p-ADIC HAHN SERIES WITH SPARSE SUPPORT

by

Shanwen Wang & Yijun Yuan

Abstract. — Let p be a prime number. We introduce a sparseness condition on the supports of p-adic Hahn series, and prove that this condition implies transcendence over $\breve{\mathbf{Q}}_p$, the completed maximal unramified extension of $\mathbf{Q}_p$. As an application, we prove a p-adic analogue of a theorem of Huang and Ştefănescu on the algebraicity of a special family of Hahn series. This yields infinitely many counterexamples to a question concerning the correspondence between $\mathbf{F}_p$-algebraicity of equal-characteristic Hahn series and $\mathbf{Q}_p$-algebraicity of p-adic Hahn series.

All results in this paper have been fully formalized in Lean 4 with Mathlib.

Contents

1. Introduction 1
2. Preliminaries on Hahn series 5
3. Sparseness and (c, n)-sparseness 6
4. T-scaled realization of $\mathbf{L}_p$ 10
5. Main theorem and its applications 12
Appendix A. AI-assisted formalization in Lean 4 16
References 21

1. Introduction

Let p be a prime number. Let $\mathbf{F}_p$ be the finite field of p elements, $\overline{\mathbf{F}}_p$ be an algebraic closure of $\mathbf{F}_p$. Let $\mathbf{Q}_p$ be the field of p-adic numbers, $\breve{\mathbf{Q}}_p = W(\overline{\mathbf{F}}_p)[p^{-1}]$ be the completed maximal unramified extension of $\mathbf{Q}_p$, $\breve{\mathbf{Z}}_p = W(\overline{\mathbf{F}}_p)$ be the ring of integers of $\breve{\mathbf{Q}}_p$, $\overline{\mathbf{Q}}_p$ be an algebraic closure of $\mathbf{Q}_p$, and $\mathbf{C}_p = \widehat{\overline{\mathbf{Q}}}_p$ be the field of p-adic complex numbers. We normalize the p-adic valuation on $\mathbf{C}_p$ and its subfields by $v_p(p) = 1$. For any set X, we denote by $\mathrm{card}(X)$ the cardinality of X.

2020 ***Mathematics Subject Classification.*** — 11J61, 11J81, 11D88, 68V20.

Key words and phrases. — p-adic transcendence, p-adic Hahn series, sparse support, formalization.

1.1. p-adic transcendence via p-adic Hahn series.

— A p-adic Hahn series is a generalized formal power series of the form $f = \sum_{q\in\mathbf{Q}}[f(q)]p^q$, where $f(q) \in \overline{\mathbf{F}}_p$ and $[\cdot]$ is the Teichmüller lift, whose support $\mathrm{Supp}(f) = \{q \in \mathbf{Q} \mid f(q) \neq 0\}$ is a well-ordered subset of $\mathbf{Q}$. Krull, Lampert, and Poonen showed that the set of p-adic Hahn series, which we denote by $\mathbf{L}_p$, forms the spherical completion of $\overline{\mathbf{Q}}_p$. In particular, $\mathbf{L}_p$ is algebraically closed and complete with respect to the p-adic valuation given by $f \mapsto \min \mathrm{Supp}(f)$.

The field $\mathbf{L}_p$ provides a natural setting in which to study transcendental number theory over p-adic fields. More precisely, a fundamental question arises:

Question 1.1. — *Given a p-adic Hahn series $f \in \mathbf{L}_p$, how can one determine whether f is algebraic over $\mathbf{Q}_p$?*

Although this question remains open in general, several necessary conditions for a p-adic Hahn series to be algebraic over $\mathbf{Q}_p$ are known:

1. In [5] and [6], Lampert and Poonen proved that if $f \in \mathbf{L}_p$ is algebraic over $\mathbf{Q}_p$, then
   (a) there exists an integer T such that $\mathrm{Supp}(f) \subset \frac{1}{T}\mathbf{Z}[1/p]$;
   (b) there exists a finite extension $\mathbf{F}_q$ of $\mathbf{F}_p$ such that $\{f(q)\}_{q\in\mathbf{Q}} \subset \mathbf{F}_q$.
   These conditions are also studied quantitatively in [8].
2. In [5], Lampert also proved that if $f \in \mathbf{L}_p$ is algebraic over $\mathbf{Q}_p$, then the accumulation points of $\mathrm{Supp}(f)$ are rational numbers.
3. In [4] and [3], Kedlaya gives a necessary and sufficient condition for an equal-characteristic Hahn series in $\mathbf{L}_p^\flat := \overline{\mathbf{F}}_p(\!(t^{\mathbf{Q}})\!)$ to lie in the algebraic closure $\overline{\mathbf{F}}_p(\!(t)\!)^{\mathrm{alg}}$ of $\overline{\mathbf{F}}_p(\!(t)\!)$, phrased in the language of automata theory. As an application, Kedlaya uses Witt vectors to lift this result to the p-adic case: he shows that the field $\mathbf{C}_p$, when viewed as a subfield of $\mathbf{L}_p$, coincides with the completion of the set $\Theta\left(\overline{\mathbf{F}}_p(\!(t)\!)^{\mathrm{alg},\wedge}\right)$, where $\overline{\mathbf{F}}_p(\!(t)\!)^{\mathrm{alg},\wedge}$ is the t-adic completion of $\overline{\mathbf{F}}_p(\!(t)\!)^{\mathrm{alg}}$ and $\Theta\colon \overline{\mathbf{F}}_p(\!(t^{\mathbf{Q}})\!) \longrightarrow \mathbf{L}_p$ is the map $\sum_{q\in\mathbf{Q}} f(q)t^q \longmapsto \sum_{q\in\mathbf{Q}}[f(q)]p^q$.

Intuitively, Kedlaya's result indicates that the algebraicity of equal-characteristic Hahn series in $\overline{\mathbf{F}}_p(\!(t^{\mathbf{Q}})\!)$ and the algebraicity of p-adic Hahn series in $\mathbf{L}_p$ are related to some extent via the map Θ, which leads to the following questions:

Question 1.2. —

1. *Suppose that $f \in \mathbf{L}_p^\flat$ is algebraic over $\mathbf{F}_p(\!(t)\!)$. Is $\Theta(f)$ algebraic over $\mathbf{Q}_p$?*
2. *Suppose that $f \in \mathbf{L}_p$ is algebraic over $\mathbf{Q}_p$. Is $\Theta^{-1}(f)$ algebraic over $\mathbf{F}_p(\!(t)\!)$?*

In [9], we gave a negative answer to the first question by showing that the p-adic Hahn series $\mathfrak{A} := \sum_{k=1}^{\infty} p^{-1/p^k}$ is transcendental over $\mathbf{Q}_p$, whereas its preimage under Θ is a root of the polynomial $X^p - X - t^{-1}$ over $\mathbf{F}_p(\!(t)\!)$.

1.2. Sparseness and main theorem.

— The key observation in [9] is that the support of $\mathfrak{A}$ is "sparse", in the sense that for any non-zero polynomial $P(X) \in \mathbf{Q}_p[X]$, the multinomial expansion of $P(\mathfrak{A})$ contains orphaned terms that cannot be cancelled by other terms, which forces $P(\mathfrak{A})$ to be nonzero. In this article, we generalize this idea to the following combinatorial definitions:

Definition 1.3. —

1. *Any rational number q can be uniquely written in the form $q = w + \sum_{i=1}^{\infty} q_i \cdot p^{-i}$, where $w \in \mathbf{Z}$, $q_i \in \{0, 1, \ldots, p-1\}$ and $q_i \neq p-1$ for infinitely many i. We call $\mathfrak{N}_p(q) := \sum_{i=1}^{\infty} q_i \in \mathbf{N} \cup \{\infty\}$ the **p-digit sum** of q.*

2. *For any set S of rational numbers, define the* ***dominant*** p***-digit sum*** *of S to be*
$$\mathrm{dom}_p(S) := \sup\{\mathfrak{N}_p(q) | q \in S\} \in \mathbf{N} \cup \{\infty\},$$
and define the p***-digit dominant part*** *of S to be the following subset of S:*
$$\mathrm{Dom}_p(S) := \{q \in S | \mathfrak{N}_p(q) = \mathrm{dom}_p(S)\}.$$

Definition 1.4. — *A set $S \subset [0,1) \cap \mathbf{Q}$ is* ***sparse*** *if:*

1. $\mathrm{dom}_p(S)$ *is finite;*
2. *for infinitely many integers $n \geq 1$, there exist n elements $d_1, \cdots, d_n \in \mathrm{Dom}_p(S)$ such that*
   (a) *there is no carry in base p when adding $d_1, \cdots, d_n$ together;*
   (b) *if $e_1, \cdots, e_n \in S$ satisfy that $d_1 + \cdots + d_n$ and $e_1 + \cdots + e_n$ differ by an integer, then up to a permutation of $\{1, 2, \cdots, n\}$, $d_i = e_i$ for every $i = 1, 2, \cdots, n$.*

Remark 1.5. — *We give some comments on Definition 1.4:*

1. *The first condition is closely related to Kedlaya's criterion for the algebraicity of equal-characteristic Hahn series: he shows that if $f \in \mathbf{L}_p^\flat$ is algebraic over $\overline{\mathbf{F}}_p(\!(t)\!)$, then* $\mathrm{dom}_p(-T \cdot \mathrm{Supp}(f))$ *is finite for some integer $T \geq 1$.*
2. *Condition (2a) indicates that the addition $d_1 + \cdots + d_n$, no matter how one parenthesizes it, behaves like the addition in the free commutative monoid $\bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$.*
3. *Condition (2b) is a combinatorial rigidity condition asserting that the chosen carry-free sum admits a unique decomposition modulo $\mathbf{Z}$. This produces the orphaned exponents in the multinomial expansion, which is the key mechanism behind the transcendence proof.*

The following example illustrates a typical situation in which the sparseness condition is satisfied:

Example 1.6 **(cf. Example 3.7).** — *Let $A_1, A_2, \cdots$ be a family of pair-wise disjoint nonempty subsets of $\mathbf{Z}_{\geq 1}$ such that $\sup_i \mathrm{card}(A_i) < \infty$, and this supremum is attained by infinitely many i. Then the set*
$$\left\{ \sum_{r \in A_i} p^{-r} \,\middle|\, i = 1, 2, \cdots \right\}$$
is sparse.

The main theorem of this article is the following:

Theorem 1.7 **(cf. Theorem 5.3).** — *Let $f \in \mathbf{L}_p$ be a p-adic Hahn series such that there exists an integer $T \geq 1$ for which $-T \cdot \mathrm{Supp}(f)$ admits a sparse set $W \neq \{0\}$ of representatives modulo $\mathbf{Z}$. Then f is transcendental over $\breve{\mathbf{Q}}_p$, and hence over $\mathbf{Q}_p$.*

Remark 1.8. — *The sparseness condition does not involve the coefficients of f, so it is natural that it does not distinguish algebraicity over $\mathbf{Q}_p$ from algebraicity over $\breve{\mathbf{Q}}_p$.*

1.3. p-adic analogue of a result of Huang and Ştefănescu.

— As mentioned in [4, Section 1], a prototype of Kedlaya's criterion for the $\overline{\mathbf{F}}_p$-algebraicity of Hahn series in $\mathbf{L}_p^\flat$ is the following result of Huang, which was independently discovered by Ştefănescu:

Theorem 1.9 **(cf. [1, 7]).** — *Let $f = \sum_{i=1}^{\infty} f(i) \cdot t^{-1/p^i} \in \mathbf{L}_p^\flat$ be a Hahn series. Then the following are equivalent:*

1. *the series f is algebraic over $\mathbf{F}_p(\!(t)\!)$;*
2. *the series f is algebraic over $\overline{\mathbf{F}}_p(\!(t)\!)$;*

3. *the sequence* $\{f(i)\}_{i\geq 1}$ *is eventually periodic.*

With the help of Theorem 1.7 (as well as Example 1.6), we prove the following p-adic analogue of Theorem 1.9:

***Theorem 1.10* (cf. Corollary 5.5).** — *Let* $f=\sum_{i=1}^{\infty}[f(i)]\cdot p^{-1/p^i}\in\mathbf{L}_p$ *be a* p*-adic Hahn series. Then the following are equivalent:*

1. *the series* f *is algebraic over* $\mathbf{Q}_p$*;*
2. *the series* f *is algebraic over* $\breve{\mathbf{Q}}_p$*;*
3. $f(i)=0$ *for all but finitely many* i*.*

Remark 1.11. — *Theorem 1.10 yields infinitely many counter-examples to the first question in Question 1.2: any Hahn series* $f=\sum_{i=1}^{\infty}f(i)\cdot t^{-1/p^i}\in\mathbf{L}_p^\flat$ *for which* $f(i)\neq 0$ *for infinitely many* i *and the sequence* $\{f(i)\}_{i\geq 1}$ *is eventually periodic is algebraic over* $\mathbf{F}_p(\!(t)\!)$ *by Theorem 1.9, while* $\Theta(f)$ *is transcendental over* $\mathbf{Q}_p$ *by Theorem 1.7.*

1.4. Formalization in Lean. — Considering the highly combinatorial nature of the sparseness condition and the transcendence proof, **we formalize all results of this article in Lean 4**, a popular proof assistant based on dependent type theory, with the help of the agentic autoformalization system Archon (cf. [2]) developed by AI4Math team at BICMR, Peking University.

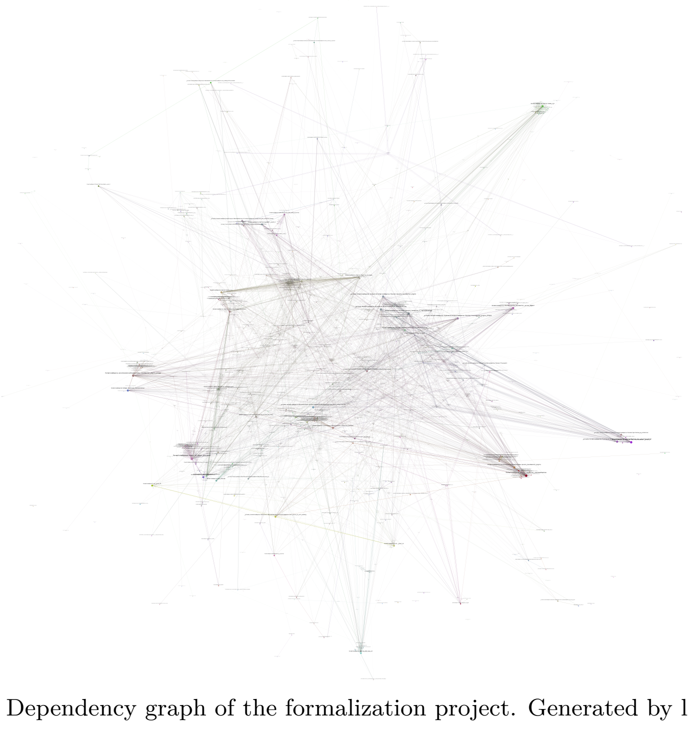

Figure 1. Dependency graph of the formalization project. Generated by lean-graph.

The formalization is available at `https://github.com/YijunYuan/FormalizedSparse`, and we refer the reader to Appendix A for a detailed discussion on the formalization process.

Acknowledgements. — The authors would like to thank Wanying He and Jiedong Jiang for their help on using Archon. The research is partially supported by the National Key R&D Program of China (Grant No. 2024YFA1014000).

2. Preliminaries on Hahn series

To keep this article self-contained, we briefly recall some basic facts about Hahn series.

Definition 2.1 **([6, Section 3]).** — *Let R be a commutative ring and G be an ordered group.*

1. *For any $f \in \mathrm{Hom}_{\mathrm{Set}}(G, R)$, we define the* ***support*** *of f to be*
$$\mathrm{Supp}(f) = \{g \in G \colon f(g) \neq 0\}.$$
2. *Define the set of* ***Hahn series*** *over R with value group G to be*
$$R(\!(G)\!) := \{f \in \mathrm{Hom}_{\mathrm{Set}}(G, R) \colon \mathrm{Supp}(f) \text{ is well-ordered}\}.$$
By introducing a formal variable t, elements in $R(\!(G)\!)$ will also be written as $\sum_{g\in G} r_g t^g$, where $r_g \in R$ for all $g \in G$.

Proposition 2.2 **([6, Lemma 1, Corollary 2]).** — *Let R be a commutative ring and G be an ordered group.*

1. *With identity $1 \cdot t^0$ and addition as well as multiplication given by*
$$\sum_{g\in G} a_g t^g + \sum_{g\in G} b_g t^g := \sum_{g\in G} (a_g + b_g) t^g, \ \sum_{g\in G} a_g t^g \cdot \sum_{g\in G} b_g t^g := \sum_{g\in G} \left(\sum_{h\in G} a_h b_{g-h} \right) t^g,$$
$R(\!(G)\!)$ forms a commutative ring.
2. *If R is a field, then so is $R(\!(G)\!)$. Moreover, with the map*
$$v \colon R(\!(G)\!) \longrightarrow G \cup \{\infty\}, \ f \longmapsto \begin{cases} \min \mathrm{Supp}(f), & \text{if } f \neq 0 \\ \infty, & \text{if } f = 0 \end{cases},$$
$R(\!(G)\!)$ becomes a valued field with value group G and residue field R.

Since $\mathrm{char}\, R(\!(G)\!) = \mathrm{char}\, R$, we call $R(\!(G)\!)$ the **equal-characteristic field of Hahn series** over R with value group G, also denoted by $R(\!(t^G)\!)$ with respect to the formal variable t.

Proposition 2.3 **([6, Proposition 3, Corollary 3, Proposition 5])**

Let k be a perfect field of characteristic p and G be an ordered group containing $\mathbb{Z}$ as a subgroup. Besides that, let
$$\mathcal{N} := \left\{ \sum_{g\in G} r_g t^g \in W(k)(\!(t^G)\!) \colon \text{for every } g \in G, \ \sum_{n\in\mathbb{Z}} r_{g+n} p^n = 0 \right\},$$
where $W(k)$ is the ring of Witt vectors of k. We call elements in $\mathcal{N}$ the ***null series*** *of $W(k)(\!(t^G)\!)$. Then*

1. *$\mathcal{N}$ is a maximal ideal of $W(k)(\!(t^G)\!)$, which makes $W(k)(\!(p^G)\!) := W(k)(\!(t^G)\!)/\mathcal{N}$ a field⁽¹⁾, called the* ***field of p-adic Hahn series****.*

---

(1) Informally, $W(k)(\!(p^G)\!)$ is obtained by replacing the formal variable t in elements of $W(k)(\!(t^G)\!)$ by the prime p.

2. *Every element x in $W(k)(\!(p^G)\!)$ can be uniquely written as*

$$x = \sum_{g \in G} [r_g] p^g,$$

*where $r_g \in k$ for all $g \in G$ and $[\cdot]\colon k \longrightarrow W(k)$ is the Teichmüller lift. We call this the **standard expansion** of the element x.*

3. *For $f = \sum_{g \in G} [r_g] p^g \in W(k)(\!(p^G)\!)$, define the **support** of f to be*

$$\operatorname{Supp}(f) = \{g \in G \colon r_g \neq 0\}.$$

Then the map

$$v\colon W(k)(\!(G)\!)/\mathcal{N} \longrightarrow G \cup \{\infty\},\ f \mapsto \begin{cases} \min \operatorname{Supp}(f), & \text{if } f \neq 0 \\ \infty, & \text{if } f = 0 \end{cases}$$

makes $W(k)(\!(G)\!)/\mathcal{N}$ a mixed-characteristic valued field with value group G and residue field k.

The most fundamental property of the field of Hahn series is the following:

***Theorem 2.4* (cf. [6, Theorem 1, Corollary 4, Corollary 6])**

Let F be an equal-characteristic (resp. mixed-characteristic) valued field with divisible value group G and algebraically closed residue field k. Then the equal-characteristic (resp. p-adic) field of Hahn series $k(\!(t^G)\!)$ (resp. $W(k)(\!(p^G)\!)$) is the unique (up to isomorphism of valued fields) minimal spherically complete extension of F. Moreover, it is algebraically closed and complete.

The following are the fields of Hahn series used in this article:

Example 2.5. — *Let $F = \overline{\mathbf{F}}_p(\!(t)\!)$ (resp. $\breve{\mathbf{Q}}_p$), which has value group $\mathbf{Q}$ and residue field $\overline{\mathbf{F}}_p$. Then the field of equal-characteristic (resp. p-adic) Hahn series $\mathbf{L}_p^\flat := \overline{\mathbf{F}}_p(\!(t^{\mathbf{Q}})\!)$ (resp. $\mathbf{L}_p := \breve{\mathbf{Z}}_p(\!(p^{\mathbf{Q}})\!) = W(\overline{\mathbf{F}}_p)(\!(p^{\mathbf{Q}})\!)$) is the spherical completion of F with the same residue field and value group, and is algebraically closed and complete.*

For brevity, we call elements of $\mathbf{L}_p$ p-adic Hahn series without specifying the residue field and value group.

3. Sparseness and (c, n)-sparseness

To give a rigorous and workable formulation of the sparseness condition in Definition 1.4, we introduce several combinatorial notions and auxiliary functions related to base-p expansions of rational numbers in $[0, 1)$.

Definition 3.1. —

1. *For $\underline{d} \in \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$, let $\Psi(\underline{d}) := \sum_{i=1}^{\infty} d_i \in \mathbf{N}$ and $\|\underline{d}\| := \sum_{i=1}^{\infty} d_i p^{-i} \in \mathbf{Q}_{\geq 0}$.*
2. *Let $\mathbb{P} := \bigoplus_{\mathbf{Z}_{\geq 1}} \{0, 1, \cdots, p-1\} \subsetneq \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$.*
3. *For $\underline{d} \in \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$ and $i \in \mathbf{Z}_{\geq 1}$, denote by $\underline{d}_i$ the i-th component of $\underline{d}$.*

Lemma 3.2. — *For any $\underline{d} \in \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$, there exists a unique element $\tau(\underline{d}) \in \mathbb{P}$ such that $\|\underline{d}\| - \|\tau(\underline{d})\| \in \mathbf{Z}$.*

Proof. — If one writes $\|\underline{d}\|$ as a decimal expansion[2] in base p:

$$\|\underline{d}\| = w.d_1 \cdots d_n \cdots ,$$

where $w \in \mathbf{Z}$ and $d_i \in \{0, 1, \cdots, p-1\}$ for any i, then $\tau(\underline{d}) = 0.d_1 \cdots d_n \cdots$. The uniqueness is trivial. ☐

We collect several properties of these concepts in the following lemma:

Lemma 3.3. — *Let $\underline{d}, \underline{e} \in \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$.*

1. *The maps $\Psi(\cdot)$ and $\|\cdot\|$ are additive. Moreover, $\|\cdot\|$ is injective when restricted to $\mathbb{P}$.*
2. *One has $\tau(\underline{d}) = \tau(\underline{e})$ if and only if $\|\underline{d}\| - \|\underline{e}\| \in \mathbf{Z}$.*
3. *$\underline{d} \in \mathbb{P}$ if and only if $\tau(\underline{d}) = \underline{d}$.*
4. *One has $\Psi(\tau(\underline{d})) \leq \Psi(\underline{d})$, with equality if and only if $\underline{d} \in \mathbb{P}$.*

Proof. — The first three statements are straightforward. We only prove the last statement.

For any $\underline{d} \in \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$, let

$$N(\underline{d}) := \begin{cases} 0, & \text{if } \underline{d} \in \mathbb{P}; \\ \max\{i | d_i \geq p\}, & \text{otherwise.} \end{cases}$$

By a descent argument on $N(\underline{d})$, it suffices to show that if $N(\underline{d}) \geq 1$, then there exists $\underline{e} \in \bigoplus_{\mathbf{Z}_{\geq 1}} \mathbf{N}$ such that $\|\underline{d}\| - \|\underline{e}\| \in \mathbf{Z}$, $\Psi(\underline{e}) < \Psi(\underline{d})$, and $N(\underline{e}) < N(\underline{d})$.

Writing $d_{N(\underline{d})} = p{\cdot}r + s$ with $r \in \mathbf{N}$ and $s \in \{0, 1, \cdots, p-1\}$, we take $\underline{e} := (d_0, \cdots, d_{N(\underline{d})-2}, d_{N(\underline{d})-1} + r, s, 0, \cdots)$. Then

$$\Psi(\underline{d}) - \Psi(\underline{e}) = d_{N(\underline{d})} - r - s = (p-1) \cdot r > 0$$

and $\|\underline{d}\| - \|\underline{e}\| = 0$. The result follows. ☐

Remark 3.4. — *For a rational number q in $[0, 1)$ with a finite-length decimal expansion in base p, the preimage $\underline{q} \in \mathbb{P}$ of q under the map $\|\cdot\|$ extracts the digits of q in base p, and $\Psi(\underline{q})$ is the p-digit sum of q.*

Definition 3.5. — *Let p be a prime number. Let $c, n \geq 1$ be integers. A subset $S \subset \mathbb{P}$ is (c, n)-**sparse** if*

1. *there exists $c \geq 1$ such that $\Psi(\underline{d}) \leq c$ for every $\underline{d} \in S$;*
2. *there exist n (not necessarily distinct) elements $\underline{d}^{(1)}, \underline{d}^{(2)}, \cdots, \underline{d}^{(n)} \in S$ such that*
   - *(a) $\Psi\left(\underline{d}^{(i)}\right) = c$ for $i = 1, 2, \cdots, n$ and $\sum_{i=1}^{n} \underline{d}_j^{(i)} < p$ for $j \in \mathbf{N}$ (that is, $\sum_{i=1}^{n} \underline{d}^{(i)} \in \mathbb{P}$);*
   - *(b) if $\underline{e}^{(1)}, \underline{e}^{(2)}, \cdots, \underline{e}^{(n)} \in S$ satisfy $\left\|\sum_{i=1}^{n} \underline{d}^{(i)}\right\| - \left\|\sum_{i=1}^{n} \underline{e}^{(i)}\right\| \in \mathbf{Z}$, then up to a permutation of $\{1, 2, \cdots, n\}$, $\underline{d}^{(i)} = \underline{e}^{(i)}$ for every $i = 1, 2, \cdots, n$.*

The following lemma reformulates the sparseness condition in Definition 1.4 in terms of the (c, n)-sparseness condition in Definition 3.5.

Lemma 3.6. — *A subset $W \neq \{0\}$ of $[0, 1) \cap \mathbf{Q}$ is sparse in the sense of Definition 1.4 if and only if there exists a subset S of $\mathbb{P}$ such that $W = \|S\|$ and there exists an integer $c \geq 1$ such that S is (c, n)-sparse for infinitely many integers $n \geq 1$.*

---

[2] We do not allow infinite strings of $(p-1)$ in the decimal expansion of $\|\underline{d}\|$ to ensure the uniqueness of $\tau(\underline{d})$.

Proof. — Suppose that $W \subset [0,1) \cap \mathbf{Q}$ is sparse. Then the p-digit sum of every element in W is finite and bounded by $\mathrm{dom}_p(W)$. Thus we may take

$$S \coloneqq \left\{ (q_i)_{i \in \mathbf{Z}_{\geq 1}} | q = \sum_{i=0}^{\infty} q_i p^{-i} \in W,\ q_i \in \{0, \cdots, p-1\} \right\}.$$

Then S is $(\mathrm{dom}_p(W), n)$-sparse for infinitely many integers $n \geq 1$.

The converse direction follows by a similar argument. ☐

We give a concrete example of the sparse set, which is the prototype of the situation in which the sparseness condition is satisfied:

Example 3.7. — *Let* $\underline{A} \coloneqq (A_i)_{i \geq 1}$ *be a family of pair-wise disjoint nonempty subsets of* $\mathbf{Z}_{\geq 1}$ *such that* $\sup_i |A_i| < \infty$ *and this supremum is attained by infinitely many* i*. Then the set*

$$M(\underline{A}) \coloneqq \left\{ \sum_{r \in A_i} p^{-r} \middle| i = 1, 2, \cdots \right\} \subset [0,1) \cap \mathbf{Q}$$

is sparse.

Proof. — For any i, one has $\mathfrak{N}_p(\sum_{r \in A_i} p^{-r}) = |A_i|$, and consequently $\mathrm{dom}_p(M(\underline{A})) = \sup_i |A_i| < \infty$. The infinity of attainment of the supremum implies that $\mathrm{Dom}_p(M(\underline{A}))$ is an infinite set:

$$\mathrm{Dom}_p(M(\underline{A})) = \left\{ \mathbf{d}_j \coloneqq \sum_{r \in A_{i_j}} p^{-r} \middle| j = 1, 2, \cdots \right\},$$

with $\mathbf{d}_k \neq \mathbf{d}_l$ for any $k \neq l$.

Take $e_1, \cdots, e_n \in M(\underline{A})$ such that $\sum_{j=1}^n e_j - \sum_{j=1}^n \mathbf{d}_j \in \mathbf{Z}$. Since $M(\underline{A}) \subset \|\mathbb{P}\|$, we may write $e_j = \left\|\underline{e}^{(j)}\right\|$ and $\mathbf{d}_j = \left\|\underline{\mathbf{d}}^{(j)}\right\|$ with $\underline{e}^{(j)}, \underline{\mathbf{d}}^{(j)} \in \mathbb{P}$ for every $j = 1, \cdots, n$. Then one has $\sum_{j=1}^n \mathbf{d}_j = \sum_{r \in \bigsqcup_{j=1}^n A_{i_j}} p^{-r}$, and consequently $\sum_{j=1}^n \underline{\mathbf{d}}^{(j)} \in \mathbb{P}$. In particular,

$$n \cdot \mathrm{dom}_p(M(\underline{A})) = \mathfrak{N}_p\left(\sum_{r \in \bigsqcup_{j=1}^n A_{i_j}} p^{-r} \right) = \mathfrak{N}_p\left(\sum_{j=1}^n \mathbf{d}_j \right) = \Psi\left(\sum_{j=1}^n \underline{\mathbf{d}}^{(j)} \right).$$

Since

$$\begin{aligned} \Psi\left(\sum_{j=1}^n \underline{\mathbf{d}}^{(j)} \right) = \Psi\left(\tau\left(\sum_{j=1}^n \underline{\mathbf{d}}^{(j)} \right) \right) &= \Psi\left(\tau\left(\sum_{j=1}^n \underline{e}^{(j)} \right) \right) \\ &\leq \Psi\left(\sum_{j=1}^n \underline{e}^{(j)} \right) = \sum_{j=1}^n \Psi\left(\underline{e}^{(j)} \right) \leq n \cdot \mathrm{dom}_p(M(\underline{A})), \end{aligned}$$

one concludes that $\Psi\left(\underline{e}^{(j)}\right) = \mathrm{dom}_p(M(\underline{A}))$ for every $j = 1, \cdots, n$ and $\sum_{j=1}^n \underline{e}^{(j)} \in \mathbb{P}$. Consequently, one has $\sum_{j=1}^n \underline{e}^{(j)} = \sum_{j=1}^n \underline{\mathbf{d}}^{(j)}$, implying that $\sum_{j=1}^n e_j = \sum_{j=1}^n \mathbf{d}_j$.

Note that there is no duplication among $e_1, \cdots, e_n$: if not, then there exists a coordinate of $\sum_{j=1}^n \underline{e}^{(j)}$ that is at least 2, contradicting the fact that every coordinate of $\sum_{j=1}^n \underline{\mathbf{d}}^{(j)}$ is 0 or 1. If one writes $e_j = \sum_{r \in A_{k_j}} p^{-r}$ for every $j = 1, \cdots, n$, then

$$\sum_{r \in \bigsqcup_{j=1}^n A_{k_j}} p^{-r} = \sum_{j=1}^n e_j = \sum_{j=1}^n \mathbf{d}_j = \sum_{r \in \bigsqcup_{j=1}^n A_{i_j}} p^{-r},$$

implying that $\bigsqcup_{j=1}^n A_{k_j} = \bigsqcup_{j=1}^n A_{i_j}$. This forces $A_{k_j} = A_{i_j}$ and consequently $e_j = \mathbf{d}_j$ for every $j = 1, \cdots, n$, up to a permutation of $\{1, \cdots, n\}$. ☐

We end this section with the following technical lemma, which will be used in the proof of Theorem 1.7 to extract the orphaned exponents in the multinomial expansion.

Lemma 3.8. — *Let $S \subset \mathbb{P}$ be a (c, n)-sparse subset for some integers $c, n \geq 1$. Let $\underline{d}^{(1)}, \cdots, \underline{d}^{(n)}$ be elements of S that satisfy the condition (2) of Definition 3.5 and set*

$$\phi_0 \colon S \longrightarrow \mathbf{N},\ \underline{d} \longmapsto \operatorname{card}\Big(\Big\{i | 1 \leq i \leq n,\ \underline{d} = \underline{d}^{(i)}\Big\}\Big).$$

Then ϕ_0 is the unique function $\phi \colon S \to \mathbf{N}$ such that $\sum_{\underline{d} \in S} \phi(\underline{d}) \leq n$ and

$$\sum_{\underline{d} \in S} \|\underline{d}\| \cdot \phi(\underline{d}) \equiv \sum_{\underline{d} \in S} \|\underline{d}\| \cdot \phi_0(\underline{d}) \pmod{\mathbf{Z}}.$$

Proof. — Suppose that $\phi_1 \colon S \to \mathbf{N}$ is another function such that $\sum_{\underline{d} \in S} \phi_1(\underline{d}) \leq n$ and

$$\sum_{\underline{d} \in S} \|\underline{d}\| \cdot \phi_1(\underline{d}) \equiv \sum_{\underline{d} \in S} \|\underline{d}\| \cdot \phi_0(\underline{d}) \pmod{\mathbf{Z}}.$$

1. If $\sum_{\underline{d} \in S} \phi_1(\underline{d}) < n$, then by Lemma 3.3 (4) we have

$$\Psi\left(\tau\left(\sum_{\underline{d} \in S} \underline{d} \cdot \phi_1(\underline{d})\right)\right) \leq \Psi\left(\sum_{\underline{d} \in S} \underline{d} \cdot \phi_1(\underline{d})\right) = \sum_{\underline{d} \in S} \Psi(\underline{d}) \cdot \phi_1(\underline{d}) < c \cdot n.$$

   On the other hand, the sparseness condition (2a) of Definition 3.5 implies that $\sum_{j=1}^n \underline{d}_i^{(j)} < p$ for any i, hence

$$\sum_{\underline{d} \in S} \underline{d} \cdot \phi_0(\underline{d}) = \left(\sum_{j=1}^n \underline{d}_i^{(j)}\right)_{i \in \mathbf{Z}_{\geq 0}} = \tau\left(\sum_{\underline{d} \in S} \underline{d} \cdot \phi_0(\underline{d})\right)$$

   and consequently

$$\Psi\left(\tau\left(\sum_{\underline{d} \in S} \underline{d} \cdot \phi_0(\underline{d})\right)\right) = \sum_{i=0}^{\infty} \left(\sum_{j=1}^n \underline{d}_i^{(j)}\right) = \sum_{j=1}^n \Psi(\underline{d}^{(j)}) = n \cdot c.$$

   Thus, we have $\tau\Big(\sum_{\underline{d} \in S} \phi_1(\underline{d}) \cdot \underline{d}\Big) \neq \tau\Big(\sum_{\underline{d} \in S} \underline{d} \cdot \phi_0(\underline{d})\Big)$, which leads to a contradiction by Lemma 3.3 (2).

2. If $\sum_{\underline{d} \in S} \phi_1(\underline{d}) = n$, then we take $\underline{e}^{(1)}, \underline{e}^{(2)}, \cdots, \underline{e}^{(n)} \in S$ such that $\phi_1(\underline{d}) = \operatorname{card}\big(\big\{i | 1 \leq i \leq n,\ \underline{d} = \underline{e}^{(i)}\big\}\big)$ for any $\underline{d} \in S$. Then

$$\sum_{\underline{d} \in S} \|\underline{d}\| \cdot \phi_1(\underline{d}) - \sum_{\underline{d} \in S} \|\underline{d}\| \cdot \phi_0(\underline{d}) = \left\|\sum_{\underline{d} \in S} \underline{d} \cdot \phi_1(\underline{d})\right\| - \left\|\sum_{\underline{d} \in S} \underline{d} \cdot \phi_0(\underline{d})\right\| = \left\|\sum_{i=1}^n \underline{e}^{(i)}\right\| - \left\|\sum_{i=1}^n \underline{d}^{(i)}\right\| \in \mathbf{Z}.$$

   By the sparseness condition (2b) of Definition 3.5, up to a permutation of $\{1, 2, \cdots, n\}$, $\underline{d}^{(i)} = \underline{e}^{(i)}$ for any $i = 1, 2, \cdots, n$. This implies that $\phi_1 = \phi_0$.

☐

4. T-scaled realization of $\mathbf{L}_p$

To prove Theorem 1.7, one has to group the terms of a p-adic Hahn series $f = \sum_{q\in\mathbf{Q}}[f(q)]p^q$ in $\mathbf{L}_p$ by the residue of the exponent modulo $\frac{1}{T}\mathbf{Z}$ for some integer $T \geq 1$. This works when $T = 1$, since a direct computation shows that the element

$$\sum_{q\in\operatorname{Supp}(f)/\mathbf{Z}}\left(\sum_{w\in\mathbf{Z}} f(q+w)p^w\right)t^q \in \breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$$

is a preimage of f under the natural projection $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!) \to \mathbf{L}_p$, where

$$\operatorname{Supp}(f)/\mathbf{Z} \coloneqq \{\inf(\operatorname{Supp}(f)\cap(q+\mathbf{Z})) | q \in \operatorname{Supp}(f)\} \subset \mathbf{Q}$$

is a well-ordered set of representatives of $\operatorname{Supp}(f)$ modulo $\mathbf{Z}$. However, when $T > 1$, the same construction fails, because the element $\sum_{w\in\frac{1}{T}\mathbf{Z}} f(q+w)p^w$ does not necessarily lie in $\breve{\mathbf{Z}}_p$, so a direct analogue of the above construction is not well-defined(3). To resolve this, we enlarge the ring $\breve{\mathbf{Z}}_p = W(\overline{\mathbf{F}}_p)$ to include $p^{1/T}$, and realize $\mathbf{L}_p$ as a quotient of $\breve{\mathbf{Z}}_p[p^{1/T}](\!(t^{\mathbf{Q}})\!)$ by a suitable ideal. This is the main content of this section.

From now on, T will be a positive integer. We set $\breve{\mathbf{Z}}_{p,T} \coloneqq \breve{\mathbf{Z}}_p[p^{1/T}]$ and $\breve{\mathbf{Q}}_{p,T} \coloneqq \breve{\mathbf{Q}}_p(p^{1/T})$.

Lemma 4.1. — *One has* $\left[\breve{\mathbf{Q}}_{p,T}\colon \breve{\mathbf{Q}}_p\right] = T$. *In particular,* $1, p^{1/T}, \cdots, p^{(T-1)/T}$ *form a basis of* $\breve{\mathbf{Q}}_{p,T}$ *over* $\breve{\mathbf{Q}}_p$.

Proof. — Since $\breve{\mathbf{Q}}_p$ is a discrete valuation field and $p^{1/T}$ is a root of the Eisenstein polynomial $X^T - p$, the result follows from the Eisenstein criterion. ☐

Remark 4.2. — $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$ *is a subfield of* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$ *via the natural inclusion* $\breve{\mathbf{Z}}_p \subseteq \breve{\mathbf{Z}}_{p,T}$.

Definition 4.3. — *Let* $\mathcal{N}_T$ *be the set of elements* $\sum_{q\in\mathbf{Q}} c_q t^q \in \breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$ *such that for any* $q \in \mathbf{Q}$,

$$\sum_{n\in\mathbf{Z}} c_{q+\frac{n}{T}}p^{\frac{n}{T}} = 0.$$

We call these elements the T***-null-series*** *in* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$.

Remark 4.4. — *When* $T = 1$, $\mathcal{N}_T$ *coincides with the ideal* $\mathcal{N}$ *of null-series in* $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$ *defined in Proposition 2.3.*

Lemma 4.5. —

1. $\mathcal{N}_T$ *is an ideal of* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$.
2. *For every element* f *of* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$, *there exists a unique element* $g = \sum_{q\in\mathbf{Q}}[g(q)]t^q$ *in* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$ *such that* $f - g \in \mathcal{N}_T$.
3. $\mathcal{N}_T$ *is a maximal ideal of* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$, *so that* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)/\mathcal{N}_T$ *is a field.*

Proof. — These three statements are generalizations of [6, Proposition 3], [6, Proposition 4] and [6, Corollary 3] respectively, and the proofs are essentially the same. ☐

Remark 4.6. — *By (2) of this lemma, we will formally write elements of* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)/\mathcal{N}_T$ *as* $\sum_{q\in\mathbf{Q}}[g(q)]p^q$.

Lemma 4.7. — *One has* $\mathcal{N}_T \cap \breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!) = \mathcal{N}$.

---

(3) This issue was detected during the formalization of the proof of the main theorem in Lean 4.

Proof. — Let $f=\sum_{q\in\mathbf{Q}}[f(q)]t^q\in\mathcal{N}$, then for any $q\in\mathbf{Q}$,

$$\sum_{n\in\mathbf{Z}}[f(q+n)]p^n=0. \tag{1}$$

Notice that for any $q\in\mathbf{Q}$, one has

$$\begin{aligned}\sum_{n\in\mathbf{Z}}\Big[f\Big(q+\frac{n}{T}\Big)\Big]p^{\frac{n}{T}}&=\sum_{u=0}^{T-1}\sum_{\substack{n\in\mathbf{Z}\\ n\equiv u\pmod T}}\Big[f\Big(q+\frac{n}{T}\Big)\Big]p^{\frac{n}{T}}\\&=\sum_{u=0}^{T-1}p^{\frac{u}{T}}\sum_{n\in\mathbf{Z}}\Big[f\Big(\Big(q+\frac{u}{T}\Big)+n\Big)\Big]p^n.\end{aligned}$$

By applying (1) to $q+\frac{u}{T}$ for $u=0,1,\cdots,T-1$, one has $\sum_{n\in\mathbf{Z}}\big[f\big(q+\frac{n}{T}\big)\big]p^{\frac{n}{T}}=0$. As a result, $f\in\mathcal{N}_T$.

Conversely, let $g=\sum_{q\in\mathbf{Q}}[g(q)]t^q\in\mathcal{N}_T$. Then for any $q\in\mathbf{Q}$, one has

$$\sum_{n\in\mathbf{Z}}\Big[g\Big(q+\frac{n}{T}\Big)\Big]p^{\frac{n}{T}}=\sum_{u=0}^{T-1}p^{\frac{u}{T}}\sum_{n\in\mathbf{Z}}\Big[g\Big(\Big(q+\frac{u}{T}\Big)+n\Big)\Big]p^n=0.$$

By Lemma 4.1, for any $u=0,1,\cdots,T-1$, one has $\sum_{n\in\mathbf{Z}}\big[g\big(\big(q+\frac{u}{T}\big)+n\big)\big]p^n=0$. The result follows from taking $u=0$. ☐

Lemma 4.8. — *One has* $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)+\mathcal{N}_T=\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$.

Proof. — Take $f=\sum_{q\in\mathbf{Q}}c_qt^q\in\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$, where

$$c_q=\sum_{n=0}^{\infty}[c_{q,n}]p^{\frac{n}{T}}\in\breve{\mathbf{Z}}_{p,T}.$$

Then one has $f=f_0+p^{\frac{1}{T}}f_1+\cdots+p^{\frac{T-1}{T}}f_{T-1}$, where

$$f_u=\sum_{t\in\mathbf{Q}}\left(\sum_{\substack{n\in\mathbf{N}\\ n\equiv u\pmod T}}[c_{q,n}]p^{\frac{n-u}{T}}\right)t^q\in\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!),\ u=0,1,\cdots,T-1.$$

It suffices to show that $p^{\frac{u}{T}}f_u$ belongs to $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)+\mathcal{N}_T$ for $u=0,1,\cdots,T-1$. This is clear if we write f_u as $f_u=(p^{\frac{u}{T}}\cdot t^0-t^{\frac{u}{T}})f_u+t^{\frac{u}{T}}f_u$, where $t^{\frac{u}{T}}f_u\in\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$ and $(p^{\frac{u}{T}}\cdot t^0-t^{\frac{u}{T}})f_u$ is a T-null-series. ☐

Proposition 4.9. — *There is an isomorphism of valued fields*

$$\sigma\colon\mathbf{L}_p\xrightarrow{\cong}\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)/\mathcal{N}_T.$$

Moreover, the following diagram commutes:

$$\begin{tikzcd}\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!) \arrow[r, hook, "\iota"] \arrow[d, two heads, "\pi"] & \breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!) \arrow[d, two heads, "\pi_T"] \\ \mathbf{L}_p \arrow[r, "\cong", "\sigma"'] & \breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)/\mathcal{N}_T\end{tikzcd}\ . \tag{2}$$

In particular, σ *sends an element* $\sum_{q\in\mathbf{Q}}[f(q)]p^q$ *in* $\mathbf{L}_p$ *to the element* $\sum_{q\in\mathbf{Q}}[f(q)]p^q$ *in* $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)/\mathcal{N}_T$.

Proof. — This is a direct consequence of the following standard fact in commutative algebra: given domains $R_1 \subset R_2$ and an ideal I of R_2, if $R_1 + I = R_2$, then $R_1/(I \cap R_1) \cong R_2/I$, where the isomorphism is induced by the inclusion $R_1 \subset R_2$. ☐

As an application, the T-scaled realization Proposition 4.9 gives the desired grouping of terms of a p-adic Hahn series by residue modulo $\frac{1}{T}\mathbf{Z}$ in the case $T > 1$, and provides an explicit lift of $\sigma(f)$ that will serve as the starting point for the proof of the main theorem. We record this lift below.

Let $f = \sum_{q \in \mathbf{Q}} [f(q)] p^q \in \mathbf{L}_p$ and let $W \subset [0,1) \cap \mathbf{Q}$ be a set of representatives of $-T \cdot \mathrm{Supp}(f)$ modulo $\mathbf{Z}$. Intuitively, the field $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)/\mathcal{N}_T$ provides a suitable setting to write f as $\sum_{q \in W} C_{-\frac{q}{T}} p^{-\frac{q}{T}}$, where $C_{-\frac{q}{T}} \in \breve{\mathbf{Z}}_{p,T}$ for every $q \in W$. To be more precise, for every $q \in W$, we set

$$n_q := -\frac{q}{T} - \inf\left(\mathrm{Supp}(f) \cap \left(-\frac{q}{T} + \frac{1}{T}\mathbf{Z}\right)\right) \in \frac{1}{T}\mathbf{Z}.$$

This element is well-defined. Indeed, the set $\mathrm{Supp}(f) \cap \left(-\frac{q}{T} + \frac{1}{T}\mathbf{Z}\right)$ is non-empty by the assumption that $-\frac{1}{T}W$ is a set of representatives of $\mathrm{Supp}(f)$ modulo $\frac{1}{T}\mathbf{Z}$. Moreover, it is a subset of the well-ordered set $\mathrm{Supp}(f)$, and hence has a minimum element. Then the set

$$\widetilde{W} := \left\{-\frac{q}{T} - n_q \,\middle|\, q \in W\right\} \subseteq \mathrm{Supp}(f)$$

is a well-ordered set of representatives of $\mathrm{Supp}(f)$ modulo $\frac{1}{T}\mathbf{Z}$, and is in bijection with W via the map

$$\mu_W \colon W \longrightarrow \widetilde{W},\ q \longmapsto -\frac{q}{T} - n_q = \inf\left(\mathrm{Supp}(f) \cap \left(-\frac{q}{T} + \frac{1}{T}\mathbf{Z}\right)\right). \tag{3}$$

For any $w \in \widetilde{W}$, set $C_w := \sum_{r \in \mathbf{Z}} [f(w + \frac{r}{T})] p^{\frac{r}{T}} \in \breve{\mathbf{Q}}_{p,T}$. By the construction of $\widetilde{W}$, we have $C_w \neq 0$ and $f(w + \frac{r}{T}) = 0$ for every $w \in \widetilde{W}$ and every integer $r < 0$. Thus $C_w \in \breve{\mathbf{Z}}_{p,T}$, and the element

$$\widehat{f} := \sum_{w \in \widetilde{W}} C_w \cdot t^w \in \breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!) \tag{4}$$

is a lift of $\sigma(f)$ in $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$.

5. Main theorem and its applications

5.1. Main theorem. — In this subsection, we prove the main theorem of this paper (cf. Theorem 5.3): a p-adic Hahn series whose sign-inverted support, after scaling by some integer $T \geq 1$, admits a sparse set of representatives modulo $\mathbf{Z}$ is transcendental over $\breve{\mathbf{Q}}_p$. The argument proceeds by contradiction: if such a Hahn series f is a root of a polynomial P over $\breve{\mathbf{Q}}_p$, then expanding the lifted element $P(\widehat{f})$ from (4) via the multinomial formula would yield a family of vanishing identities on its coefficients (Lemma 5.1). The sparseness condition then forces these identities, when specialized at a carefully chosen exponent, to collapse to a single nonzero term (Lemma 5.2), and will lead to a contradiction.

We isolate the two key inputs as separate lemmas before assembling the proof.

Lemma 5.1. — *Let $f \in \mathbf{L}_p$ be a p-adic Hahn series and let $T \geq 1$ be an integer. Let $W \subset [0,1) \cap \mathbf{Q}$ be a set of representatives of $-T \cdot \mathrm{Supp}(f)$ modulo $\mathbf{Z}$, and let $\widetilde{W} \subseteq \mathrm{Supp}(f)$*

and $\widehat{f} = \sum_{w\in\widetilde{W}} C_w \cdot t^w \in \breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$ *be as in* (4)*. Suppose that* f *is a root of a polynomial* $P(X) = \sum_{i=0}^{n} a_i X^i \in \breve{\mathbf{Z}}_p[X]$*. Then for every* $q \in \mathbf{Q}$*, one has*

$$\sum_{w\in\mathbf{Z}} p^{\frac{w}{T}} \sideset{}{'}\sum_{\substack{\widetilde{\phi}\colon \widetilde{W}\to\mathbf{N}\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\le n\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\cdot s = q+\frac{w}{T}}} a_{\sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)} \frac{\left(\sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\right)!}{\prod''_{s\in\widetilde{W}}\widetilde{\phi}(s)!} \sideset{}{''}\prod_{s\in\widetilde{W}} C_s^{\widetilde{\phi}(s)} = 0, \tag{5}$$

where $\sum'$ *emphasizes that the summation is taken over a finite index set, and* $\sum''$ *(resp.* $\prod''$*) emphasizes that the function inside the summation (resp. product) is finitely supported.*

Proof. — By multinomial expansion, for any integer $i \ge 0$, one has

$$\widehat{f}^i = \sum_{q\in\mathbf{Q}} \left(\sideset{}{'}\sum_{\substack{s_1,\cdots,s_i\in\widetilde{W}\\ s_1+\cdots+s_i=q}} \prod_{k=1}^{i} C_{s_k} \right) t^q = \sum_{q\in\mathbf{Q}} \left(\sideset{}{'}\sum_{\substack{\widetilde{\phi}\colon \widetilde{W}\to\mathbf{N}\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)=i\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\cdot s=q}} \frac{i!}{\prod''_{s\in\widetilde{W}}\widetilde{\phi}(s)!} \sideset{}{''}\prod_{s\in\widetilde{W}} C_s^{\widetilde{\phi}(s)} \right) t^q.$$

Since $\sigma\left(\sum_{i=0}^{n} a_i f^i\right) = \sum_{i=0}^{n} a_i \sigma(f)^i = 0$ and $\widehat{f}$ is a lift of $\sigma(f)$, the element

$$\begin{aligned}\sum_{i=0}^{n} a_i \widehat{f}^i &= \sum_{q\in\mathbf{Q}} \left(\sum_{i=0}^{n} a_i \sideset{}{'}\sum_{\substack{\widetilde{\phi}\colon \widetilde{W}\to\mathbf{N}\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)=i\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\cdot s=q}} \frac{i!}{\prod''_{s\in\widetilde{W}}\widetilde{\phi}(s)!} \sideset{}{''}\prod_{s\in\widetilde{W}} C_s^{\widetilde{\phi}(s)} \right) t^q \\ &= \sum_{q\in\mathbf{Q}} \left(\sideset{}{'}\sum_{\substack{\widetilde{\phi}\colon \widetilde{W}\to\mathbf{N}\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\le n\\ \sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\cdot s=q}} a_{\sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)} \frac{\left(\sum''_{s\in\widetilde{W}}\widetilde{\phi}(s)\right)!}{\prod''_{s\in\widetilde{W}}\widetilde{\phi}(s)!} \sideset{}{''}\prod_{s\in\widetilde{W}} C_s^{\widetilde{\phi}(s)} \right) t^q\end{aligned}$$

is a T-null-series in $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$. The equation (5) follows from the definition of T-null-series. □

Lemma 5.2. — *Let* $f \in \mathbf{L}_p$ *be a* p*-adic Hahn series. Let* $S \subset \mathbb{P}$ *be a* (c,n)*-sparse subset for some integers* $c, n \ge 1$*, and let* $\underline{d}^{(1)}, \cdots, \underline{d}^{(n)} \in S$ *satisfy condition (2) of Definition 3.5. Set* $\phi_0\colon S \longrightarrow \mathbf{N}$, $\underline{d} \longmapsto \operatorname{card}\left(\left\{ i \,\middle|\, \underline{d} = \underline{d}^{(i)} \right\}\right)$*. Suppose that* $\|S\|$ *is a set of representatives of* $-T\cdot\operatorname{Supp}(f)$ *modulo* $\mathbf{Z}$*, and let* $\widetilde{S} := \widetilde{\|S\|}$ *and* $\mu := \mu_{\|S\|}\circ\|\cdot\|\colon S \longrightarrow \widetilde{S}$ *be as in* (3)*. Set*

$$q_0 := -\frac{1}{T}\sum_{\underline{d}\in S} \|\underline{d}\| \cdot \phi_0(\underline{d}).$$

Then $\widetilde{\phi} = \phi_0 \circ \mu^{-1}$ *is the unique function* $\widetilde{\phi} \colon \widetilde{S} \to \mathbf{N}$ *satisfying*

$$\sideset{}{''}\sum_{s\in\widetilde{S}} \widetilde{\phi}(s) \leq n \quad \text{and} \quad \sideset{}{''}\sum_{s\in\widetilde{S}} \widetilde{\phi}(s)\cdot s \equiv q_0 \pmod{\tfrac{1}{T}\mathbf{Z}}.$$

Proof. — Take $\widetilde{\phi} \colon \widetilde{S} \to \mathbf{N}$ that satisfies $\sum''_{s\in\widetilde{S}} \widetilde{\phi}(s) \leq n$ and $\sum''_{s\in\widetilde{S}} \widetilde{\phi}(s)\cdot s \equiv q_0 \pmod{\frac{1}{T}\mathbf{Z}}$, and set $\phi \coloneqq \widetilde{\phi}\circ\mu$. Then

$$\sideset{}{''}\sum_{\underline{d}\in S} \phi(\underline{d}) = \sideset{}{''}\sum_{s\in\widetilde{S}} \widetilde{\phi}(s) \leq n$$

and

$$\begin{aligned}\sideset{}{''}\sum_{\underline{d}\in S} \phi(\underline{d})\cdot\|\underline{d}\| &= -T\sideset{}{''}\sum_{s\in\widetilde{S}} \widetilde{\phi}(s)\cdot s - \sideset{}{''}\sum_{\underline{d}\in S} \widetilde{\phi}(\mu(\underline{d}))\big(T\cdot n_{\|\underline{d}\|}\big)\\ &\equiv -T\sideset{}{''}\sum_{s\in\widetilde{S}} \widetilde{\phi}(s)\cdot s\\ &\equiv -T\cdot q_0 = \sum_{\underline{d}\in S}\|\underline{d}\|\cdot\phi_0(\underline{d}) \pmod{\mathbf{Z}}.\end{aligned}$$

By Lemma 3.8, one knows that $\phi = \phi_0$, i.e. $\widetilde{\phi} = \phi_0\circ\mu^{-1}$. □

The main theorem now follows from the above two lemmas and the construction of $\widehat{f}$ in (4).

Theorem 5.3. — *Let* $f \in \mathbf{L}_p$ *be a p-adic Hahn series such that there exists an integer* $T \geq 1$ *for which* $-T\cdot\operatorname{Supp}(f)$ *admits a sparse set* $W \neq \{0\}$ *of representatives modulo* $\mathbf{Z}$*. Then* f *is transcendental over* $\breve{\mathbf{Q}}_p$*, and hence over* $\mathbf{Q}_p$*.*

Proof. — Suppose for contradiction that f is a root of a polynomial $P(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \breve{\mathbf{Z}}_p[X]$ with $a_n \neq 0$. By multiplying $P(X)$ by a suitable power of X, we may assume that there exists $S \subset \mathbb{P}$ such that $\|S\| \coloneqq \{\|\underline{d}\| \in \mathbf{Q} | \underline{d} \in S\}$ is a set of representatives of $-T\cdot\operatorname{Supp}(f)$ modulo $\mathbf{Z}$, and S is (c,n)-sparse for some integer $c \geq 1$ (cf. Lemma 3.6), where n equals the degree of $P(X)$. Let $\underline{d}^{(1)}, \cdots, \underline{d}^{(n)} \in S$ be elements satisfying condition (2) of Definition 3.5 for n and c, set

$$\phi_0 \colon S \longrightarrow \mathbf{N},\ \underline{d} \longmapsto \operatorname{card}\Big(\Big\{i \Big| 1 \leq i \leq n,\ \underline{d} = \underline{d}^{(i)}\Big\}\Big),$$

and let $\widetilde{S} \coloneqq \widetilde{\|S\|}$ and $\mu \colon S \to \widetilde{S}$ be as in (3) (with $W = \|S\|$).

By Lemma 5.1 applied to $W = \|S\|$, for every $q \in \mathbf{Q}$ one has the identity (5). Specializing to

$$q = q_0 \coloneqq -\frac{1}{T}\sum_{\underline{d}\in S}\|\underline{d}\|\cdot\phi_0(\underline{d}),$$

Lemma 5.2 tells us that the only $\widetilde{\phi} \colon \widetilde{S} \to \mathbf{N}$ contributing to the interior summation of (5) at $q = q_0$ is $\widetilde{\phi} = \phi_0\circ\mu^{-1}$. Hence (5) at $q = q_0$ reduces to the single term

$$\frac{p^{\frac{w}{T}}\cdot n!}{\prod''_{\underline{d}\in S}\phi_0(\underline{d})!}\cdot a_n\cdot\sideset{}{''}\prod_{\underline{d}\in S} C_{\mu(\underline{d})}^{\phi_0(\underline{d})} = 0,$$

where $w = T\Big(\sum''_{s\in\widetilde{S}}\phi_0\big(\mu^{-1}(s)\big)\cdot s - q_0\Big) \in \mathbf{Z}$. This forces $a_n = 0$ or $C_{\mu(\underline{d})} = 0$ for some $\underline{d} \in S$, contradicting $a_n \neq 0$ and $C_s \neq 0$ for every $s \in \widetilde{S}$. □

5.2. Applications. — A typical application of Theorem 1.7 is to show that if the decimal parts of any two elements of the sign-inverted support $-\operatorname{Supp}(f)$ of a p-adic Hahn series f have non-overlapping base-p digits, then f is transcendental over $\breve{\mathbf{Q}}_p$:

Proposition 5.4. — *Let* $\underline{A} := (A_i)_{i\geq 1}$ *be a family of pair-wise disjoint nonempty subsets of* $\mathbf{N}$ *such that* $\sup_i \operatorname{card}(A_i) < \infty$. *If the set of the form*

$$M_c(\underline{A}) := \left\{ \frac{1}{T}\left(c_i - \sum_{r\in A_i} p^{-r} \right) \middle| c_i \in \mathbf{Z},\ i = 1, 2, \cdots \right\},\ T \in \mathbf{Z}_{\geq 1}$$

is well-ordered, then any p-adic Hahn series with support $M_c(\underline{A})$ is transcendental over $\breve{\mathbf{Q}}_p$.

Proof. — Let $f \in \mathbf{L}_p$ be a p-adic Hahn series with support $M_c(\underline{A})$. Let

$$\lambda\colon M(\underline{A}) \longrightarrow M_c(\underline{A}),\ \sum_{r\in A_i} p^{-r} \longmapsto \frac{1}{T}\left(c_i - \sum_{r\in A_i} p^{-r} \right),$$

which is obviously a bijection.

Since there is at most one A_i that contains 0, and removing finitely many terms from a p-adic Hahn series does not change its algebraicity, we may assume that $0 \notin A_i$ for every $i \geq 1$.

Let $C \in \mathbf{Z}_{\geq 1}$ be the maximal integer such that $\operatorname{card}(A_i) = C$ is attained for infinitely many i, and let $\underline{A}' := \{A_i \mid i = 1, 2, \cdots,\ \operatorname{card}(A_i) \leq C\}$. Since $C \leq \sup_i \operatorname{card}(A_i) < \infty$, the set $\underline{A}\backslash\underline{A}'$ is finite. If we set

$$M(\underline{A}') := \left\{ \sum_{r\in A_i} p^{-r} \middle| A_i \in \underline{A}' \right\},$$

then $M(\underline{A}')$ is a sparse set by Example 3.7, and $M(\underline{A})\backslash M(\underline{A}')$ is a finite set. Note that $\lambda(M(\underline{A}'))$ is a well-ordered subset of $\mathbf{Q}$, we can set

$$f' := \sum_{q\in\lambda(M(\underline{A}'))} [f(q)]p^q,$$

which is a well-defined element of $\mathbf{L}_p$. By Theorem 1.7, f' is transcendental over $\breve{\mathbf{Q}}_p$. Since the support of $f - f'$ is finite, f is also transcendental over $\breve{\mathbf{Q}}_p$. □

As a simple corollary of Proposition 5.4, we prove the following mixed-characteristic analogue of the result of Huang and Ştefănescu (cf. Theorem 1.9):

Corollary 5.5. — *Let* $f = \sum_{i=1}^{\infty} [f(i)] \cdot p^{-1/p^i} \in \mathbf{L}_p$ *be a p-adic Hahn series. Then the following are equivalent:*

1. *the series f is algebraic over $\mathbf{Q}_p$;*
2. *the series f is algebraic over $\breve{\mathbf{Q}}_p$;*
3. *$f(i) = 0$ for all but finitely many i.*

Proof. — One only needs to verify (2) implies (3).

If $\operatorname{Supp}(f)$ is infinite, then we may write $\operatorname{Supp}(f) = \{q_1, \cdots, q_n, \cdots\}$, where $q_i = -1/p^{k_i}$ for some integer $k_i \in \mathbf{Z}_{\geq 1}$. The result follows from taking $A_i = \{k_i\}$, $c_i = 0$ for every $i \geq 1$ and $T = 1$ in Proposition 5.4. □

Appendix A. AI-assisted formalization in Lean 4

A.1. Overview of the formalization project. — The formalization of this paper, which we refer to as the `FormalizedSparse` project, contains approximately 19,300 lines of Lean code (including docstrings) and is organized into the following files:

A.1.1. `WittVector.lean`. — This file contains our realization of $\breve{\mathbf{Q}}_p$, the completed maximal unramified extension of $\mathbf{Q}_p$, and its ring of integers $\breve{\mathbf{Z}}_p$.

Although Mathlib already contains a formalization of $\mathbf{Q}_p$, the lack of infrastructure on ramification theory in Mathlib makes it difficult to define $\breve{\mathbf{Q}}_p$ literally as the union of all finite unramified extensions of $\mathbf{Q}_p$. Instead, we start with $\breve{\mathbf{Z}}_p$, which we define as the ring of Witt vectors over $\overline{\mathbf{F}}_p$:

```
-- The algebraic closure of `𝔽ₚ`
abbrev Fpbar (p : ℕ) [Fact (Nat.Prime p)] := AlgebraicClosure (ZMod p)
notation "𝔽ᵃ_[" p "]" => Fpbar p

-- The ring of integers of completed maximal unramified extension of `ℚₚ`,
abbrev OQpUn (p : ℕ) [Fact (Nat.Prime p)] := WittVector p (Fpbar p)
notation "ℤᵘⁿ_[" p "]" => OQpUn p
```

Then we define $\breve{\mathbf{Q}}_p$ as the fraction field of $\breve{\mathbf{Z}}_p$, with the induced valuation:

```
abbrev QpUn (p : ℕ) [Fact (Nat.Prime p)] :=
  WithVal ((IsDiscreteValuationRing.maximalIdeal (ℤᵘⁿ_[p])).valuation
    ((FractionRing (ℤᵘⁿ_[p]))))
notation "ℚᵘⁿ_[" p "]" => QpUn p
```

We establish several `instance` around `ℚᵘⁿ_[p]`, such as the fact that it is a complete rank-1 valued field. Besides that, we define the embedding from $\mathbf{Q}_p$ to $\breve{\mathbf{Q}}_p$: one has $\mathbf{Z}_p \cong W(\mathbf{F}_p)$, which injects into $\breve{\mathbf{Z}}_p := W(\overline{\mathbf{F}}_p)$. This extends to an embedding $\mathbf{Q}_p \hookrightarrow \breve{\mathbf{Q}}_p$:

```
-- The embedding from ℚ_[p] to ℚᵘⁿ_[p].
noncomputable def Qp_embd {p : ℕ} [Fact (Nat.Prime p)] : ℚ_[p] →+* ℚᵘⁿ_[p] :=
  @IsFractionRing.map ℤ_[p] ℤᵘⁿ_[p] ℚ_[p] ℚᵘⁿ_[p] _ _ _ _ _ _ _ _ _
    ((WittVector.map (algebraMap (ZMod p) (𝔽ᵃ_[p]))).comp
      (WittVector.fromPadicInt p)) ... -- The map $\mathbf{Z}_p \longrightarrow \breve{\mathbf{Z}}_p$ is injective. Omitted.
```

Finally, we formalize a lemma to show that this embedding preserves the valuation.

A.1.2. `PAdicHahnSeries.lean`. — This file contains the formalization of $\mathbf{L}_p$, the field of p-adic Hahn series. The material is mostly taken from [6].

Since the equal-characteristic Hahn series is already available in Mathlib, we formalize the ring $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$ as:

```
abbrev LiftedPAdicHahnSeries (p : ℕ) [Fact (Nat.Prime p)] := HahnSeries ℚ (ℤᵘⁿ_[p])
```

We define the null series condition (cf. Proposition 2.3) as a predicate `IsNullSeries` on `LiftedPAdicHahnSeries`, and define the ideal $\mathcal{N}$ as the set of null series:

```
def NullSeriesIdeal (p : ℕ) [Fact (Nat.Prime p)] : Ideal (LiftedPAdicHahnSeries p)
↪  where
  carrier := {x | IsNullSeries x}
  ... -- Omitted.
```

We provide an `instance` to show that $\mathcal{N}$ is a maximal ideal of $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$, so that the quotient $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)/\mathcal{N}$ is a field (cf. (1) of Proposition 2.3).

Instead of directly defining $\mathbf{L}_p$ as the quotient of $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)$ by $\mathcal{N}$, we first formalize (2) of Proposition 2.3:

```
-- Every element of $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)/\mathcal{N}$ has a unique lift of the form $\sum_{q\in\mathbf{Q}}[f(q)]t^q$,
-- with the support a well-ordered set.
theorem exists_canonical_expansion {p : ℕ} [Fact (Nat.Prime p)] :
  ∀ A : (LiftedPAdicHahnSeries p) / (NullSeriesIdeal p),
    ∃! (s : {f : ℚ → 𝔽ᵃ_[p] // (Function.support f).IsPWO}),
      Ideal.Quotient.ringCon (NullSeriesIdeal p)
      A.out (LiftedPAdicHahnSeries.from_coeff s.val s.prop)
```

This allows us to define a valuation on $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)/\mathcal{N}$ by considering the minimum of the support of this unique lift:

```
def val (p : ℕ) [Fact (Nat.Prime p)] :
  AddValuation ((LiftedPAdicHahnSeries p) / (NullSeriesIdeal p)) (WithTop ℚ) where
  toFun x :=
    if h : x = 0 then (⊤ : WithTop ℚ) --If $x = 0$, then its valuation is $\infty$.
    else ((support_IsPWO x).isWF.min (support_nonempty_of_nonzero p x h) : WithTop ℚ)
  ... -- Omitted.
```

And finally we define $\mathbf{L}_p$ as the fraction field of $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)/\mathcal{N}$ with this valuation:

```
abbrev pAdicHahnSeries (p : ℕ) [Fact (Nat.Prime p)] : Type _ := WithVal (val p)
notation "𝕃_[" p "]" => pAdicHahnSeries p
```

Several facilities around $\mathbf{L}_p$, such as its support and the coefficients (as a function of type `ℚ → 𝔽ᵃ_[p]`), are also formalized in this file.

To deliver the $\mathbf{Q}_p$-transcendence and $\breve{\mathbf{Q}}_p$-transcendence results, we provide the embedding from $\breve{\mathbf{Q}}_p$ to $\mathbf{L}_p$. It is induced by the embedding from $\breve{\mathbf{Z}}_p$ to $\mathbf{L}_p$, which maps an element w to the image of $w \cdot t^0$ in the quotient $\breve{\mathbf{Z}}_p(\!(t^{\mathbf{Q}})\!)/\mathcal{N}$.

```
def ZpUn_embd {p : ℕ} [Fact (Nat.Prime p)] : ℤᵘⁿ_[p] →+* 𝕃_[p] where
  toFun a := Ideal.Quotient.mk (NullSeriesIdeal p) (HahnSeries.single 0 a)
  ... -- Omitted.

def QpUn_embd {p : ℕ} [Fact (Nat.Prime p)] : ℚᵘⁿ_[p] →+* 𝕃_[p] :=
  IsFractionRing.map (j := ZpUn_embd (p := p)) ZpUn_embd_injective
```

Finally, we show by induction on the cardinality of the support that an element of $\mathbf{L}_p$ with finite support must be algebraic over $\mathbf{Q}_p$:

```
lemma alg_of_fin_supp (p : ℕ) [Fact (Nat.Prime p)] (f : 𝕃_[p])
  (hf : f.support.Finite) : IsAlgebraic ℚ_[p] f
```

A.1.3. `Tscaled.lean`. — This file corresponds to Section 4 of this paper. We use `ℤᵘⁿ_[p,T]` (resp. `ℚᵘⁿ_[p,T]`) to denote the ring $\breve{\mathbf{Z}}_{p,T}$ (resp. $\breve{\mathbf{Q}}_{p,T}$), and use `𝕃_[p,T]` to denote the quotient of $\breve{\mathbf{Z}}_{p,T}(\!(t^{\mathbf{Q}})\!)$ by $\mathcal{N}_T$. After we formalize Lemma 4.7 and Lemma 4.8, the isomorphism σ in Proposition 4.9 can be delivered:

```
def σ : 𝕃_[p] ≃+* 𝕃_[p,T] := ... -- Omitted.

-- Remark 4.10, $\sigma$ is given by $\sum_{q\in\mathbf{Q}}[f(q)]p^q \mapsto \sum_{q\in\mathbf{Q}}[f(q)]p^q$.
theorem σ_coeff_compat (f : 𝕃_[p]) : (σ p T f).coeff = f.coeff
```

A.1.4. `Sparse.lean`. — This file corresponds to Definition 1.3, Definition 1.4 and Section 3 of this paper, which is about the sparseness, (c, n)-sparseness and the related infrastructure.

Fortunately, Mathlib already contains the function `Real.digits`, which extracts the digits of a real number in a given base. This allows us to formalize the p-digit sum of a rational number, and consequently $\mathrm{dom}_p(S)$ and $\mathrm{Dom}_p(S)$ of a set $S \subseteq \mathbf{Q}$ with a little effort:

```
abbrev decDigits (p : ℕ) [Fact (Nat.Prime p)] (q : ℚ) : ℕ+ → Fin p :=
  fun n => Real.digits (Int.fract q) p ((n : ℕ) - 1)

/- The `p-digit sum`, `𝔑ₚ(q)` in `Definition 1.3 (1)`-/
def pDigitSum (p : ℕ) [Fact (Nat.Prime p)] (q : ℚ) : WithTop ℕ :=
  if h : (Function.support (decDigits p q)).Infinite then ⊤
  else ∑ n ∈ (Set.not_infinite.1 h).toFinset, (decDigits p q n).val

/- `dominant p-digit sum` of S, `Definition 1.3 (2)`-/
def dom (p : ℕ) [Fact (Nat.Prime p)] (S : Set ℚ) : WithTop ℕ :=
  sSup {pDigitSum p  q | q ∈ S}

/- `p-digit dominant part` of S, `Definition 1.3 (2)`-/
def Dom (p : ℕ) [Fact (Nat.Prime p)] (S : Set ℚ) : Set ℚ :=
  {q ∈ S | pDigitSum p q = dom p S}
```

With these preparations, we present the formalized version of Definition 1.4:

```
/- `Definition 1.4`: The sparse condition -/
def IsSparse (p : ℕ) [Fact (Nat.Prime p)] (S : Set ℚ) : Prop :=
  S ⊆ Set.Ico 0 1 ∧ dom p S < ⊤ ∧
  ∃ D : Set ℕ+, D.Infinite ∧ ( ∀ n ∈ D, ∃ d : Fin n → Dom p S, (
      ∀ i : ℕ+, ∑ (j : Fin n), (decDigits p (d j) i).val < p
      -- No carrying when adding d₁, d₂, ..., dₙ together.
    ) ∧ ( ∀ e : Fin n → Dom p S,
      (∑ i, (d i).val -∑ i, (e i).val).isInt →
        ∃ perm : Equiv.Perm (Fin n), ∀ i, d i = e (perm i)
      -- Unique up to a permutation of $\{1, \cdots n\}$.
    ) )
```

Since Section 3 contains mostly implementation-level details, we will not demonstrate most of the formalization here, except for Example 3.7: the p-digit disjoint subset of $[0, 1) \cap \mathbf{Q}$ is sparse:

```
lemma IsSparse_of_digit_disjoint (p : ℕ) [Fact (Nat.Prime p)] (A : ℕ → Set ℕ+)
(hA1 : ∀ n, (A n).Nonempty) (hA2 : ∀ i j, (A i) ∩ (A j) ≠ ∅ → i = j)
(hA3 : ∀ n, (A n).Finite)
(hAsup : ∃ K : ℕ, (∀ n, (hA3 n).toFinset.card ≤ K) ∧
          {n | (hA3 n).toFinset.card = K}.Infinite) :
  IsSparse p {∑ r ∈ (hA3 i).toFinset, (p : ℚ) ^ (-(r: ℤ)) | i : ℕ }
```

A.1.5. `MainTheorem.lean`. — The single objective of this file is to formalize the proof of Theorem 1.7:

```
theorem main_theorem (p : ℕ) [Fact (Nat.Prime p)] (f : 𝕃_[p]) (T : ℕ+)
(W : Set ℚ) (hW1 : W ≠ {0}) (hW2 : IsSparse p W)
(hf : IsRepModZ W {-1 * T * q | q ∈ f.support}) :
  ¬ IsAlgebraic ℚᵘⁿ_[p] f
```

Here `IsRepModZ` is a predicate to express the condition that W is a set of representatives of $-T \cdot \mathrm{Supp}(f)$ modulo $\mathbf{Z}$.

A variant of this theorem, which replaces the algebraicity over $\breve{\mathbf{Q}}_p$ by the algebraicity over $\mathbf{Q}_p$, is also formalized in this file.

A.1.6. `Application.lean`. — This file contains the formalization of Proposition 5.4 and Theorem 1.10, which are the application of Theorem 1.7 and Example 3.7. Among them, Theorem 1.10, the p-adic analogue of the result of Huang and Ştefănescu, is formalized as:

```
theorem pAdicHuangStefanescu (p : ℕ) [Fact (Nat.Prime p)]
(f : 𝕃_[p]) (hf : f.support ⊆ {-(p : ℚ) ^ (-(i : ℤ)) | i : ℕ+}) :
List.TFAE [
  f.support.Finite,
  IsAlgebraic ℚᵘⁿ_[p] f,
  IsAlgebraic ℚ_[p] f
]
```

The formalized statement looks very similar to the informal one.

A.2. AI-assisted formalization. — The agentic auto-formalization system Archon, which is based on Claude Opus 4.7 and developed by the AI4Math team of BICMR of Peking University, greatly accelerated our formalization process. By design, once the blueprint or the corresponding informal proof of the project is provided, Archon works fully autonomously to complete the project-level informal-to-formal translation of the statements (as well as mathematical definitions) and to formalize the proof.

A.2.1. *Good translation of the statements and definitions.* — Specific to our project, our experience suggests that it is better to write all the formalized definitions and statements by a human mathematician with solid experience in Lean(4), and then let Archon fill in the proofs. While Archon is capable of independently formalizing mathematical concepts, it, like other artificial intelligence systems, lacks an adequate mathematical intuition to formulate definitions in a manner that facilitates their subsequent application.

---

(4) In the `FormalizedSparse` project, we formalized all the definitions and statements by hand.

For example, at the beginning of the project, we let Archon formalize $\breve{\mathbf{Q}}_p$, the completed maximal unramified extension of $\mathbf{Q}_p$. As we observed, Archon tried very hard to literally formalize the definition of $\breve{\mathbf{Q}}_p$, without success. In fact, the mathematical insight here is that the ramification information is not what we really need. Instead, the key property of $\breve{\mathbf{Q}}_p$ that is repeatedly used in this project is that every element of $\breve{\mathbf{Q}}_p$ can be uniquely written as a p-adic Laurent series whose coefficients are Teichmüller representatives. This is the perfect scenario for applying the ring of Witt vectors, which is already formalized in Mathlib.

A.2.2. *Automatic formalization: proof and revision.* — After the formalized statements and definitions are settled(5), we let Archon fill in the proofs unsupervisedly. The result is quite satisfactory: Archon is able to complete most proofs in a way that closely follows the informal proof, and the generated code (as well as the docstrings) is mostly readable. We learned several things from this process:

1. During its work, the formalization by Archon helped us find several mistakes (which are now fixed) in the original version of the informal proof:
    (a) Theorem 1.7 does not hold for the sparse set $W = \{0\}$.
    (b) We did not require the supremum in Example 1.6 to be attained by infinitely many i in the original version, which is also necessary for the proof to work.
    (c) As we mentioned at the beginning of Section 4, some elements in $\mathbf{L}_p$ were casually written in the form $\sum_{q\in[0,1/T)} c_q p^q$, with $c_q \in \breve{\mathbf{Q}}_{p,T}$ for every q. This is not a well-defined element of $\mathbf{L}_p$, unless we consider the T-scaled realization (i.e., `𝕃_[p,T]`) of $\mathbf{L}_p$ and the isomorphism σ in Proposition 4.9.

   These mistakes, which are all fixed in the current version, are all related to the technical details of the proof, and would not have been easy to find without the help of formalization. Archon records these mistakes in the docstrings. For the first two mistakes, Archon added the necessary condition by itself and continued the formalization process without any human intervention. For the third mistake, Archon failed to complete the formalization and terminated with a detailed report.
2. Archon's ability to backtrack is impressive. When formally proving Theorem 1.7, which is highly combinatorial, we gave no hint to Archon about the structure of the proof, except for the informal proof itself. Archon made multi-level plans to divide the proof into several lemmas and assemble them together to complete the proof. During the work of Archon, we observed that some of its intermediate lemmas were not correct, and Archon was able to backtrack and revise the proof plan to fix the mistake without any human intervention.

A.3. How formalization helps mathematical research. — For mathematicians with limited experience in formalization, agent-based systems such as Archon may eventually provide a practical way to validate proofs in a largely black-box manner. One can envision a future workflow in which a paper written in natural language is automatically translated into a formal proof object and subsequently verified by the system after extensive computation.

For mathematicians with some experience in formalization, we believe that mathematical research can benefit substantially from a human-in-the-loop workflow. In such a workflow, researchers formulate definitions and statements of intermediate results in formal language, while

---

(5) We also built up several helper results (with `sorry`) that we expected to be useful during the formalization process.

AI systems assist with the labor-intensive formalization process. Successfully formalized intermediate results then provide verified foundations for subsequent arguments, while failed formalizations may help detect errors at an early stage of the research process.

References

[1] Huang M F. On the Algebraic Closure of the Field of Meromorphic Functions Over An Algebraically Closed Field of Characteristic p. PhD thesis. Purdue University, 1968

[2] Ju H C, Gao G X, Jiang J D, et al. Automated Conjecture Resolution with Formal Verification. 2026. arXiv: `2604.03789 [cs.LG]`

[3] Kedlaya K S. On the Algebraicity of Generalized Power Series. Beitr Algebra Geom, 2017, 58.3: 499–527

[4] Kedlaya K S. The Algebraic Closure of the Power Series Field in Positive Characteristic. Proc. Amer. Math. Soc., 2001, 129.12: 3461–3470

[5] Lampert D. Algebraic p-Adic Expansions. Journal of Number Theory, 1986, 23.3: 279–284

[6] Poonen B. Maximally Complete Fields. L'Enseignement Mathématique, 1993, 39.1-2: 87

[7] Ştefănescu D. On Meromorphic Formal Power Series. Bulletin mathématique de la Société des Sciences Mathématiques de la République Socialiste de Roumanie, 1983, 27 (75).2: 169–178. JSTOR: `43684000`

[8] Wang S W, Yuan Y J. Hyper-algebraic invariants of p-adic algebraic numbers. 2024. arXiv: `2402.15947 [math.NT]`

[9] Wang S W, Yuan Y J. On the p-adic transcendence of $\sum_{k=1}^{\infty} p^{-1/p^k}$. 2025. arXiv: `2509.24609 [math.NT]`

Shanwen Wang, School of Mathematics, Renmin University of China, No. 59 Zhongguancun Street, Haidian District, Beijing, 100872, China • *E-mail :* `s_wang@ruc.edu.cn`
Url : `https://shanwenwang.github.io/`

Yijun Yuan, Institute for Theoretical Sciences, Westlake University, No. 600 Dunyu Road, Sandun town, Xihu district, Hangzhou, Zhejiang Province, 310030, China • *E-mail :* `941201yuan@gmail.com`
Url : `https://yijunyuan.github.io/`