0) - \delta_{m-1}s(z^m) = \sum_{i=0}^{m-1} (\delta_{m-1}s(z^i) - \delta_{m-1}s(z^{i+1}))$ shows that

$$\begin{aligned} |\delta_{m-1}s(x) - \delta_{m-1}s(y)|_p &\leq \max_{0 \leq i < m} |\delta_{m-1}s(z^i) - \delta_{m-1}s(z^{i+1})|_p \\ &\leq \max_{0 \leq i < m} |x_{m-i} - y_{m-i}|_p \\ &= d(x, y). \end{aligned}$$

We have shown that $\delta_{m-1}s$ is Lipschitz continuous on a dense subset of \mathbb{Z}_p^m , so there is a unique Lipschitz continuous extension $\widetilde{\delta_{m-1}s}: \mathbb{Z}_p^m \rightarrow \mathbb{C}_p$ of $\delta_{m-1}s$. Therefore s extends to a $(m-1)$ -times strictly differentiable function $f: \mathbb{Z}_p \rightarrow \mathbb{Q}_p$ satisfying

$$f^{(m-1)}(a) = (m-1)! \cdot \widetilde{\delta_{m-1}s}(x_a)$$

where $x_a := (a, a, \dots, a)$ (cf. [9], §29). By Lipschitz continuity of $\widetilde{\delta_{m-1}s}$ we get that

$$|f^{(m-1)}(a) - f^{(m-1)}(b)| \leq |(m-1)!|_p d(x_a, x_b) = |(m-1)!|_p |a - b|_p. \quad \square$$

Corollary. *Let $s: \mathbb{N} \rightarrow \mathbb{Q}$ be a function and suppose that $\delta_{m+1}s$ is \mathbb{Z} -valued for a non-negative integer m . Then for every prime p , the function s extends to an m -times strictly differentiable function $f_p: \mathbb{Z}_p \rightarrow \mathbb{Q}_p$ such that $f_p^{(m)}$ is Lipschitz continuous with Lipschitz constant $|m!|_p$.*

Acknowledgments

The author is grateful to Kartik Prasanna for a suggestion which simplified the proof of Theorem 2. The author is also grateful to the anonymous referee for several valuable suggestions. The author would also like to thank Angus Chung and Trevor Hyde for helpful conversations.

References

- [1] Y. Amice, Interpolation p -adique, *Bull. Soc. Math. Fr.* 92 (1964) 117–180.
- [2] M. Bhargava, On P -orderings, rings of integer-valued polynomials, and ultrametric analysis, *J. Am. Math. Soc.* 22 (4) (2009) 963–993.
- [3] B. Dwork, On the rationality of the zeta function of an algebraic variety, *Am. J. Math.* 82 (1960) 631–648.
- [4] D.L. Hilliker, E.G. Straus, Some p -adic versions of Polya's theorem on integer valued analytic functions, *Proc. Am. Math. Soc.* 26 (1970) 395–400.
- [5] K. Iwasawa, Lectures on p -adic L -Functions, *Annals of Mathematics Studies*, vol. 74, Princeton University Press/University of Tokyo Press, Princeton, N.J./Tokyo, 1972.
- [6] K. Mahler, An interpolation series for continuous functions of a p -adic variable, *J. Reine Angew. Math.* 199 (1958) 23–34.
- [7] L.M. Milne-Thomson, *The Calculus of Finite Differences*, Macmillan and Co., Ltd., London, 1951.
- [8] J.B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Ill. J. Math.* 6 (1962) 64–94.
- [9] W.H. Schikhof, *Ultrametric Calculus*, *Cambridge Studies in Advanced Mathematics*, vol. 4, Cambridge University Press, Cambridge, 2006.