FUNCTIONAL DEGREES AND ARITHMETIC APPLICATIONS III: BEYOND PRIME EXPONENT

PETE L. CLARK AND UWE SCHAUZ

Abstract. This is a continuation of our prior work on group-theoretic generalizations of the prime Ax-Katz Theorem. In this work we give a lower bound on the p-adic divisibility of the solution set $Z(f_1, \dots, f_r)$ for maps $f_1, \dots, f_r$ between two finite commutative p-groups A and B in terms of the invariant factors of A, the exponent of B and the *functional degrees* of the maps in the sense of the Aichinger-Moosbauer calculus. The case of maps between arbitrary finite commutative groups follows easily from this using prior work of Aichinger-Moosbauer and the present authors. We get in particular a version of Ax-Katz for polynomial functions over any finite rng.

Contents

1. Introduction and Main Results

1.1. **Notation and Terminology.** Throughout this paper, p is a fixed but arbitrary (positive) prime number. We denote by ord_p the **p-adic valuation** on $\mathbb{Q}$, and with $\mathbb{Z}_{(p)}$ for the ring of rational numbers of non-negative p-adic valuation. We set

$$\mathbb{Z}^+ := \{n \in \mathbb{Z} \mid n > 0\}\ ,\ \ \mathbb{N} := \{n \in \mathbb{Z} \mid n \geq 0\} \quad \text{and} \quad \tilde{\mathbb{N}} := \mathbb{N} \cup \{-\infty, \infty\}\,,$$

and endow $\tilde{\mathbb{N}}$ with the total ordering extending the usual one on $\mathbb{N}$ in which $-\infty$ is the smallest and ∞ is the largest element.

If $R, R_1, \ldots, R_r$ are sets, such that each of the sets $R_1, \ldots, R_r$ contains a distinguished element denoted 0, and if $f_1 : R^n \to R_1$, $\ldots$, $f_r : R^n \to R_r$ are functions (possibly given as polynomials), we define

$$Z(f_1, \ldots, f_r) \,=\, Z_{R^n}(f_1, \ldots, f_r) := \big\{x \in R^n \mid f_1(x) = 0, \ldots, f_r(x) = 0\big\}.$$

Let $N \in \mathbb{Z}^+$. For $s, \alpha_1, \ldots, \alpha_N \in \mathbb{N}$ and $r \in \mathbb{R}$, we put

$$\begin{aligned}
\overline{r} &:= \max(r, 0) && \text{and} && \underline{\alpha} &:= (\alpha_1, \alpha_2, \ldots, \alpha_N)\\
r\underline{\alpha} &:= (r\alpha_1, r\alpha_2, \ldots, r\alpha_N) && \text{and} && r^{\underline{\alpha}} &:= (r^{\alpha_1}, r^{\alpha_2}, \ldots, r^{\alpha_N})\\
[s) &:= \{0, \ldots, s-1\} && \text{and} && [s] &:= \{0, \ldots, s\}\\
[\underline{\alpha}) &:= [\alpha_1) \times [\alpha_2) \times \cdots \times [\alpha_N) && \text{and} && [\underline{\alpha}] &:= [\alpha_1] \times [\alpha_2] \times \cdots \times [\alpha_N]\\
|\underline{\alpha}| &:= \alpha_1 + \alpha_2 + \cdots + \alpha_N && \text{and} && \alpha'_s &:= \#\big\{1 \leq t \leq N \mid \alpha_t \geq s\big\}
\end{aligned}$$

If the sequence $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ is monotone decreasing with positive entries, then the sequence $(\alpha'_1, \alpha'_2, \ldots, \alpha'_{\alpha_1})$ is called its **conjugate sequence**; but it will be convenient to also have defined the next term α'_{α_1+1} and sums of the form $\sum_{j=1}^{\alpha'_{\alpha_1+1}} c_j$ (both as equal to 0 if α_1 is maximal among the α_j). The conjugate sequence is again monotone decreasing with positive entries, and the conjugate of the conjugate is the original sequence. This can be seen by interpreting passing to the conjugate sequence as reflecting Ferrers' diagram through the main diagonal.

1.2. **The Story so far.** This paper is a direct continuation of our prior works [CS21] and [CS23a]; in these papers as well as in the present paper, our goal is to synthesize, further develop and apply work of Wilson [Wi06] and Aichinger-Moosbauer [AM21].

In [AM21], Aichinger and Moosbauer develop a calculus of finite differences for maps $f : A \to B$ between arbitrary commutative groups A and B. To every such map f they attach a **functional degree** $\mathrm{fdeg}(f)$ taking values in $\tilde{\mathbb{N}}$ [AM21, §2]. See also the exposition in our prior work [CS21, §2.3].[1] Taking B^A to be the set of all functions $f : A \to B$, which naturally has the structure of a module over the group ring $\mathbb{Z}[A]$, we put

$$\mathcal{F}(A, B) := \{f \in B^A \mid \mathrm{fdeg}(f) < \infty\},$$

which then is a $\mathbb{Z}[A]$-submodule of B^A. One of the key insights of [AM21] is that it is often fruitful to view the elements of $\mathcal{F}(A, B)$ as the "polynomial functions from A to B." If P is a polynomial expression in n variables with coefficients in a (not necessarily

[1] In [AM21] $\mathrm{fdeg}(0) := 0$, and $\mathrm{fdeg}(f) \in \mathbb{N} \cup \{\infty\}$ for all f, but we set $\mathrm{fdeg}(0) := -\infty$.

commutative) rng[2] R, and $E(P) : R^n \to R$ is the corresponding polynomial function, then [AM21, Lemma 12.5]

$$\mathrm{fdeg}(E(P)) \leq \deg(P).$$

It is an interesting problem to precisely understand the discrepancy between these two kinds of degree. After work of Aichinger-Moosbauer [AM21, §10] and work of the present authors [CS23a, Prop. 2.19 and Thm. 4.9], we know how to compute $\mathrm{fdeg}(E(P))$ in terms of the support of the monomial of P when R is any commutative integral domain. Equality holds in all cases if and only if R has characteristic 0.

For commutative groups A and B, we put

$$\delta(A, B) \coloneqq \sup_{f \in B^A} \mathrm{fdeg}(f).$$

When A and B are nontrivial finite commutative groups, Aichinger-Moosbauer showed that $\delta(A, B)$ is finite if and only if A and B are both p-groups for the same prime number p, and they raised the question of determining the exact value of $\delta(A, B)$ in this case. This was answered by the present authors:

Theorem 1.1. *Let $N, \beta, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$, and let B be a finite commutative p-group of exponent p^β. Then*

$$\delta(\bigoplus_{i=1}^{N} \mathbb{Z}/p^{\alpha_i}\mathbb{Z}, B) = \delta_p(\underline{\alpha}, \beta) \coloneqq \sum_{i=1}^{N} (p^{\alpha_i} - 1) + (\beta - 1)(p - 1)p^{\max\{\alpha_1, \ldots, \alpha_N\} - 1}.$$

Proof. This is [CS21, Thm. 4.9 c)]. □

For finite commutative p-groups A and B, the quantity $\delta(A, B)$ can be interpreted as the "maximum complexity" for a map $f : A \to B$. For instance, if $A = (\mathbb{Z}/p\mathbb{Z})^n$ and $B = \mathbb{Z}/p\mathbb{Z}$, then (as Aichinger-Moosbauer knew) the largest possible functional degree is $(p - 1)n$, and one function of this degree is given by evaluating the polynomial $t_1^{p-1} \cdots t_n^{p-1}$. This is related to an observation of Chevalley: over a finite field $\mathbb{F}_q$, the function $x \mapsto x^q - x$ is identically zero, so for any polynomial $P \in \mathbb{F}_q[t_1, \ldots, t_n]$ there is another "reduced polynomial" $P_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ consisting of monomial terms $t_1^{a_1} \cdots t_n^{a_n}$ with $0 \leq a_i \leq q - 1$ and such that $E(P) = E(P_r)$, i.e., the two polynomials determine the same polynomial function. The largest degree of a reduced monomial is therefore $(q-1)n$.

Already this hints that the Aichinger-Moosbauer functional calculus should have numerous theoretic connections, in particular to the following celebrated results.

Theorem 1.2. *Let $q \coloneqq p^N$. Let $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ be nonzero polynomials. If $\sum_{j=1}^{r} \deg(f_j) < n$, then*

a) $\mathrm{ord}_p(\#Z_{\mathbb{F}_q^n}(f_1, \ldots, f_r)) \geq 1$ (Chevalley-Warning Theorem [Ch35], [Wa35]),

b) $\mathrm{ord}_q(\#Z_{\mathbb{F}_q^n}(f_1, \ldots, f_r)) \geq \left\lceil \frac{n - \sum_{j=1}^{r} \deg(f_j)}{\max_{j=1}^{r} \deg(f_j)} \right\rceil$ (Ax-Katz Theorem [Ax64], [Ka71]).

Indeed, Aichinger-Moosbauer used their functional calculus to prove the following result, a striking generalization of Theorem 1.2 a).

[2]Not a typo: a ring has a multiplicative identity, a rng may not.

Theorem 1.3. *(Group-Theoretic Chevalley-Warning Theorem) Let*

$$A := \bigoplus_{i=1}^{m} \mathbb{Z}/p^{\alpha_i}\mathbb{Z}, \quad B := \bigoplus_{i=1}^{n} \mathbb{Z}/p^{\beta_i}\mathbb{Z}$$

be finite commutative p-groups, and let $f_1, \ldots, f_r : A^N \to B$ be functions. If

$$\left(\sum_{j=1}^{r} \operatorname{fdeg}(f_j)\right)\left(\sum_{i=1}^{n}(p^{\beta_i}-1)\right) < \left(\sum_{i=1}^{m} p^{\alpha_i} - 1\right)N$$

then

$$\operatorname{ord}_p(\#Z_{A^N}(f_1, \ldots, f_r)) \geq 1.$$

Proof. This is [AM21, Thm. 12.2]. □

The same work [AM21] gave a group-theoretic generalization of Warning's Second Theorem [AM21, Thm. 14.2] but left open the problem of applying their calculus to higher p-adic congruences. However, a 2006 work of R. Wilson [Wi06] gave a strikingly new and elementary proof of Theorem 1.2 b) over the prime field $\mathbb{F}_p$ using, in particular, the difference operator $\Delta_1 : f \mapsto f(x+1) - f(x)$. Comparing the work of Wilson with that of Aichinger-Moosbauer, we found that – notwithstanding some differences in perspective and presentation – they are deeply related. Our proof of Theorem 1.1 makes use either of Wilson's work (or, alternately, earlier related work of Weisman [We77]). Moreover, with some further development of the Achinger-Moosbauer calculus – especially that for any commutative group B, elements of $\mathcal{F}(\mathbb{Z}^N, B)$ have series expansions (this will be recalled later as Theorem 2.1) – we were able [CS23a, Cor. 1.9] to refine Wilson's argument to give the following group-theoretic generalization of Theorem 1.2 b) over $\mathbb{F}_p$.

Theorem 1.4 (Group-Theoretic Prime Ax-Katz Theorem)**.** *Let $N, n, r \in \mathbb{Z}^+$, and put $A := (\mathbb{Z}/p\mathbb{Z})^N$. Let $f_1, \ldots, f_r \in A^{A^n}$ be nonzero functions. Then*

$$\operatorname{ord}_p(\#Z_{A^n}(f_1, \ldots, f_r)) \geq \left\lceil \frac{N\left(n - \sum_{j=1}^{r} \operatorname{fdeg}(f_j)\right)}{\max_{j=1}^{r} \operatorname{fdeg}(f_j)} \right\rceil.$$

We emphasize that like Theorem 1.3 , Theorem 1.4 is a purely group-theoretic result. When $A = \mathbb{F}_p$ we recover Theorem 1.2b) over the prime field $\mathbb{F}_p$. When $A = \mathbb{F}_q$, because of the connection between the functional degree and the p-weight degree, it recovers Moreno-Moreno's strengthening of the prime Ax-Katz Theorem [MM95], which however does not imply the full Ax-Katz Theorem over $\mathbb{F}_q$ (cf. [CS23a, Remark 1.4]).

After seeing a related manuscript of Grynkiewicz [Gr22], we noticed that the argument that proves Theorem 1.4 can be adapted to prove a more general result for maps $f_j : (\mathbb{Z}/p\mathbb{Z})^N \to B_j$ where each B_j is a finite commutative p-group. This setting is indeed more general than the one in Theorem 1.4 , as the new N may play the role of the old nN. It is, however, more general than actually needed, as one can always reduce (Remark 1.6) to the case of cyclic p-groups B_j. For cyclic p-groups B_j the result is:

Theorem 1.5. [CS23a, Thm. 1.7] *Let $N, r, \beta_1, \ldots, \beta_r \in \mathbb{Z}^+$, and put $A := (\mathbb{Z}/p\mathbb{Z})^N$. For each $1 \leq j \leq r$, let $f_j \in (\mathbb{Z}/p^{\beta_j}\mathbb{Z})^A$ be a nonzero function. Then*

$$\operatorname{ord}_p(\#Z_A(f_1, \ldots, f_r)) \geq \left\lceil \frac{N - \sum_{j=1}^{r} \frac{p^{\beta_j}-1}{p-1} \operatorname{fdeg}(f_j)}{\max_{j=1}^{r} p^{\beta_j - 1} \operatorname{fdeg}(f_j)} \right\rceil.$$

Remark 1.6. *For a finite commutative p-group $B = \bigoplus_{j=1}^{K} B_j$ and each $1 \leq k \leq K$, let $\pi_k : \bigoplus_{j=1}^{K} B_j \to B_k$ be the k^{th} coordinate projection. Because we can exchange a map $f : A \to \bigoplus_{j=1}^{K} B_j$ for the tuple $(f_k)_{k=1}^{K}$ of maps $f_k \coloneqq \pi_k \circ f$, and because we have*

$$\operatorname{fdeg}(f) \;=\; \max_{1 \leq k \leq K} \operatorname{fdeg}(f_k)\,,$$

Theorem 1.5 applies to treat the case of arbitrary finite commutative p-groups B_j as mentioned above: see [CS23a, Rem. 1.8 & Cor. 1.9] *for the details.*

1.3. **The Main Theorem.** Now we generalize to arbitrary finite commutative p-groups, to obtain a generalized Ax-Katz type lower bound on $\operatorname{ord}_p(\#Z(f_1, \ldots, f_r))$ as main result of this paper. Without loss of generality (Remark 1.6), we presume that all codomains B_j are cyclic p-groups. Our generalization concerns the common domain A of our functions $f_j : A \to B_j$. We consider the groups

$$A := \bigoplus_{i=1}^{N} \mathbb{Z}/p^{\alpha_i}\mathbb{Z}$$

and

$$B_1 := \mathbb{Z}/p^{\beta_1}\mathbb{Z}\,,\; \ldots\;,\; B_r := \mathbb{Z}/p^{\beta_r}\mathbb{Z}$$

where $r, \beta_1, \ldots, \beta_r, N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$. We presume that each f_j is a nonconstant of functional degree at most $d_j \in \mathbb{Z}^+$. In other words, for each $1 \leq j \leq r$ we have a function

$$f_j : A \to B_j \quad \text{with} \quad 0 \,<\, \operatorname{fdeg}(f_j) \,\leq\, d_j\,,$$

and we may order these f_j and the α_i such that

$$d_1 p^{\beta_1} \geq d_2 p^{\beta_2} \geq \cdots \geq d_r p^{\beta_r} \quad \text{and} \quad \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_N\,.$$

To express our complicated result, we use the notational definitions of Section 1.1, including conjugate numbers and over-bar notation. We also need several new parameters. We set

$$\begin{aligned} \alpha \;&:=\; \alpha_1 + \alpha_2 + \cdots + \alpha_N \\ &=\; \alpha'_1 + \alpha'_2 + \cdots + \alpha'_{\alpha_1} \quad \text{(by Ex. 3.3)} \end{aligned}$$

and

$$\begin{aligned} \breve{\alpha} \;&:=\; \breve{\alpha}_1 + \breve{\alpha}_2 + \cdots + \breve{\alpha}_N \\ &=\; \alpha'_1 + \alpha'_2 + \cdots + \alpha'_{\breve{\alpha}_1} \quad \text{(by Ex. 3.4)} \end{aligned}$$

where

$$\begin{aligned} \breve{\alpha}_i \;&:=\; \min\bigl\{\alpha_i\,, L\bigr\} \\ L \;&:=\; \beta_1 + \lfloor \log_p\bigl(d_1\bigr) \rfloor\,. \end{aligned}$$

Using that $\alpha'_1 + \alpha'_2 + \cdots + \alpha'_{\alpha_1} = \alpha$, we define numbers $D_1, D_2, \ldots, D_\alpha$ by setting

$$\bigl(D_1, D_2, \ldots, D_\alpha\bigr) \;:=\; \bigl(\underbrace{1, 1, \ldots, 1}_{\alpha'_1 \text{ times}}, \underbrace{p, p, \ldots, p}_{\alpha'_2 \text{ times}}, \;\ldots,\; \underbrace{p^{\alpha_1 - 1}, p^{\alpha_1 - 1}, \ldots, p^{\alpha_1 - 1}}_{\alpha'_{\alpha_1} \text{ times}}\bigr).$$

We further put

$$\begin{aligned}\breve{\mathcal{A}} &:= \sum_{i=1}^{N} \frac{p^{\breve{\alpha}_i}-1}{p-1} \\ &= \alpha'_1 p^0 + \cdots + \alpha'_{\breve{\alpha}_1} p^{\breve{\alpha}_1 - 1} \quad \text{(by Ex. 3.3)} \\ &= D_1 + \cdots + D_{\breve{\alpha}}\end{aligned}$$

and

$$\mathcal{B} := \sum_{j=1}^{r} d_j \frac{p^{\beta_j}-1}{p-1}\,.$$

Theorem 1.7. *With the parameters and settings above,*

$$\operatorname{ord}_p\big(\#Z_A(f_1,\dots,f_r)\big) \geq \begin{cases} \left\lceil \dfrac{\breve{\mathcal{A}}-\mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil + \alpha - \breve{\alpha} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \alpha - \max\big\{1 \leq t \leq \alpha \mid D_1 + \cdots + D_t \leq \mathcal{B}\big\} & \text{if } \breve{\mathcal{A}} \leq \mathcal{B}. \end{cases}$$

Not that our lower bound is equal to 0 if $\mathcal{A} \leq \mathcal{B}$, where $\mathcal{A} := \sum_{i=1}^{N} \frac{p^{\alpha_i}-1}{p-1} \geq \breve{\mathcal{A}}$, because then

$$D_1 + \cdots + D_\alpha \;=\; \alpha'_1 p^0 + \cdots + \alpha'_{\alpha_1} p^{\alpha_1 - 1} \;=\; \sum_{i=1}^{N} \sum_{j=0}^{\alpha_i - 1} p^j \;=\; \mathcal{A} \;\leq\; \mathcal{B}\,.$$

We also want to mention that FFFFFalse

$$\left\lceil \frac{\breve{\mathcal{A}}-\mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil + \alpha - \breve{\alpha} \;\leq\; \left\lceil \frac{\mathcal{A}-\mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil,$$

which makes one wonder whether the second simpler term still is a lower bound.

The result takes a somewhat simpler form when $\alpha_1 = \cdots = \alpha_N$. In that case,

$$\alpha'_1 \;=\; \cdots \;=\; \alpha'_{\alpha_1} \;=\; N \quad \text{and} \quad \breve{\mathcal{A}} \;=\; N \frac{p^{\breve{\alpha}_1}-1}{p-1}\,.$$

With the parameters

$$Q := \big\lfloor \log_p\big((p-1)\mathcal{B}/N + 1\big) \big\rfloor \quad \text{and} \quad R := \left\lfloor \frac{\mathcal{B} - N\frac{p^Q-1}{p-1}}{p^Q} \right\rfloor$$

we may also restate the second expression in the theorem:

$$\alpha - \max\big\{1 \leq t \leq \alpha \mid D_1 + \cdots + D_t \leq \mathcal{B}\big\} \;=\; \mathcal{V}_p(\underline{\alpha}, (p-1)\mathcal{B}) \;=\; \overline{N(\alpha_1 - Q) - R}\,,$$

by the last formula in Theorem 4.3 and the first formula in Corollary 4.5. In the case $\alpha_1 = \cdots = \alpha_N$,

$$D_1 + \cdots + D_{NQ} = N\frac{p^Q-1}{p-1} \leq N\frac{\big((p-1)\mathcal{B}/N+1\big)-1}{p-1} = \mathcal{B}$$

$$D_1 + \cdots + D_{NQ+R} = N\frac{p^Q-1}{p-1} + Rp^Q \leq N\frac{p^Q-1}{p-1} + \mathcal{B} - N\frac{p^Q-1}{p-1} = \mathcal{B}$$

Corollary 1.8. *If* $\alpha_1 = \cdots = \alpha_N$ *then*

$$\mathrm{ord}_p\big(\#Z_A(f_1,\ldots,f_r)\big) \geq \begin{cases} \left\lceil \frac{N\frac{p^{\breve{\alpha}_1}-1}{p-1} - \mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil + N(\alpha_1 - \breve{\alpha}_1) & \text{if } N\frac{p^{\breve{\alpha}_1}-1}{p-1} > \mathcal{B}, \\ N(\alpha_1 - Q) - R & \text{if } N\frac{p^{\breve{\alpha}_1}-1}{p-1} \leq \mathcal{B}. \end{cases}$$

Remark 1.9.

a) *The special case of Corollary 1.8 in which* $\alpha_1 = \ldots = \alpha_N = 1$ *(and* $d_j = \mathrm{fdeg}(f_j)$ *for* $1 \leq j \leq r$*) is Theorem 1.5. Indeed, in that case* $\breve{\alpha}_1 = 1$*, and then both lower bounds (the one in Theorem 1.5 and the one in Corollary 1.8) are obviously equal if* $N > \mathcal{B}$*, while they are both zero or negative if* $N \leq \mathcal{B}$*.*

b) *Before* [CS23a]*, Ax-Katz type p-adic congruences on the solution set of a polynomial system over a finite rng were only known for finite commutative rings in which every ideal is principal* [Ax64]*,* [Ka71]*,* [MR75]*,* [Ka12]*.*

Now let R *be a finite rng with order a power of* p*, so there are* $N, \alpha_1, \ldots, \alpha_N$ *such that*

$$(R,+) \cong \bigoplus_{i=1}^{N} \mathbb{Z}/p^{\alpha_i}\mathbb{Z} =: A_1\,.$$

Let $P_1, \ldots, P_r$ *be polynomials in* n *variables over* R *with* $\deg(P_j) \leq d_j$ *for each* $1 \leq j \leq r$*. Then Theorem 1.7 with* $A := A_1^n$ *and Remark 1.6 apply to give an Ax-Katz type lower bound on* $\mathrm{ord}_p(\#Z(P_1,\ldots,P_r))$*. In particular, as* $\breve{\mathcal{A}} \geq n$*, one sees the following* ***asymptotic Ax-Katz phenomenon****: if* r *and* $d_1, \ldots, d_r$ *remain fixed, then* $\mathrm{ord}_p(\#Z(P_1,\ldots,P_r))$ *approaches infinity with* n*.*

c) *Let* $A, B_1, \ldots, B_r$ *be any nontrivial finite commutative groups, and write out the primes dividing* $\#\left(A \times \prod_{i=1}^r B_i\right)$ *as* $\ell_1 < \ldots < \ell_s$*. Then for each* $1 \leq j \leq r$ *we have a canonical* $\mathbb{Z}$*-module injection*

$$\prod_{k=1}^{s} B_j[\ell_k^\infty]^{A[\ell_k^\infty]} \to B_j^A$$

in which we send the vector $\big(g_k : A[\ell_k^\infty] \to B_j[\ell_k^\infty]\big)_{1\leq k\leq s}$ *to the function*

$$(x_1,\ldots,x_s) \mapsto (g_1(x_1),\ldots,g_s(x_s))\,.$$

By [CS21, Cor. 3.15c]*, upon restriction to functions of finite functional degree, we get canonical isomorphy that we write as an equality:*

$$\mathcal{F}(A, B_j) \;=\; \prod_{k=1}^{s} \mathcal{F}(A[\ell_k^\infty], B_j[\ell_k^\infty])$$

in which moreover $\mathrm{fdeg}(g_1,\ldots,g_s) = \max\{\mathrm{fdeg}(g_k) \mid 1 \leq k \leq s\}$*. In other words, a map of finite functional degree between any two finite commutative groups is determined by its restrictions to the primary components of its domain and its functional degree is simply the largest functional degree of any primary component.*

If we now consider all r *maps* $\big(f_j : A \to B_j\big)_{1\leq j\leq r}$ *then we get* $s \times r$ *primary component maps*

$$\big(g_{j,l} : A[\ell_k^\infty] \to B_j[\ell_k^\infty]\big)_{\substack{1\leq j\leq r \\ 1\leq k\leq s}}$$

and it is immediate that

$$\#Z(f_1,\dots,f_r) \;=\; \prod_{k=1}^{s} \#Z(g_{1,k},\dots,g_{r,k})\,.$$

So, Theorem 1.7 gives lower bounds on $\mathrm{ord}_{\ell_k}(\#Z(f_1,\dots,f_r))$ *for all* $1 \le k \le s$ *in terms of* $A, B_1, \dots, B_r$ *and* $\mathrm{fdeg}(f_1),\dots,\mathrm{fdeg}(f_r)$.

d) *The two previous remarks can be combined to address the case of polynomial expressions in* n *variables of degrees* $d_1,\dots,d_r$ *over any finite rng* R. *(At least in the case when* R *is a commutative ring, the appeal to* [CS21, Cor. 3.15c]*) can be replaced by the more familiar observation that* R *has a canonical direct product decomposition* $R = \prod_{k=1}^{s} \mathfrak{r}_l$ *into rings of prime power order, such that*

$$R[t_1,\dots,t_n] \;=\; \prod_{k=1}^{s} \mathfrak{r}_k[t_1,\dots,t_n],$$

which already gives "Cartesian decomposition" for polynomial functions.) In this fully general case, the asymptotic Ax-Katz phenomenon can be expressed as follows: keeping the number and degrees of the polynomial expressions $f_1,\dots,f_r$ *fixed, we find that* $\#Z(f_1,\dots,f_r)$ *becomes divisible by an arbitrarily large power of* $\#R$ *as the number of variables becomes sufficiently large.*

1.4. **Schedule of Remaining Tasks.** In Section 2 we will explicitly evaluate $\nu_p(\underline{\alpha},\underline{n})$ and $\mathcal{V}_p(D)$ and thereby complete Parts 1 and 2. In Section 3 we compute $\min_{\underline{n}\in[\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n})$ and thereby complete Part 3. In the brief final Section 4, we will show that this completes the proof of Theorem 1.7, and we will prove Corollary 1.8.

2. Reduction to Discrete Optimization

In this section we start the proof of Theorem 1.7. It is a full proof, except that some results about minima of certain discrete functions from later sections are cited. So, this main part of the proof reduces us to some discrete optimization problems, which can be stated and solved completely independent from the original problem. We first recall some basics from and our earlier work, then introduce some basic number theoretic results, and then start that reductionistic proof.

2.1. **Some Recalled Results.** We denote by ord_p the p-**adic valuation** on $\mathbb{Q}$, and with $\mathbb{Z}_{(p)}$ for the ring of rational numbers of non-negative p-adic valuation. We set

$$\mathbb{Z}^+ := \{n \in \mathbb{Z} \mid n > 0\}\,,\quad \mathbb{N} := \{n \in \mathbb{Z} \mid n \ge 0\} \quad\text{and}\quad \tilde{\mathbb{N}} := \mathbb{N} \cup \{-\infty, \infty\}\,,$$

and endow $\tilde{\mathbb{N}}$ with the total ordering extending the usual one on $\mathbb{N}$ in which $-\infty$ is the smallest and ∞ is the largest element.

Let $N \in \mathbb{Z}^+$. For $s, \alpha_1, \dots, \alpha_N \in \mathbb{N}$ and $r \in \mathbb{R}$, we put

$$\begin{aligned}
\overline{r} &:= \max(r,0) && \text{and} && \underline{\alpha} &:= (\alpha_1,\alpha_2,\dots,\alpha_N)\\
r\underline{\alpha} &:= (r\alpha_1, r\alpha_2, \dots, r\alpha_N) && \text{and} && r^{\underline{\alpha}} &:= (r^{\alpha_1}, r^{\alpha_2}, \dots, r^{\alpha_N})\\
[s) &:= \{0,\dots,s-1\} && \text{and} && [s] &:= \{0,\dots,s\}\\
[\underline{\alpha}) &:= [\alpha_1) \times [\alpha_2) \times \cdots \times [\alpha_N) && \text{and} && [\underline{\alpha}] &:= [\alpha_1] \times [\alpha_2] \times \cdots \times [\alpha_N]\\
|\underline{\alpha}| &:= \alpha_1 + \alpha_2 + \cdots + \alpha_N && \text{and} && \alpha'_s &:= \#\big\{1 \le t \le N \mid \alpha_t \ge s\big\}
\end{aligned}$$

Theorem 2.1. *Let B be a commutative group, and let $f \in B^{\mathbb{Z}^N}$.*

a) *There is a unique function $a_\bullet : \mathbb{N}^N \to B$ such that*

$$f(\underline{x}) = \sum_{\underline{n} \in \mathbb{N}^N} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N} a_{\underline{n}} \quad \text{for all } \underline{x} \in \mathbb{N}^N.$$

The function values of $a_\bullet$ are given by the formula $a_{\underline{n}} = \Delta^{\underline{n}} f(\underline{0})$.

b) *If $d := \mathrm{fdeg}(f) < \infty$, then*

$$f(\underline{x}) = \sum_{\substack{\underline{n} \in \mathbb{N}^N \\ |\underline{n}| \leq d}} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N} \Delta^{\underline{n}} f(\underline{0}) \quad \text{for all } \underline{x} \in \mathbb{Z}^N.$$

Proof. This is [CS23a, Thm. 2.8]. □

We now recall some terminology and results concerning proper lifts. Let $\mu : B \to B'$ be a surjective homomorphism of commutative groups, and let $f \in \mathcal{F}(\mathbb{Z}^N, B')$. By Theorem 2.1 there is a unique function $a_\bullet : \mathbb{N}^N \to B'$ that is finitely nonzero (i.e., its support $\{\underline{n} \in \mathbb{N}^N \mid a_{\underline{n}} \neq 0\}$ is finite) such that

$$f(\underline{x}) = \sum_{\substack{\underline{n} \in \mathbb{N}^N \\ |\underline{n}| \leq d}} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N} \Delta^{\underline{n}} f(\underline{0}) \quad \text{for all } \underline{x} \in \mathbb{Z}^N.$$

Then a **proper lift** of $a_\bullet$ to B is a finitely nonzero function $\tilde{a}_\bullet : \mathbb{N}^N \to B$ such that $\mu \circ \tilde{a}_\bullet = a_\bullet$ and for all $\underline{n} \in \mathbb{N}^N$, $\tilde{a}_{\underline{n}} = 0 \iff a_{\underline{n}} = 0$. Proper lifts always exist. To a proper lift $\tilde{a}_\bullet$ of $a_\bullet$ we attach the function $\tilde{f} \in B^{\mathbb{Z}^N}$ defined by

$$\tilde{f}(\underline{x}) := \sum_{\underline{n} \in \mathbb{N}^N} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N} \tilde{a}_{\underline{n}}\,.$$

We then have $f = \mu \circ \tilde{f}$ and $\mathrm{fdeg}(\tilde{f}) = \mathrm{fdeg}(f)$.

Corollary 2.2. *Let $N, \beta, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$. Let $f : \bigoplus_{i=1}^N \mathbb{Z}/p^{\alpha_i}\mathbb{Z} \to \mathbb{Z}/p^\beta\mathbb{Z}$ be any function, $F : \mathbb{Z}^N \to \mathbb{Z}/p^\beta\mathbb{Z}$ be the pullback of f, and $\tilde{F} : \mathbb{Z}^N \to \mathbb{Z}$ be a proper lift of F.*

a)

$$\tilde{F}(\underline{x}) = \sum_{\substack{\underline{n} \in \mathbb{N}^N \\ |\underline{n}| \leq \delta_p(\underline{\alpha}, \beta)}} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N} \Delta^{\underline{n}} \tilde{F}(\underline{0}) \quad \text{for all } \underline{x} \in \mathbb{Z}^N.$$

b) *For all $h \in \mathbb{Z}^+$ and all $\underline{n} \in \mathbb{N}^n$ with*

$$|\underline{n}| > \delta_p(\underline{\alpha}, h) := \sum_{i=1}^N (p^{\alpha_i} - 1) + (h-1)(p-1)p^{\max\{\alpha_1, \ldots, \alpha_N\} - 1},$$

we have

$$p^h \mid \Delta^{\underline{n}} \tilde{F}(\underline{0})\,.$$

Proof. This is [CS23a, Cor. 2.25]. □

2.2. **The numbers $\nu_p(\underline{\alpha}, \underline{n})$ and $\mathcal{V}_p(\underline{\alpha}, D)$.** For $\alpha \in \mathbb{Z}^+$ and $n \in \mathbb{N}$, we put

$$\nu_p(\alpha, n) := \operatorname{ord}_p\left(\sum_{x \in [p^\alpha)} \binom{x}{n}\right).$$

Lemma 2.3. *For each $\alpha \in \mathbb{Z}^+$ and $n \in \mathbb{N}$,*

$$\nu_p(\alpha, n) = \begin{cases} \alpha - \operatorname{ord}_p(n+1) & \text{if } n \leq p^\alpha - 1, \\ \infty & \text{otherwise.} \end{cases}$$

Proof. The case $n = 0$ is handled by Remark 2.4a), while if $n \geq p^\alpha$ then $\sum_{x \in [p^\alpha)} \binom{x}{n} = 0$, so $\nu_p(\alpha, n) = \infty$. So we may assume that $1 \leq n \leq p^\alpha - 1$. Using Pascal's rule $\binom{a}{b} = \binom{a-1}{b} + \binom{a-1}{b-1}$ we see that

$$\begin{aligned} \sum_{x=0}^{p^\alpha - 1} \binom{x}{n} &= \binom{n+1}{n+1} + \binom{n+1}{n} + \binom{n+2}{n} + \binom{n+3}{n} + \cdots + \binom{p^\alpha - 1}{n} \\ &= \binom{n+2}{n+1} + \binom{n+2}{n} + \binom{n+3}{n} + \cdots + \binom{p^\alpha - 1}{n} \\ &\vdots \\ &= \binom{p^\alpha - 1}{n+1} + \binom{p^\alpha - 1}{n} \\ &= \binom{p^\alpha}{n+1}. \end{aligned}$$

Now we can apply Kummer's insight [Ku52] that the p-adic valuation of a binomial coefficient $\binom{a}{b}$ is the number of carries when b and $a-b$ are added in base p. Since the base p representation of p^α is $1000 \cdots 0$ with α zeros, this number of carries is $\alpha - \operatorname{ord}_p(n+1)$ in our case. □

For $N \in \mathbb{Z}^+$, $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N) \in (\mathbb{Z}^+)^N$ and $\underline{n} = (n_1, \ldots, n_N) \in \mathbb{N}^N$, we put

$$\nu_p(\underline{\alpha}, \underline{n}) := \operatorname{ord}_p\left(\sum_{\underline{x} \in [p^{\underline{\alpha}})} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N}\right) = \sum_{i=1}^N \nu_p(\alpha_i, n_i).$$

To any $D \in \mathbb{N} \cup \{\infty\}$, we also define

$$\mathcal{V}_p(\underline{\alpha}, D) := \min\left\{\nu_p(\underline{\alpha}, \underline{n}) \;\middle|\; |\underline{n}| \leq D\right\},$$

which is always finite and zero if $D = \infty$, as we see next:

Proposition 2.4. *Let $\alpha \in \mathbb{Z}^+$, $\underline{\alpha} \in (\mathbb{Z}^+)^N$, and $D \in \mathbb{N} \cup \{\infty\}$.*

a)
$$v_p(\alpha, 0) = \operatorname{ord}_p\left(\sum_{x \in [p^\alpha)} \binom{x}{0}\right) = \operatorname{ord}_p(p^\alpha) = \alpha,$$

and thus
$$v_p(\underline{\alpha}, \underline{0}) = |\underline{\alpha}|$$

and
$$\mathcal{V}_p(\underline{\alpha}, D) \leq |\underline{\alpha}|.$$

b)
$$v_p(\alpha, p^\alpha - 1) \;=\; \mathrm{ord}_p\Big(\sum_{x\in[p^\alpha)} \binom{x}{p^\alpha - 1}\Big) \;=\; \mathrm{ord}_p(1) \;=\; 0\,,$$

and thus
$$v_p\big(\underline{\alpha}, (p^{\alpha_1}-1, \ldots, p^{\alpha_N}-1)\big) \;=\; 0$$

and
$$D \geq \sum_{i=1}^N (p^{\alpha_i}-1) \quad\Longrightarrow\quad \mathcal{V}_p(\underline{\alpha}, D) = 0\,.$$

c) *Keeping $\underline{\alpha}$ fixed, $\mathcal{V}_p(\underline{\alpha}, D)$ is monotonically decreasing in D.*

As we already determined $\nu_p(\underline{\alpha}, \underline{n}) = \sum_{i=1}^N \nu_p(\alpha_i, n_i)$ in the lemma above, the precise calculation of $\mathcal{V}_p(\underline{\alpha}, D)$ is mere discrete optimization. We will do that in the next section, in Theorem 4.3. The result can be stated in various different forms. The one that we can read without further definitions, just the ones already used in Theorem 1.7, is the following:

$$\mathcal{V}_p(\underline{\alpha}, D) \;=\; \alpha - \max\big\{0 \leq j \leq \alpha \mid D_1 + \cdots + D_j \leq \tfrac{D}{p-1}\big\}\,. \tag{1}$$

2.3. **The integral $\int_S f$.** Let A and B be commutative groups, let $f \in B^A$, and let $S \subseteq A$ be a finite subset. Following [KP12], we set
$$\int_S f \;\coloneqq\; \sum_{x\in S} f(x) \;\in\; B \quad \text{and} \quad \int f \;\coloneqq\; \int_A f\,.$$

Here we are mostly interested in the case $A = \mathbb{Z}^N$, $B = \mathbb{Z}$ and $S = [p^{\underline{\alpha}})$. The following results generalize work of Wilson [Wi06, Lemma 4] that treats the case $\alpha_1 = \cdots = \alpha_N = 1$.

Proposition 2.5. *Let $D \in \mathbb{N} \cup \{\infty\}$ and $N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$. If $f \in \mathbb{Z}^{\mathbb{Z}^N}$ has functional degree* $\mathrm{fdeg}(f) \leq D$*, then*
$$\mathrm{ord}_p\bigg(\int_{[p^{\underline{\alpha}})} f\bigg) \;\geq\; \mathcal{V}_p\big(\underline{\alpha}, D\big).$$

Proof. For commutative groups A and B and a finite subset $S \subseteq A$, the map $\int_S : B^A \to B$ is a $\mathbb{Z}$-module homomorphism - and this also holds when $A = \mathbb{Z}^N$, $B = \mathbb{Z}$, and $S = [p^{\underline{\alpha}})$. By Theorem 2.1, it therefore suffices to prove the inequality for functions of the form
$$\underline{x} \;\mapsto\; \binom{x_1}{n_1}\cdots\binom{x_N}{n_N}$$
with $|\underline{n}| \leq D$. This, however, is easy:
$$\mathrm{ord}_p\bigg(\int_{[p^{\underline{\alpha}})} \binom{x_1}{n_1}\cdots\binom{x_N}{n_N}\bigg) = \prod_{j=1}^N \sum_{x_j\in[p^{\alpha_j})} \binom{x_j}{n_j} = \nu_p\big(\underline{\alpha}, \underline{n}\big) \;\geq\; \mathcal{V}_p\big(\underline{\alpha}, |\underline{n}|\big) \;\geq\; \mathcal{V}_p\big(\underline{\alpha}, D\big)\,.$$

□

2.4. **The Proof of Theorem 1.7.** Below is the proof of Theorem 1.7, modulo two main discrete optimization tasks. On one side, it shows that the broad outline of the argument is the same as that of Theorem 1.5, using the key ideas from Wilson's proof of Ax-Katz over $\mathbb{F}_p$. On the other side, it motivates and sets up the new work of the present paper, those two optimization tasks, which are needed to complete the argument.

Proof of Theorem 1.7. Recall that $A = \bigoplus_{i=1}^N \mathbb{Z}/p^{\alpha_i}\mathbb{Z}$. The quotient map $q : \mathbb{Z}^N \to A$ restricted to $[p^{\underline{\alpha}})$ induces a bijection $[p^{\underline{\alpha}}) \to A$, $\underline{x} \mapsto q(\underline{x})$.

For each $\beta \in \mathbb{Z}^+$ and $1 \le j \le r$ put

$$\hat{n}_j(\beta) := (p^{\beta_j} - 1) + (\beta - 1)p^{\beta_j - 1}(p-1)\,,$$

define $\chi_j : \mathbb{Z} \to \mathbb{Z}/p^\beta\mathbb{Z}$ by

$$\chi_j(x) := \begin{cases} 1 & \text{if } x \equiv 0 \pmod{p^{\beta_j}}, \\ 0 & \text{otherwise,} \end{cases}$$

and let $\widetilde{\chi}_j : \mathbb{Z} \to \mathbb{Z}$ be a proper lift of χ_j from $\mathbb{Z}/p^\beta\mathbb{Z}$ to $\mathbb{Z}$.

If $A_1, \dots, A_n$ are commutative groups and R is a rng, then (as in [AM21, §6]) the tensor product $\bigotimes_{i=1}^n h_i$ of maps $h_i : A_i \to R$ is the map

$$\bigotimes_{i=1}^n h_i : \bigoplus_{i=1}^n A_i \to R\,, \ \ (x_1, \dots, x_n) \mapsto h_1(x_1) \cdots h_n(x_n)\,.$$

Let $\chi : \mathbb{Z}^r \to \mathbb{Z}/p^\beta\mathbb{Z}$ be the tensor product $\bigotimes_{j=1}^r \chi_j$ of the χ_j, and let $\widetilde{\chi} : \mathbb{Z}^r \to \mathbb{Z}$ be the tensor product $\bigotimes_{j=1}^r \widetilde{\chi}_j$ of the $\widetilde{\chi}_j$. Moreover let $\tilde{f}_j : A \to \mathbb{Z}$ be a proper lift of $f_j : A \to \mathbb{Z}/p^{\beta_j}\mathbb{Z}$, and let $\tilde{F}_j : \mathbb{Z}^N \to \mathbb{Z}$ be obtained from $\tilde{f}_j$ by pulling back from A to $\mathbb{Z}^N$. Then, for each $\underline{x} \in [p^{\underline{\alpha}})$,

$$\chi(\tilde{F}_1(\underline{x}), \dots, \tilde{F}_r(\underline{x})) = \begin{cases} 1 & \text{if } q(\underline{x}) \in Z(f_1, \dots, f_r), \\ 0 & \text{otherwise,} \end{cases}$$

and thus, with the function $\tilde{\chi}(\tilde{F}_1, \dots, \tilde{F}_r) : \underline{x} \mapsto \tilde{\chi}(\tilde{F}_1(\underline{x}), \dots, \tilde{F}_r(\underline{x}))$,

$$\#Z(f_1, \dots, f_r) = kp^\beta + \int_{[p^{\underline{\alpha}})} \widetilde{\chi}(\tilde{F}_1, \dots, \tilde{F}_r) \quad \text{for some } k \in \mathbb{Z}\,.$$

We may certainly assume that $Z(f_1, \dots, f_r)$ is nonempty, so that $\mathrm{ord}_p\big(\#Z(f_1, \dots, f_r)\big)$ is finite. Hence, it is possible to choose $\beta \in \mathbb{Z}^+$ such that

$$\beta > \mathrm{ord}_p\big(\#Z(f_1, \dots, f_r)\big) \in \mathbb{N}\,, \tag{2}$$

and then we get

$$\mathrm{ord}_p\big(\#Z(f_1, \dots, f_r)\big) = \mathrm{ord}_p\bigg(\int_{[p^{\underline{\alpha}})} \widetilde{\chi}(\tilde{F}_1, \dots, \tilde{F}_r)\bigg)\,.$$

Now, for each $1 \le j \le r$, Corollary 2.2 provides a function $c_j : [\hat{n}_j(\beta)] \to \mathbb{Z}$ such that, for each $x \in \mathbb{Z}$,

$$\tilde{\chi}_j(x) = \sum_{n \in [\hat{n}_j(\beta)]} \binom{x}{n} c_j(n)\,.$$

In particular, with $\hat{\underline{n}}(\beta) \coloneqq (\hat{n}_1(\beta), \dots, \hat{n}_r(\beta))$, for each $\underline{x} \in \mathbb{Z}^N$,

$$\begin{aligned}\tilde{\chi}(\tilde{F}_1(\underline{x}), \dots, \tilde{F}_r(\underline{x})) &= \tilde{\chi}_1(\tilde{F}_1(\underline{x})) \cdots \tilde{\chi}_r(\tilde{F}_r(\underline{x})) \\ &= \sum_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \binom{\tilde{F}_1(\underline{x})}{n_1} \cdots \binom{\tilde{F}_r(\underline{x})}{n_r} c_1(n_1) \cdots c_r(n_r)\,.\end{aligned}$$

Hence, with functions $\binom{\tilde{F}_j}{n_j} : \underline{x} \mapsto \binom{\tilde{F}_1(\underline{x})}{n_j}$,

$$\int_{[p^{\underline{\alpha}})} \tilde{\chi}(\tilde{F}_1, \dots, \tilde{F}_r) = \sum_{\underline{n} \in [\hat{\underline{n}}(\beta)]} c_1(n_1) \cdots c_r(n_r) \int_{[p^{\underline{\alpha}})} \binom{\tilde{F}_1}{n_1} \cdots \binom{\tilde{F}_r}{n_r}.$$

Thus if we put

$$\mathbf{m} \coloneqq \min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \Big(\mathrm{ord}_p\big(c_1(n_1)\big) + \cdots + \mathrm{ord}_p\big(c_r(n_r)\big) + \mathrm{ord}_p\Big(\int_{[p^{\underline{\alpha}})} \binom{\tilde{F}_1}{n_1} \cdots \binom{\tilde{F}_r}{n_r} \Big) \Big),$$

it follows that

$$\mathrm{ord}_p\big(\#Z(f_1, \dots, f_r)\big) = \mathrm{ord}_p\Big(\int_{[p^{\underline{\alpha}})} \widetilde{\chi}(\tilde{F}_1, \dots, \tilde{F}_r) \Big) \geq \mathbf{m}\,.$$

Thus the matter of it is to give a good lower bound on the quantity $\mathbf{m}$, using that $\mathrm{fdeg}(\tilde{F}_j) = \mathrm{fdeg}(f_j) \leq d_j$ for all $1 \leq j \leq r$ (cf. [CS23a, Cor. 2.13, §2.4 and §2.5]). Part of this can be quickly done in the same way as in [CS23a]: Corollary 2.2 also says that the functions $c_j : [\hat{n}_j(\beta)] \to \mathbb{Z}$ can be chosen such that, for each $h \in \mathbb{Z}^+$ and $n \in [\hat{n}_j(\beta)]$,

$$p^{\beta_j} - 1 + p^{\beta_j - 1}(p-1)(h-1) < n \implies p^h \mid c_j(n)\,.$$

Thus, taking

$$h_j = h_j(n_j) \coloneqq \left\lceil \frac{n_j - (p^{\beta_j} - 1)}{p^{\beta_j - 1}(p-1)} \right\rceil < \frac{n_j - (p^{\beta_j} - 1)}{p^{\beta_j - 1}(p-1)} + 1,$$

we have

$$p^{\beta_j} - 1 + (h_j - 1)p^{\beta_j - 1}(p-1) < n_j\,,$$

and thus Corollary 2.2 yields

$$\mathrm{ord}_p\big(c_j(n_j)\big) \geq \overline{h_j} = \overline{\left\lceil \frac{n_j - (p^{\beta_j} - 1)}{p^{\beta_j - 1}(p-1)} \right\rceil}\,.$$

Moreover, using [AM21, Thm. 4.3 and Lem. 6.1], we have

$$\mathrm{fdeg}\Big(\binom{\tilde{F}_1}{n_1} \cdots \binom{\tilde{F}_r}{n_r} \Big) \leq \sum_{j=1}^{r} d_j n_j\,,$$

and Proposition 2.5 shows that

$$\mathrm{ord}_p\Big(\int_{[p^{\underline{\alpha}})} \binom{\tilde{F}_1}{n_1} \cdots \binom{\tilde{F}_r}{n_r} \Big) \geq \mathcal{V}_p\big(\underline{\alpha}, \textstyle\sum_{j=1}^{r} d_j n_j\big)\,.$$

We deduce that

$$\mathrm{ord}_p(\#Z(f_1, \dots, f_r)) \geq \mathbf{m} \geq \min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n})$$

where

$$\mathcal{N}(\underline{n}) := \sum_{j=1}^{r} \overline{\left\lceil \frac{n_j - (p^{\beta_j} - 1)}{p^{\beta_j - 1}(p-1)} \right\rceil} + \mathcal{V}_p\Big(\underline{\alpha},\, \textstyle\sum_{j=1}^{r} d_j n_j\Big).$$

The precise calculation of $\min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n})$ is mere discrete optimization. We will do that in the next section, in Lemma 5.2. After increasing β if necessary[3], it yields

$$\min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n}) = \begin{cases} \left\lceil \frac{\breve{\mathcal{A}} - \mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil + \sum_{i=1}^{N} \overline{\alpha_i - \breve{\alpha}_1} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \alpha - \max\big\{1 \le j \le \alpha \mid D_1 + \cdots + D_j \le \mathcal{B}\big\} & \text{if } \breve{\mathcal{A}} \le \mathcal{B}. \end{cases}$$

Notice that the answer obtained is independent of β. □

The rest of the proof of Theorem 1.7 then breaks into the following parts:

PART 1: Compute $\nu_p(\underline{\alpha}, \underline{n})$ for all p, $\underline{\alpha}$ and $\underline{n}$.

PART 2: Minimize $\underline{n} \mapsto \nu_p(\underline{\alpha}, \underline{n})$ over all $\underline{n}$ with $|\underline{n}| \le D$ to compute $\mathcal{V}_p(\underline{\alpha}, D)$.

PART 3: For large β, minimize $\underline{n} \mapsto \mathcal{N}(\underline{n})$ over $\underline{n} \in [\hat{\underline{n}}(\beta)]$ to compute our lower bound for $\mathbf{m}$.

It turns out that Part 1 follows easily from some classical number theory. Solving the discrete optimization problems in Parts 2 and 3 requires siginficantly more work and forms the main content of this paper.

3. Conjugate Sequences

As preparation for the discrete optimization tasks ahead, we provide here more results about conjugate sequences. Again, $N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$ with $\alpha_1 \ge \cdots \ge \alpha_N$ are fixed given, and the conjugate numbers $\alpha'_1, \ldots, \alpha'_{\alpha_1}$ (and $\alpha'_{\alpha_1+1} = \alpha'_{\alpha_1+2} \cdots := 0$) are defined as in Section 1.1. The finite sequence $(\alpha'_j) \in (\mathbb{Z}^+)^{\alpha_1}$ is also called the **conjugate partition** of $(\alpha_i) \in (\mathbb{Z}^+)^N$, since both sequences partition the number $\alpha := \alpha_1 + \alpha_2 + \cdots + \alpha_N$, as we will see. Exactly as transposed matrices, conjugate partitions are simply obtained by reflecting the Ferrers diagram about the main diagonal:

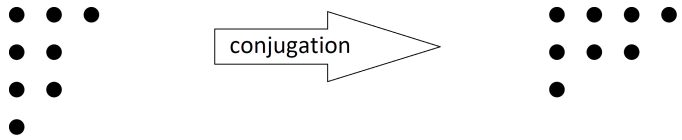

FIGURE 1. The conjugate of $(3, 2, 2, 1)$ is $(4, 3, 1)$.

[3]We need $\beta > s_0$ in Lemma 5.2. Within the full proof of Theorem 1.7, however, we presume $\beta > \mathrm{ord}_p(\#Z)$ already in (2), and $\mathrm{ord}_p(\#Z) \ge s_0$ by the findings of this paper.

3.1. **Two General Conjugation Lemmas.** The following lemma is formulated in a way that is helpful in our calculations.

Lemma 3.1. *Let $(a_i) \in (\mathbb{Z}^+)^N$ be monotone decreasing, and $1 \le m \le a_1$ ($1 \le m \le \max_{1\le i\le N} a_i$). We have the following identity in $\mathbb{Z}[x]$:*

$$a'_m x^m + a'_{m+1}x^{m+1} + \cdots + a'_{a_1}x^{a_1} \;=\; \sum_{i=1}^{a'_m} \big(x^m + x^{m+1} + \cdots + x^{a_i}\big)\,.$$

Proof. Both polynomials have degree at most a_1, and there are no monomials of degree less than m. Let $m \le j \le a_1$. Then the coefficient of x^j in the standard form of the right polynomial is

$$\#\big\{1 \le i \le a'_m \mid a_i \ge j\big\} \;=\; \#\big\{1 \le i \le N \mid a_i \ge j\big\} \;=\; a'_j\,,$$

because

$$a_i \ge j \quad\Longrightarrow\quad a_i \ge m \quad\Longrightarrow\quad a_1, \ldots, a_i \ge m \quad\Longrightarrow\quad i \le a'_m\,.$$

This is the same as the coefficient of x^j in the left polynomial. □

The following lemma is obvious if we imagine taking the minimum as intersecting two Ferrers diagrams, as "intersecting" and "reflecting" commute.

Lemma 3.2. *If the two sequences $(a_i), (b_i) \in (\mathbb{Z}^+)^N$ are monotone decreasing, then the sequence $(c_i) := \big(\min(a_i, b_i)\big) \in (\mathbb{Z}^+)^N$ is also monotone decreasing, its conjugate sequence (c'_j) has length $\min(a_1, b_1)$, and*

$$c'_j \;=\; \min(a'_j, b'_j) \quad \textit{for all } 1 \le j \le \min(a_1, b_1)\,.$$

3.2. **Special Cases.**

Example 3.3. *Setting x equal to 1 in Lemma 3.1, we obtain*

$$\sum_{j=m}^{a_1} a'_j \;=\; \sum_{i=1}^{a'_m} \big(a_i - m + 1\big)\,,$$

and, if also $m = 1$, we get

$$\sum_{j=1}^{a_1} a'_j \;=\; \sum_{i=1}^{N} a_i\,.$$

Example 3.4. *If the sequence (b_i) in Lemma 3.2 is constant equal to $\breve{\alpha}_1 \le a'_1$, we obtain as the conjugate of the sequence $(c_i) := \big(\min(a_i, \breve{\alpha}_1)\big) \in (\mathbb{Z}^+)^N$ the sequence $(a'_1, a'_2, \ldots, a'_{\breve{\alpha}_1})$. So, if we apply Lemma 3.1 to (c_i) with $m = 1$ and $x = p$, we get*

$$a'_1 p^0 + \cdots + a'_{\breve{\alpha}_1} p^{\breve{\alpha}_1 - 1} \;=\; \sum_{i=1}^{N} \frac{p^{\min\{a_i, \breve{\alpha}_1\}} - 1}{p - 1}\,.$$

Example 3.5. *(X Const. with one step ...) If the sequence (b_i) in Lemma 3.2 is constant equal to $\breve{\alpha}_1 \le a'_1$, we obtain as the conjugate of the sequence $(c_i) := \big(\min(a_i, \breve{\alpha}_1)\big) \in (\mathbb{Z}^+)^N$ the sequence $(a'_1, a'_2, \ldots, a'_{\breve{\alpha}_1})$. So, if we apply Lemma 3.1 to (c_i) with $m = 1$ and $x = p$, we get*

$$a'_1 p^0 + \cdots + a'_{\breve{\alpha}_1} p^{\breve{\alpha}_1 - 1} \;=\; \sum_{i=1}^{N} \frac{p^{\min\{a_i, \breve{\alpha}_1\}} - 1}{p - 1}\,.$$

4. Minimization of $\nu_p(\underline{\alpha}, \bullet)$

In this section, we determine the minimum value $\mathcal{V}_p(\underline{\alpha}, D)$ of the function $\nu_p(\underline{\alpha}, \bullet)$ over the restricted domain

$$\mathcal{D}(N, D) := \big\{ \underline{n} \in \mathbb{N}^N \,\big|\, |\underline{n}| \le D \big\},$$

where the numbers $D \in \mathbb{N}$, and $N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$ with $\alpha_1 \ge \cdots \ge \alpha_N$ are fixed given. In this regard, the original definition of $\nu_p(\underline{\alpha}, \bullet)$ does not matter. We may view the formula in Lemma 2.3 as the definition. More precisely, for $\underline{n} \in \mathbb{N}^N$,

$$\nu_p(\underline{\alpha}, \underline{n}) := \sum_{i=1}^{N} \nu_p(\alpha_i, n_i) \quad \text{with} \quad \nu_p(\alpha_i, n_i) := \begin{cases} \alpha_i - \operatorname{ord}_p(n_i + 1) & \text{if } n_i \le p^{\alpha_i} - 1, \\ \infty & \text{otherwise.} \end{cases}$$

As mentioned earlier, the case $D = \infty$ is trivial and can be excluded. We have $\mathcal{V}_p(\underline{\alpha}, \infty) = 0$, and more generally

$$D \ge (p^{\alpha_1} - 1) + \cdots + (p^{\alpha_N} - 1) \quad \Longrightarrow \quad \mathcal{V}_p(\underline{\alpha}, D) \,=\, 0\,,$$

because $\nu_p\big(\underline{\alpha}, (p^{\alpha_1} - 1, \ldots, p^{\alpha_N} - 1)\big) = 0$.

4.1. **The Minimum Value $\mathcal{V}_p(\underline{\alpha}, D)$ of $\nu_p(\underline{\alpha}, \bullet)$ over $\mathcal{D}(N, D)$.** Next, we determine the minimum value $\mathcal{V}_p(\underline{\alpha}, D)$ for $D \in \mathbb{N}$. With $\alpha'_{\alpha_1+1} := 0$, and with sums of the form $\sum_{j=1}^{\alpha'_{\alpha_1+1}}$ regarded as empty with value zero, we first deduce a formula for $\mathcal{V}_p(\underline{\alpha}, D)$ in terms of the parameters D_j of Theorem 1.7. Again, $N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$ are such that $\alpha_1 \ge \cdots \ge \alpha_N$.

Theorem 4.1. *For each fixed $D \in \mathbb{N}$, the function*

$$\nu_p(\underline{\alpha}, \bullet)\big|_{\mathcal{D}(N,D)} \colon \mathcal{D}(N, D) \longrightarrow \mathbb{N} \cup \{\infty\}, \quad \underline{n} \longmapsto \nu_p(\underline{\alpha}, \underline{n})$$

has minimum value

$$\mathcal{V}_p(\underline{\alpha}, D) \,=\, \alpha - \max\big\{ 0 \le j \le \alpha \mid D_1 + \cdots + D_j \le \tfrac{D}{p-1} \big\}\,.$$

Proof. We may restrict the domain of $\nu_p(\underline{\alpha}, \bullet)$ from $\mathcal{D}(N, D)$ to $\mathcal{D}(N, D) \cap [p^{\underline{\alpha}})$ with $[p^{\underline{\alpha}}) := \prod_{i=1}^{N} [p^{\alpha_i})$, because $\nu_p(\underline{\alpha}, \bullet)$ is finite inside but positive infinite outside of $[p^{\underline{\alpha}})$. Inside $[p^{\underline{\alpha}})$, however,

$$\nu_p(\underline{\alpha}, \underline{n}) = \alpha - \sum_{i=1}^{N} \operatorname{ord}_p(n_i + 1)\,.$$

So, we need to find a maximum point of the function

$$\mathcal{D}(N, D) \cap [p^{\underline{\alpha}}) \longrightarrow \mathbb{N}\,, \quad \underline{n} \longmapsto \sum_{i=1}^{N} \operatorname{ord}_p(n_i + 1)\,.$$

Now, if $\underline{m} = (m_i)_{i=1}^N \in \mathcal{D}(N, D) \cap [p^{\underline{\alpha}})$ is a maximum point, then the point $\underline{\tilde{m}} = (\tilde{m}_i)_{i=1}^N$ with $\tilde{m}_i + 1 := p^{\operatorname{ord}_p(m_i+1)}$ is also a maximum point, because $\operatorname{ord}_p(\tilde{m}_i + 1) = \operatorname{ord}_p(m_i + 1)$ and $0 \le \tilde{m}_i \le m_i$ for all $1 \le i \le N$. Hence, we may restrict our attention to points $\underline{m}$ such that each $m_i + 1$ is a power of p, say $m_i = p^{\mu_i} - 1$, and then $\operatorname{ord}_p(m_i + 1) = \mu_i$. Hence, with the substitutions $m_i := p^{\mu_i} - 1$ in mind, we just have to find the maximum of the function

$$\sigma : \big\{ \underline{\mu} \in [\underline{\alpha}] \,\big|\, \omega(\underline{\mu}) \le \tfrac{D}{p-1} \big\} \longrightarrow \mathbb{N}\,, \quad \underline{\mu} \longmapsto \sigma(\underline{\mu}) := \sum_{i=1}^{N} \mu_i\,,$$

where

$$\omega(\underline{\mu}) \,:=\, \sum_{i=1}^{N} \frac{p^{\mu_i}-1}{p-1} \,=\, \sum_{i=1}^{N} \sum_{j=0}^{\mu_i-1} p^j \,=\, \sum_{i=1}^{N} \sum_{j=1}^{\mu_i} p^{j-1}.$$

Now, if we draw Ferrers diagram for a potential argument $\underline{\mu}$ of σ, as sub-diagram of Ferrers diagram of $\underline{\alpha}$, then $\sigma(\underline{\mu})$ is the number of dots in that sub-diagram, while $\omega(\underline{\mu})$ gives a weighted count of those dots: a dot in the j^{th} column is counted with weight p^{j-1}, as shown in Figure 2. Hence, to find the maximum of σ, we need to maximize the number of dots in the sub-diagram corresponding to $\underline{\mu}$, while keeping their weight $\omega(\underline{\mu})$ below $\frac{D}{p-1}$.

Our constructive idea is to start from zero and increase the number of dots inside the diagram of $\underline{\mu}$ one by one, following a certain order, till the threshold $\frac{D}{p-1}$ for $\omega(\underline{\mu})$ is reached or would next be exceeded. The number of dots that can be selected within that limit only depends on the order in which we pick the dots. It is possible to reach the maximum value of σ with the right order. Obviously, the outcome of our construction is optimal (in this sense) if, in each step, the newly added dot has lowest possible weight among all remaining unselected dots (since this keeps $\omega(\underline{\mu})$ as small as possible). This is an optimality criteria that may apply to some orders of selection. In the order that we describe next, it is obviously met.

In our situation of column-wise increasing weights, we select the dots column by column, from left to right, starting with the left-most column of lowest weight. Insight a column the order of selection does not matter, as long as the column is completely finished before we move to the next column. We may just go top-down inside columns, as in Figure 2. Following that order, we collect in step t a dot of weight D_t, because that is how we defined D_t. Hence, after t steps we obtain a $\underline{\mu} = \underline{\mu}(t)$ with

$$\omega(\underline{\mu}) = D_1 + \cdots + D_t \quad \text{and} \quad \sigma(\underline{\mu}) = t\,.$$

Our selection process has to stop when the limit $\frac{D}{p-1}$ for $\omega(\underline{\mu})$ is reached, i.e. when

$$t \,=\, \max\big\{0 \le t \le \alpha \mid D_1 + \cdots + D_t \le \tfrac{D}{p-1}\big\}\,.$$

At that point, $\underline{\mu}$ is a maximum point of σ, and the associated $\underline{m} := (p^{\mu_i}-1)_{i=1}^N$ is a minimum point of $\nu_p(\underline{\alpha}, \bullet)$ in $\mathcal{D}(N,D)$. The minimum value is

$$\mathcal{V}_p(\underline{\alpha}, D) \,=\, \nu_p(\underline{\alpha}, \underline{m}) \,=\, \alpha - \sigma(\underline{\mu}) \,=\, \alpha - \max\big\{0 \le t \le \alpha \mid D_1 + \cdots + D_t \le \tfrac{D}{p-1}\big\}\,.$$

□

Remark 4.2.

a) *In the last proof we also constructed a minimum point $\underline{m} = \underline{m}(D)$ of the function $\mathcal{D}(N,D) \to \mathbb{N} \cup \{\infty\}$, $\underline{n} \mapsto \nu_p(\underline{\alpha}, \underline{n})$, for each given $\alpha_1, \ldots, \alpha_N$ and $D \in \mathbb{N}$. The minimum value is assumed at the point*

$$\underline{m}(D) \,:=\, \big(p^{\alpha_i(D)} - 1\big)_{i=1}^{N}\,,$$

where

$$\alpha_i(D) \,:=\, \begin{cases} Q(D)+1 & \text{if } 1 \le i \le R(D) \\ Q(D) & \text{if } R(D) < i \le \alpha'_{Q(D)+1} \\ \alpha_i & \text{if } \alpha'_{Q(D)+1} < i \le N \end{cases}$$

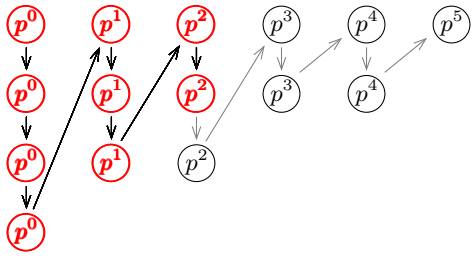

FIGURE 2. The minimum weight of a set of 9 dots inside $\underline{\alpha} = (6, 5, 3, 1)$ is $D_1 + D_2 + \cdots + D_9 = 4 + 3p + 2p^2$.

with

$$\begin{aligned} Q(D) &:= \max\Big\{0 \le Q \le \alpha_1 \;\Big|\; \textstyle\sum_{j=1}^{Q} \alpha'_j p^{j-1} \le \frac{D}{p-1}\Big\} \\ &= \max\Big\{0 \le Q \le \alpha_1 \;\Big|\; D_1 + \cdots + D_{\alpha'_1+\cdots+\alpha'_Q} \le \frac{D}{p-1}\Big\} \end{aligned}$$

and

$$\begin{aligned} R(D) &:= \max\Big\{0 \le R \le \alpha'_{Q(D)+1} \;\Big|\; \textstyle\sum_{j=1}^{Q(D)} \alpha'_j p^{j-1} + R\, p^{Q(D)} \le \frac{D}{p-1}\Big\} \\ &= \max\Big\{0 \le R \le \alpha'_{Q(D)+1} \;\Big|\; D_1 + \cdots + D_{\alpha'_1+\cdots+\alpha'_{Q(D)}+R} \le \frac{D}{p-1}\Big\}. \end{aligned}$$

Note that if $Q(D) = \alpha_1$ then $\alpha'_{Q(D)+1} = 0$ by definition, and $R(D) = 0 = \alpha'_{Q(D)+1}$ in this case. In all other cases $R(D) < \alpha'_{Q(D)+1}$ by the maximality of $Q(D)$.

b) *Using the notations above, it is not hard to see that our formula for $\mathcal{V}_p(\underline{\alpha}, D)$ can be expressed in the following forms:*

$$\begin{aligned} \mathcal{V}_p(\underline{\alpha}, D) &= \sum_{i=1}^{N} \big(\alpha_i - \alpha_i(D)\big) \\ &= \sum_{i=1}^{\alpha'_{Q(D)+1}} \alpha_i \; - \alpha'_{Q(D)+1} Q(D) - R(D) \\ &= \sum_{i=1}^{\alpha'_{Q(D)}} \alpha_i \; - \alpha'_{Q(D)} Q(D) - R(D) \\ &= \sum_{j=Q(D)+1}^{\alpha_1} \alpha'_j \; - R(D) \end{aligned}$$

4.2. **X (remove partially) The Minimum Value $\mathcal{V}_p(\underline{\alpha}, D)$ of $\nu_p(\underline{\alpha}, \bullet)$ over $\mathcal{D}(N, D)$.** Now we are ready to determine the minimum value

$$\mathcal{V}_p(\underline{\alpha}, D) \;=\; \nu_p(\underline{\alpha}, \underline{m}(D))\,,$$

for $D \in \mathbb{N}$. With $\alpha'_{\alpha_1+1} := 0$, and with sums of the form $\sum_{j=1}^{\alpha'_{\alpha_1+1}}$ regarded as empty with value zero, we can deduce four expressions for $\mathcal{V}_p(\underline{\alpha}, D)$ in terms of the parameters of Theorem 4.1. In addition, the next theorem contains one formula for $\mathcal{V}_p(\underline{\alpha}, D)$ in

terms of the parameters D_j of Theorem 1.7. Again, $N, \alpha_1, \dots, \alpha_N \in \mathbb{Z}^+$ are such that $\alpha_1 \geq \cdots \geq \alpha_N$.

Theorem 4.3.

$$\begin{aligned}
\mathcal{V}_p(\underline{\alpha}, D) &= \sum_{i=1}^{N} \big(\alpha_i - \alpha_i(D)\big) \\
&= \sum_{i=1}^{\alpha'_{Q(D)+1}} \alpha_i \; - \alpha'_{Q(D)+1} Q(D) - R(D) \\
&= \sum_{i=1}^{\alpha'_{Q(D)}} \alpha_i \; - \alpha'_{Q(D)} Q(D) - R(D) \\
&= \sum_{j=Q(D)+1}^{\alpha_1} \alpha'_j \; - R(D) \\
&= \alpha - \max\big\{0 \leq j \leq \alpha \mid D_1 + \cdots + D_j \leq \tfrac{D}{p-1}\big\}
\end{aligned}$$

Proof. The first expression for $\mathcal{V}_p(\underline{\alpha}, D)$ can be calculated based on Theorem 4.1 with the help of Lemma 2.3:

$$\mathcal{V}_p(\underline{\alpha}, D) \,=\, \nu_p\big(\underline{\alpha}, \underline{m}(D)\big) \,=\, \sum_{i=1}^{N} \big(\alpha_i - \mathrm{ord}_p(p^{\alpha_i(D)} - 1 + 1)\big) \,=\, \sum_{i=1}^{N} \big(\alpha_i - \alpha_i(D)\big).$$

To obtain the second expression, observe that, by the definition of the $\alpha_i(D)$,

$$\begin{aligned}
\sum_{i=1}^{N} \big(\alpha_i - \alpha_i(D)\big) &= \sum_{i=1}^{R(D)} \big(\alpha_i - Q(D) - 1\big) \,+ \sum_{i=R(D)+1}^{\alpha'_{Q(D)+1}} \big(\alpha_i - Q(D)\big) \\
&= \sum_{i=1}^{\alpha'_{Q(D)+1}} \big(\alpha_i - Q(D)\big) \; - R(D) \\
&= \sum_{i=1}^{\alpha'_{Q(D)+1}} \alpha_i \; - \alpha'_{Q(D)+1} Q(D) - R(D).
\end{aligned}$$

The third expression for $\mathcal{V}_p(\underline{\alpha}, D)$ can directly be derived from the second expression, as $\alpha_i = Q(D)$ for all $\alpha'_{Q(D)+1} < i \leq \alpha'_{Q(D)}$.

We deduce the fourth expression from the expression with the summands $(\alpha_i - Q(D))$ five lines above, as by Lemma 3.1 with $m := Q(D) + 1$ and $x := 1$ we have that (and this is vacuously true if $Q(D) = \alpha_1$)

$$\sum_{i=1}^{\alpha'_{Q(D)+1}} \big(\alpha_i - Q(D)\big) \,=\, \sum_{j=Q(D)+1}^{\alpha_1} \alpha'_j \,.$$

The last expression, involving the sequence (D_j) with $\alpha_1 + \cdots + \alpha_N = \alpha'_1 + \cdots + \alpha'_{\alpha_1}$ entries, we get from our first expression $\sum_{i=1}^{N} \big(\alpha_i - \alpha_i(D)\big) = \sum_{i=1}^{N} \alpha_i - \sum_{i=1}^{N} \alpha_i(D)$ in two

steps. First, as the sequence $\big(\alpha_1(D),\dots,\alpha_N(D)\big)$ has the conjugate $\big(\alpha'_1,\dots,\alpha'_{Q(D)},R(D)\big)$, Lemma 3.1 with $m := 1$ and $x := 1$ tells us that

$$\sum_{i=1}^{N}\alpha_i(D) \;=\; \alpha'_1+\cdots+\alpha'_{Q(D)}+R(D)\,.$$

Second,

$$\alpha'_1+\cdots+\alpha'_{Q(D)}+R(D) \;=\; \max\big\{0\le j\le \textstyle\sum_{i=1}^{N}\alpha_i \mid D_1+\cdots+D_j\le \frac{D}{p-1}\big\}\,,$$

i.e. $\alpha'_1+\cdots+\alpha'_{Q(D)}+R(D)$ is the maximum of all $0\le j\le \sum_{i=1}^{N}\alpha_i=\sum_{j=1}^{\alpha_1}\alpha'_j$ with

$$D_1+\cdots+D_j \;\le\; \frac{D}{p-1}\,,$$

because $Q(D)$ is the maximum of all $0\le Q\le\alpha_1$ with

$$D_1+\cdots+D_{\alpha'_1+\cdots+\alpha'_Q} \;\le\; \frac{D}{p-1}\,,$$

and $R(D)$ is the maximum of all $0\le R\le\alpha'_{Q(D)+1}$ with

$$D_1+\cdots+D_{\alpha'_1+\cdots+\alpha'_{Q(D)}+R} \;\le\; \frac{D}{p-1}\,.$$

□

Using the last formula in Theorem 4.3, we can now show that $D<\sum_{i=1}^{N}\frac{p^{\alpha_i}-1}{p-1}$ is not just necessary for $\mathcal{V}_p(\underline{\alpha},D)>0$, as we already have seen at the start of this section, but it is also sufficient:

Corollary 4.4. *Maintain the setup of Theorem 4.3, we have*

$$\mathcal{V}_p(\underline{\alpha},D)>0 \quad\Longleftrightarrow\quad D<\sum_{i=1}^{N}\big(p^{\alpha_i}-1\big)\,.$$

Proof. It follows from Lemma 3.1 with $m := 1$ and $x := p$ that

$$D_1+\cdots+D_\alpha \;=\; \alpha'_1p^0+\cdots+\alpha'_Np^{N-1} \;=\; \sum_{i=1}^{N}\frac{p^{\alpha_i}-1}{p-1}\,.$$

So, by the last formula of Theorem 4.3,

$$\mathcal{V}_p(\underline{\alpha},D)>0 \;\Leftrightarrow\; \max\big\{0\le j\le\alpha\mid D_1+\cdots+D_j\le\tfrac{D}{p-1}\big\}<\alpha \;\Leftrightarrow\; D<\sum_{i=1}^{N}\big(p^{\alpha_i}-1\big)\,.$$

□

Of particular interest are the following cases:

Corollary 4.5. *If $\alpha_1=\cdots=\alpha_N$ and if we set*

$$Q \;:=\; \big\lfloor\log_p(D/N+1)\big\rfloor \quad\text{ and }\quad R \;:=\; \left\lfloor\frac{D-N(p^Q-1)}{(p-1)p^Q}\right\rfloor,$$

then

$$\mathcal{V}_p(\underline{\alpha},D) \;=\; \overline{N(\alpha_1-Q)-R}\,.$$

If $\alpha_1=\cdots=\alpha_N=1$ then

$$\mathcal{V}_p(\underline{\alpha},D) \;=\; \overline{N-\left\lfloor\frac{D}{p-1}\right\rfloor}\,.$$

Proof. Suppose $\alpha_1 = \cdots = \alpha_N$, i.e. $\alpha'_1 = \cdots = \alpha'_{\alpha_1} = N$.

Case 1, $D < N(p^{\alpha_1} - 1)$**:** In this case, it follows with the same reasoning as at the end of the proof of the theorem that $Q(D) < \alpha_1$ and $R(D) < \alpha'_{Q(D)+1}$. Hence,

$$Q(D) \,=\, \big\lfloor \log_p(D/N + 1) \big\rfloor \,=:\, Q \quad \text{and} \quad R(D) \,=\, \left\lfloor \frac{D - N(p^Q - 1)}{(p-1)p^Q} \right\rfloor \,=:\, R\,.$$

So,

$$\mathcal{V}_p(\underline{\alpha}, D) \,=\, \sum_{j=Q+1}^{\alpha_1} \alpha'_j \;- R \,=\, \sum_{j=Q+1}^{\alpha_1} N \;- R \,=\, N(\alpha_1 - Q) - R \,=\, \overline{N(\alpha_1 - Q) - R}\,,$$

where the last equality follows from $N(\alpha_1 - Q) - R = \mathcal{V}_p(\underline{\alpha}, D) \geq 0$.

In the subcase $\alpha_1 = \cdots = \alpha_N = 1$, this further simplifies to

$$\mathcal{V}_p(\underline{\alpha}, D) \,=\, \overline{N(1-0) - \left\lfloor \frac{D - N(p^0 - 1)}{(p-1)p^0} \right\rfloor} \,=\, \overline{N - \left\lfloor \frac{D}{p-1} \right\rfloor}\,.$$

Case 2, $D \geq N(p^{\alpha_1} - 1)$**:** In this case, $Q \geq Q(D) = \alpha_1$ and thus $\overline{N(\alpha_1 - Q) - R} = 0$. By Corollary 4.4, this is the correct value for $\mathcal{V}_p(\underline{\alpha}, D)$ if $D \geq N(p^{\alpha_1} - 1)$.

The formula for the subcase $\alpha_1 = \cdots = \alpha_N = 1$ also gives the correct value 0 . □

5. Minimization of $\mathcal{N}$

In this section we determine the minimum $\min_{\underline{n} \in [\underline{\hat{n}}(\beta)]} \mathcal{N}(\underline{n})$ of the function

$$\mathcal{N} : [\underline{\hat{n}}(\beta)] \longrightarrow \mathbb{N}\,, \quad \underline{n} \longmapsto \mathcal{N}(\underline{n}) \,:=\, \sum_{j=1}^{r} \overline{\left\lceil \frac{n_j - (p^{\beta_j} - 1)}{p^{\beta_j - 1}(p-1)} \right\rceil} + \mathcal{V}_p\Big(\underline{\alpha},\, \textstyle\sum_{j=1}^r d_j n_j\Big),$$

where (by the last formula in Theorem 4.3)

$$\mathcal{V}_p(\underline{\alpha}, \textstyle\sum_{j=1}^r d_j n_j) \,=\, \alpha - \max\left\{ 0 \leq j \leq \alpha \;\middle|\; D_1 + \cdots + D_j \leq \frac{\sum_{j=1}^r d_j n_j}{p-1} \right\},$$

and where the numbers $\beta, r, \beta_1, \ldots, \beta_r, d_1, \ldots, d_r, N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$ with

$$d_1 p^{\beta_1} \geq d_2 p^{\beta_2} \geq \cdots \geq d_r p^{\beta_r} \quad \text{and} \quad \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_N$$

are fixed given (and β is large enough). Also recall that $\alpha := \sum_{i=1}^N \alpha_i$,

$$[\underline{\hat{n}}(\beta)] \,:=\, \prod_{\ell=1}^{r} \{0, 1, \ldots, \hat{n}_\ell(\beta)\} \quad \text{with} \quad \hat{n}_\ell(\beta) \,:=\, (p^{\beta_\ell} - 1) + (\beta - 1)p^{\beta_\ell - 1}(p-1)\,,$$

and

$$\big(D_1, D_2, \ldots, D_\alpha\big) \,:=\, \big(\underbrace{1, 1, \ldots, 1}_{\alpha'_1 \text{ times}}, \underbrace{p, p, \ldots, p}_{\alpha'_2 \text{ times}}, \ldots, \underbrace{p^{\alpha_1 - 1}, p^{\alpha_1 - 1}, \ldots, p^{\alpha_1 - 1}}_{\alpha'_{\alpha_1} \text{ times}}\big).$$

5.1. **A Preparatory Lemma.** It turns out that the minimization of $\mathcal{N}(\underline{n})$ leads to another optimization problem that can be stated and solved in more general terms as follows:

Lemma 5.1. *Assume $D \in \mathbb{N}$, and let $\alpha, \Lambda_1, \Lambda_2, \ldots, \Lambda_\alpha, V_1, V_2, \ldots \in \mathbb{Z}^+$. Suppose that $(\Lambda_j)_{j=1}^{\alpha}$ is monotone increasing, that $(V_j)_{j=1}^{\infty}$ is monotone decreasing, and that $\Lambda_1 \le V_1$. Also presume that $V_j = V_1$ for all $1 \le j \le s_0$, where*

$$s_0 := \overline{\lceil (\Lambda_1 + \cdots + \Lambda_{i_0} - D)/V_1 \rceil}\,, \quad \text{with } i_0 := \max\{1 \le i \le \alpha \mid \Lambda_i \le V_1\}\,.$$

Then the function $\mathbf{S} : \mathbb{N} \longrightarrow \mathbb{Z}$ given by

$$\mathbf{S}(s) := s - i\big(V_1 + V_2 + \cdots + V_s + D\big),$$

with

$$i(x) := \max\big\{0 \le i \le \alpha \mid \Lambda_1 + \cdots + \Lambda_i \le x\big\},$$

has a minimum at the point s_0, and

$$\mathbf{S}(s_0) = \begin{cases} s_0 - i_0 & \text{if } s_0 > 0, \\ -i(D) & \text{if } s_0 = 0. \end{cases}$$

Proof. By definition, s_0 is the smallest element of $\mathbb{N}$ with $s_0 \ge (\Lambda_1 + \cdots + \Lambda_{i_0} - D)/V_1$, i.e. with

$$\Lambda_1 + \cdots + \Lambda_{i_0} \le s_0 V_1 + D\,. \tag{3}$$

We calculate $\mathbf{S}(s_0)$, $\mathbf{S}(s_0 - s)$ and $\mathbf{S}(s_0 + s)$, for all permissible $s \in \mathbb{Z}^+$, to show that $\mathbf{S}(s_0)$ is a minimum of $\mathbf{S}$. For this purpose it is convenient to extend the sequence $(\Lambda_j)_{j=1}^{\alpha}$ to an infinite sequence by setting $\Lambda_{\alpha+1}, \Lambda_{\alpha+2}, \ldots := \infty$. With that extension

$$i_0 = \max\{i \in \mathbb{Z}^+ \mid \Lambda_i \le V_1\} \quad \text{and} \quad i(x) = \max\big\{i \in \mathbb{N} \mid \Lambda_1 + \cdots + \Lambda_i \le x\big\}.$$

Case 1, $s_0 > 0$**:** In this case, by (3),

$$\Lambda_1 + \cdots + \Lambda_{i_0} \le V_1 + \cdots + V_{s_0} + D \tag{4}$$

but, by the minimality of s_0 in (3), also

$$\Lambda_1 + \cdots + \Lambda_{i_0} > V_1 + \cdots + V_{s_0 - 1} + D\,. \tag{5}$$

In the last inequality, if $s_0 \ge 2$, each summand V_j on the right is at least as large as each of the summands Λ_i on the left, because $\Lambda_1 \le \cdots \le \Lambda_{i_0} \le V_1 = \cdots = V_{s_0 - 1}$. Therefore, we can remove an equal number of those summands on both sides without destroying the inequality. Also, the bigger left sum must contain more of the smaller Λ-summands than the smaller right sum contains of the bigger V-summands, because $D \ge 0$. In particular, for each $0 < s \le s_0$,

$$\Lambda_1 + \cdots + \Lambda_{i_0 - s + 1} > V_1 + \cdots + V_{s_0 - s} + D\,. \tag{6}$$

But, also $\cdots \ge \Lambda_{i_0+2} \ge \Lambda_{i_0+1} > V_1 = V_{s_0} \ge V_{s_0+1} \ge \cdots$. So, we can also add an equal number of subsequent summands on both sides of (5). For each $s \in \mathbb{N}$,

$$\Lambda_1 + \cdots + \Lambda_{i_0 + s + 1} > V_1 + \cdots + V_{s_0 + s} + D\,. \tag{7}$$

Based on these inequalities, we can now calculate $\mathbf{S}(s_0)$, $\mathbf{S}(s_0-s)$ and $\mathbf{S}(s_0+s)$. It follows from (4) and (7) with $s=0$ that

$$\mathbf{S}(s_0) \;=\; s_0 - i_0\,.$$

It follows from (6) that, for each $0 < s \le s_0$,

$$\mathbf{S}(s_0 - s) \;\ge\; s_0 - s - (i_0 - s) \;=\; s_0 - i_0 \;=\; \mathbf{S}(s_0)\,.$$

And, it follows from (7) that, for each $s \in \mathbb{N}$,

$$\mathbf{S}(s_0 + s) \;\ge\; s_0 + s - (i_0 + s) \;=\; \mathbf{S}(s_0)\,.$$

We see that $\mathbf{S}$ attains a minimum at s_0 and $\mathbf{S}(s_0) = s_0 - i_0$.

Case 2, $s_0 = 0$**:** In this case, by (3),

$$\Lambda_1 + \cdots + \Lambda_{i(D)} D \;\le\; D \tag{8}$$

and

$$\Lambda_1 + \cdots + \Lambda_{i(D)+1} \;>\; D\,. \tag{9}$$

So,

$$\mathbf{S}(0) \;=\; 0 - i\big(0 + D\big) \;=\; -i(D)\,.$$

We also have $\cdots \ge \Lambda_{i(D)+2} \ge \Lambda_{i(D)+1} \ge \Lambda_{i_0+1} > V_1 \ge V_2 \ge \cdots$ since $i(D) \ge i_0$, as in this case $\Lambda_1 + \cdots + \Lambda_{i_0} \le D$. Hence, we can add summands to (9), in the same way as we did it to get (7) from (5). For each $s \in \mathbb{N}$,

$$\Lambda_1 + \cdots + \Lambda_{i(D)+s+1} \;>\; V_1 + \cdots + V_s + D\,,$$

and thus

$$\mathbf{S}(s) \;\ge\; s - (i(D) + s) \;=\; \mathbf{S}(0)\,.$$

We see that $\mathbf{S}$ attains a minimum at 0 and $\mathbf{S}(0) = -i(D)$. □

5.2. **The Minimum Value of $\mathcal{N}$ over $[\underline{\hat{n}}(\beta)]$.** Now we are ready to determine the minimum value $\min_{\underline{n} \in [\underline{\hat{n}}(\beta)]} \mathcal{N}(\underline{n})$.

Lemma 5.2. *With the parameters and settings above, and the derived values $\breve{\alpha}_1, \dots, \breve{\alpha}_N, \breve{\alpha}, \breve{\mathcal{A}}, \mathcal{B}$ as in Theorem 1.7. Also let $\beta \in \mathbb{Z}^+$ be such that*

$$\beta \;>\; s_0 \;:=\; \overline{\left\lceil \frac{\breve{\mathcal{A}} - \mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil}\,.$$

If $\breve{\mathcal{A}} > \mathcal{B}$, then

$$\begin{aligned} \min_{\underline{n} \in [\underline{\hat{n}}(\beta)]} \mathcal{N}(\underline{n}) \;&=\; s_0 \;+\; \alpha - \alpha'_1 + \cdots + \alpha'_{\breve{\alpha}_1} \\ &=\; s_0 \;+\; \alpha - \breve{\alpha}\,. \end{aligned}$$

If $\breve{\mathcal{A}} \le \mathcal{B}$, then

$$\begin{aligned} \min_{\underline{n} \in [\underline{\hat{n}}(\beta)]} \mathcal{N}(\underline{n}) \;&=\; \mathcal{V}_p\big(\underline{\alpha}\,,\, (p-1)\mathcal{B}\big) \\ &=\; \alpha \;-\; \max\big\{1 \le j \le \alpha \mid D_1 + \cdots + D_j \le \mathcal{B}\big\}\,. \end{aligned}$$

$$\min_{\underline{n} \in [\underline{\hat{n}}(\beta)]} \mathcal{N}(\underline{n}) \;=\; \begin{cases} s_0 + \alpha - \alpha'_1 + \cdots + \alpha'_{\breve{\alpha}_1} \;=\; s_0 + \alpha - \breve{\alpha} & \text{if } s_0 > 0, \\ \mathcal{V}_p\big(\underline{\alpha}\,,\, (p-1)\mathcal{B}\big) \;=\; \alpha - \max\big\{1 \le j \le \alpha \mid D_1 + \cdots + D_j \le \mathcal{B}\big\} & \text{if } s_0 = 0. \end{cases}$$

Proof. We shrink the domain $[\hat{\underline{n}}(\beta)]$ of the variable $\underline{n}$ till we reach a single point where the minimum is attained and can be calculated. We proceed in four steps.

Step 1: If $n_1 \le p^{\beta_1}-1$ then $\overline{\left\lceil \frac{n_1-(p^{\beta_1}-1)}{(p-1)p^{\beta_1-1}} \right\rceil} = 0$. So, as $\mathcal{V}_p(\underline{\alpha}, \bullet)$ is monotone decreasing,

$$n_1 \le p^{\beta_1}-1 \implies \mathcal{N}(n_1, n_2, \dots, n_r) \ge \mathcal{N}(p^{\beta_1}-1, n_2, \dots, n_r).$$

This shows that, in order to find a minimum, we may replace values of n_1 below $p^{\beta_1}-1$ with $p^{\beta_1}-1 \in [\hat{n}_1(\beta)]$. More generally, for each $1 \le j \le r$, we may assume $n_j \ge p^{\beta_j}-1$ and write n_j as $u_j + p^{\beta_j}-1$ with $u_j \ge 0$, which leads to the simplifications

$$\overline{\left\lceil \frac{n_j-(p^{\beta_j}-1)}{(p-1)p^{\beta_j-1}} \right\rceil} = \overline{\left\lceil \frac{u_j}{(p-1)p^{\beta_j-1}} \right\rceil} = \left\lceil \frac{u_j}{(p-1)p^{\beta_j-1}} \right\rceil$$

and

$$\mathcal{V}_p\Big(\underline{\alpha}, \textstyle\sum_{j=1}^r d_j n_j\Big) = \mathcal{V}_p\Big(\underline{\alpha}, \textstyle\sum_{j=1}^r d_j u_j + (p-1)\mathcal{B}\Big).$$

So, with

$$\mathcal{U}(\underline{u}) := \sum_{j=1}^r \left\lceil \frac{u_j}{(p-1)p^{\beta_j-1}} \right\rceil + \mathcal{V}_p\Big(\underline{\alpha}, \textstyle\sum_{j=1}^r d_j u_j + (p-1)\mathcal{B}\Big)$$

and updated ranges

$$\hat{u}_j(\beta) := \hat{n}_j(\beta) - (p^{\beta_j}-1) = (\beta-1)p^{\beta_j-1}(p-1) \quad \text{for } j = 1, \dots, r,$$

we have $\mathcal{N}(\underline{n}) = \mathcal{U}(\underline{u})$ and

$$\min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n}) = \min_{\underline{u} \in [\hat{\underline{u}}(\beta)]} \mathcal{U}(\underline{u}).$$

Step 2: To find a minimum of $\mathcal{U}$ over $[\hat{\underline{u}}(\beta)] = \prod_{j=1}^r [\hat{u}_j(\beta)]$, we can replace the domain $[\hat{u}_j(\beta)]$ of each u_j with the smaller domain

$$\begin{aligned}[\hat{u}_j(\beta)] \cap p^{\beta_j-1}(p-1)\mathbb{Z} &= \{0, p^{\beta_j-1}(p-1), \dots, (\beta-1)p^{\beta_j-1}(p-1)\} \\ &= p^{\beta_j-1}(p-1)[\beta-1].\end{aligned}$$

If the jth argument $u_j \in [\hat{u}_j]$ of $\mathcal{U}(u_1, \dots, u_r)$ is replaced with the next larger multiple of $p^{\beta_j-1}(p-1)$, which still lies inside $[\hat{u}_j(\beta)] = [(\beta-1)p^{\beta_j-1}(p-1)]$ and can be written as $p^{\beta_j-1}(p-1)\left\lceil \frac{u_j}{p^{\beta_j-1}(p-1)} \right\rceil$, then the summand $\left\lceil \frac{u_j}{p^{\beta_j-1}(p-1)} \right\rceil$ of $\mathcal{U}(\underline{u})$ stays the same and $\mathcal{U}(\underline{u})$ certainly does not increase, i.e.

$$\mathcal{U}\Big(p^{\beta_1-1}(p-1)\Big\lceil \frac{u_1}{p^{\beta_1-1}(p-1)} \Big\rceil, \dots, p^{\beta_r-1}(p-1)\Big\lceil \frac{u_r}{p^{\beta_r-1}(p-1)} \Big\rceil\Big) \le \mathcal{U}(u_1, \dots, u_r).$$

The minimum is already attained over the smaller domain $\prod_{j=1}^r \big(p^{\beta_j-1}(p-1)[\beta-1]\big)$. Hence, with

$$\begin{aligned}\mathcal{T}(t_1, \dots, t_r) &:= \mathcal{U}\big(p^{\beta_1-1}(p-1)t_1, \dots, p^{\beta_r-1}(p-1)t_r\big) \\ &= t_1 + \cdots + t_r + \mathcal{V}_p\Big(\underline{\alpha}, \textstyle\sum_{j=1}^r d_j p^{\beta_j-1}(p-1)t_j + (p-1)\mathcal{B}\Big)\end{aligned}$$

we have

$$\min_{\underline{n} \in [\hat{\underline{n}}]} \mathcal{N}(\underline{n}) = \min_{\underline{u} \in [\hat{\underline{u}}]} \mathcal{U}(\underline{u}) = \min_{\underline{t} \in [\beta-1]^r} \mathcal{T}(\underline{t}).$$

Step 3: In our search for a minimum of $\mathcal{T}(t_1,\dots,t_r)$, we can now modify any two arguments t_i and t_j with $i<j$ by replacing t_j with t_j-1 and t_i with t_i+1. If we view the expression $d_jp^{\beta_j-1}(p-1)t_j$ as sum of t_j equal summands $d_jp^{\beta_j-1}(p-1)$, this step changes one of the t_j summands $d_jp^{\beta_j-1}(p-1)$ inside the argument of $\mathcal{V}_p(\underline{\alpha},\bullet)$ into one additional summand $d_ip^{\beta_i-1}(p-1)$, of which we then have t_i+1. Since

$$d_1p^{\beta_1-1}\geq d_2p^{\beta_2-1}\geq\dots\geq d_rp^{\beta_r-1}$$

and $\mathcal{V}_p(\underline{\alpha},\bullet)$ is monotone decreasing, we have

$$\mathcal{T}(\dots,t_i+1,\dots,t_j-1,\dots)\;\leq\;\mathcal{T}(\dots,t_i,\dots,t_j,\dots).$$

The only restriction to such modifications is that each argument t_j must stay within its domain $[\beta-1]$. Through repeated applications of this modification, we can empty some t_j and fill others. This shows that the minimum is attained at a point $(t_1,\dots,t_r)$ of the form $(\beta-1,\beta-1,\dots,\beta-1,x,0,0,\dots,0)$. At such points, we have

$$\mathcal{T}(t_1,t_2,\dots,t_r)\;=\;s+\mathcal{V}_p\Big(\underline{\alpha}\,,\,(p-1)(V_1+V_2+\dots+V_s+\mathcal{B})\Big),$$

where $s=t_1+t_2+\dots+t_r=\beta-1+\beta-1+\dots+\beta-1+x\leq r(\beta-1)$, and where

$$\big(V_j\big)_{j=1}^{r(\beta-1)}:=\big(\underbrace{d_1p^{\beta_1-1},\dots,d_1p^{\beta_1-1}}_{\beta-1\text{ times}},\;\dots,\;\underbrace{d_rp^{\beta_r-1},\dots,d_rp^{\beta_r-1}}_{\beta-1\text{ times}}\big).$$

Hence, with the function

$$\mathcal{S}:[r(\beta-1)]\to\mathbb{N}\,,\quad\mathcal{S}(s)\;:=\;s+\mathcal{V}_p\Big(\underline{\alpha}\,,\,(p-1)(V_1+V_2+\dots+V_s+\mathcal{B})\Big),$$

we have

$$\min_{\underline{n}\in[\hat{\underline{n}}(\beta)]}\mathcal{N}(\underline{n})\;=\;\min_{\underline{t}\in[\beta-1]^r}\mathcal{T}(\underline{t})\;=\;\min_{s\in[r(\beta-1)]}\mathcal{S}(s).$$

Step 4: To find the minimum of $\mathcal{S}$, we use Lemma 5.1 with $D:=\mathcal{B}$, $\alpha:=\alpha_1+\dots+\alpha_N$, and $\Lambda_j:=D_j$ for all $1\leq j\leq\alpha$. We also use the values V_j as defined above for all $j\leq r(\beta-1)$, and set $V_j:=V_{r(\beta-1)}$ for all $j>r(\beta-1)$. With that infinite sequence $(V_j)_{j=1}^{\infty}$ the domain of $\mathcal{S}$ can be extended to $\mathbb{N}$, with the hope not to alter its minimum in doing so, as the expression

$$\mathcal{V}_p\Big(\underline{\alpha}\,,\,(p-1)(V_1+\dots+V_s+\mathcal{B})\Big)\;=\;\alpha-\max\big\{0\leq i\leq\alpha\mid\Lambda_1+\dots+\Lambda_i\leq V_1+\dots+V_s+\mathcal{B}\big\}$$

makes sense for or all $s\in\mathbb{N}$. The extended function $\mathcal{S}:\mathbb{N}\to\mathbb{N}$ is then almost the same as the function $\mathbf{S}:\mathbb{N}\to\mathbb{Z}$ in Lemma 5.1. For all $s\in\mathbb{N}$,

$$\mathcal{S}(s)\;:=\;\mathbf{S}(s)+\alpha\,.$$

We also have $\Lambda_1\leq V_1$ as required in Lemma 5.1. Moreover, as in our situation the sequence (Λ_i) contains repetitions of length $\alpha_1',\alpha_2',\dots$, the parameter

$$i_0\;:=\;\max\{1\leq i\leq\alpha\mid\Lambda_i\leq V_1\}$$

in Lemma 5.1 can be written as

$$i_0\;=\;\alpha_1'+\dots+\alpha_{\breve{\alpha}_1}',$$

where

$$\breve{\alpha}_1\;:=\;\min\big\{\alpha_i,\beta_1+\lfloor\log_p\big(d_1\big)\rfloor\big\}\;=\;\max\{1\leq j\leq\alpha_1\mid p^{j-1}\leq d_1p^{\beta_1-1}\}\,.$$

Applying Lemma 3.1 with $m := 1$ and $x := p$ to the sequence $(\breve{\alpha}_1, \dots, \breve{\alpha}_N)$ and its conjugate $(\alpha'_1, \dots, \alpha'_{\breve{\alpha}_1})$, we further see that

$$\Lambda_1 + \cdots + \Lambda_{\breve{\alpha}_1} \;=\; \alpha'_1 p^0 + \cdots + \alpha'_{\breve{\alpha}_1} p^{\breve{\alpha}_1 - 1} \;=\; \sum_{i=1}^{N} \frac{p^{\breve{\alpha}_i} - 1}{p-1} \;=:\; \breve{\mathcal{A}}.$$

In particular, the definition of s_0 in Lemma 5.1 coincides with the current one:

$$s_0 \;:=\; \overline{\lceil (\Lambda_1 + \cdots + \Lambda_{i_0} - \mathcal{B})/V_1 \rceil} \;=\; \overline{\left\lceil \frac{\breve{\mathcal{A}} - \mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil} \;\le\; \beta - 1\,.$$

This shows that $V_j = V_1$ for all $1 \le j \le s_0$, as required in that lemma, but it also shows that the minimum point s_0 of $\mathbf{S}$ lies inside $[r(\beta-1)]$. Hence, the minimum point s_0 of $\mathbf{S}$ is also a minimum point of $\mathbf{S}|_{[r(\beta-1)]}$ and of $\mathcal{S}|_{[r(\beta-1)]}$. Thus, Lemma 5.1 yields

$$\begin{aligned}
\min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n}) \;&=\; \min_{s \in [r(\beta-1)]} \mathcal{S}(s) \\
&=\; \mathbf{S}(s_0) \,+\, \alpha \\
&=\; \begin{cases} s_0 + \alpha - i_0 & \text{if } s_0 > 0, \\ \alpha - i(\mathcal{B}) & \text{if } s_0 = 0, \end{cases} \\
&=\; \begin{cases} s_0 + \alpha - \alpha'_1 + \cdots + \alpha'_{\breve{\alpha}_1} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \alpha - \max\{1 \le i \le \alpha \mid D_1 + \cdots + D_i \le \mathcal{B}\} & \text{if } \breve{\mathcal{A}} \le \mathcal{B}, \end{cases} \\
&=\; \begin{cases} s_0 + \alpha - \breve{\alpha} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \mathcal{V}_p(\underline{\alpha}\,,\, (p-1)\mathcal{B}) & \text{if } \breve{\mathcal{A}} \le \mathcal{B}, \end{cases}
\end{aligned}$$

where we used the last formula in Theorem 4.3 and the equation

$$\alpha - \alpha'_1 + \cdots + \alpha'_{\breve{\alpha}_1} \;=\; \sum_{i=1}^{N} \big(\alpha_i - \breve{\alpha}_i\big) \;=\; \sum_{i=1}^{N} \overline{\alpha_i - \breve{\alpha}_1}$$

which follows from Lemma 3.1 with $m := 1$ and $x := 1$ applied to the conjugate $(\alpha'_1, \dots, \alpha'_{\breve{\alpha}_1})$ of the sequence $(\breve{\alpha}_1, \dots, \breve{\alpha}_N)$. □

6. Appendix

When A is finite we define the **summation invariant**

$$\sigma\big(A, B\big) \;:=\; \sup\big\{d \in \mathbb{N} \cup \{-\infty\} \;\big|\; \textstyle\int f = 0 \text{ for all } f \in \mathcal{F}^d(A, B)\big\}.$$

We will not need the following three lemmata, but they shows some connections to Pete's "GeneralizedAx".

Lemma 6.1. *Let $\alpha, \beta \in \mathbb{Z}^+$, then*

$$\eta_p(\alpha, \beta) \;:=\; \min\{n \in \mathbb{N} \mid \nu_p(\alpha, n) < \beta\} \;=\; \begin{cases} p^{\alpha - \beta + 1} - 1 & \text{if } \beta \le \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. It suffices to consider the case $\beta \le \alpha$, as otherwise $\nu_p(\alpha, 0) = \alpha < \beta$. Assume $n \in \mathbb{N}$ is such that $\nu_p(\alpha, n) < \beta$. By our new formula for $\nu_p(\alpha, n)$, this can only be if $n \le p^\alpha - 1$ and $\nu_p(n+1) > \alpha - \beta$. But, $\nu_p(n+1) > \alpha - \beta$ can only be if $n + 1 \ge p^{\alpha - \beta + 1}$,

i.e. if $n \geq p^{\alpha-\beta+1}-1$. Therefore, $\eta_p(\alpha,\beta) \geq p^{\alpha-\beta+1}-1$. Now, if $n := p^{\alpha-\beta+1}-1$ then $\nu_p(\alpha,n) = \beta - 1 < \beta$. So, we also have $\eta_p(\alpha,\beta) \leq p^{\alpha-\beta+1}-1$. □

Lemma 6.2. *Let $\alpha \in \mathbb{Z}^+$ and $n \in \mathbb{N}$. If $n \leq p^\alpha - 1$ then*

$$\nu_p(\alpha,n) \;\geq\; \max\{\beta \in \mathbb{Z} \mid n < p^{\alpha-\beta+1}-1\} \;=\; \alpha - \lfloor \log_p(n+1) \rfloor.$$

Proof. By definition, we have $\nu_p(\alpha,n) \geq \beta$ for at least all $\beta \in \mathbb{Z}$ with $n < \eta(\beta)$. So,

$$\begin{aligned}
\nu_p(\alpha,n) &= \max\{\beta \mid \nu_p(\alpha,n) \geq \beta\} \\
&\geq \max\{\beta \mid n < \eta(\beta)\} \\
&= \max\{\beta \mid n < p^{\alpha-\beta+1}-1\} \\
&= \max\{\beta \mid \beta \leq \alpha - \lfloor \log_p(n+1) \rfloor\} \\
&= \alpha - \lfloor \log_p(n+1) \rfloor.
\end{aligned}$$

□

Lemma 6.3. *Let $\alpha, \beta \in \mathbb{Z}^+$. Then*

$$\sigma\big(\mathbb{Z}/\alpha\mathbb{Z}, \mathbb{Z}/\beta\mathbb{Z}\big) \;=\; \begin{cases} p^{\alpha-\beta+1}-2 & \text{if } \beta \leq \alpha, \\ -\infty & \text{otherwise.} \end{cases}$$

Proof. If $\beta > \alpha$, then every constant nonzero function has nonvanishing integral, and $\sigma\big(\mathbb{Z}/\alpha\mathbb{Z}, \mathbb{Z}/\beta\mathbb{Z}\big) = -\infty$ follows. So, we may assume $\beta \leq \alpha$. Then, as $n < \eta_p(\alpha,\beta)$ implies $\nu_p(\alpha,n) \geq \beta$, Theorem 2.1b) and Lemma 2.3 tell us that

$$\sigma\big(\mathbb{Z}/\alpha\mathbb{Z}, \mathbb{Z}/\beta\mathbb{Z}\big) \;\geq\; \eta_p(\alpha,\beta) - 1 \;=\; p^{\alpha-\beta+1} - 2\,.$$

That also

$$\sigma\big(\mathbb{Z}/\alpha\mathbb{Z}, \mathbb{Z}/\beta\mathbb{Z}\big) < p^{\alpha-\beta+1}-1$$

follows from $\nu_p(\alpha, p^{\alpha-\beta+1}-1) = \beta - 1 < \beta$. So, $\sigma\big(\mathbb{Z}/\alpha\mathbb{Z}, \mathbb{Z}/\beta\mathbb{Z}\big) = p^{\alpha-\beta+1}-2$, indeed. □

References

[AM21] E. Aichinger and J. Moosbauer, *Chevalley-Warning type results on abelian groups.* J. Algebra 569 (2021), 30–66.

[Ax64] J. Ax, *Zeroes of polynomials over finite fields.* Amer. J. Math. 86 (1964), 255–261.

[Ch35] C. Chevalley, *Démonstration d'une hypothèse de M. Artin.* Abh. Math. Sem. Univ. Hamburg 11 (1935), 73–75.

[CS21] P.L. Clark and U. Schauz, *Functional Degrees and Arithmetic Applications I: The Set of Functional Degrees.* To appear, J. of Algebra. `https://arxiv.org/abs/2201.02763`

[CS23a] P.L. Clark and U. Schauz, *Functional Degrees and Arithmetic Applications II: the group-theoretic prime Ax-Katz theorem.* `http://alpha.math.uga.edu/~pete/Clark-Schauz_Part2.pdf`

[Gr22] D. Grynkiewicz, *A generalization of the Chevalley-Warning and Ax-Katz theorems with a view towards combinatorial number theory.* `https://arxiv.org/abs/2208.12895`

[Ka71] N.M. Katz, *On a theorem of Ax.* Amer. J. Math. 93 (1971), 485–499.

[Ka09] D.J. Katz, *Point count divisibility for algebraic sets over $\mathbb{Z}/p^\ell\mathbb{Z}$ and other finite principal rings.* Proc. Amer. Math. Soc. 137 (2009), 4065–4075.

[Ka12] D.J. Katz, *On theorems of Delsarte-McEliece and Chevalley-Warning-Ax-Katz.* Des. Codes Cryptogr. 65 (2012), 291–324.

[KP12] R.N. Karasev and F.V. Petrov, *Partitions of nonzero elements of a finite field into pairs.* Israel J. Math. 192 (2012), 143–156.

[Ku52] E. Kummer, *Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen.* Journal für die reine und angewandte Mathematik. 1852 (44): 93–146.

[MM95] O. Moreno and C.J. Moreno, *Improvements of the Chevalley-Warning and the Ax-Katz theorems*. Amer. J. Math. 117 (1995), 241–244.

[MR75] M. Marshall and G. Ramage, *Zeros of polynomials over finite principal ideal rings*. Proc. Amer. Math. Soc. 49 (1975), 35–38.

[Sc14] U. Schauz, *Classification of polynomial mappings between commutative groups*. J. Number Theory 139 (2014), 1–28.

[Wa35] E. Warning, *Bemerkung zur vorstehenden Arbeit von Herrn Chevalley*. Abh. Math. Sem. Hamburg 11 (1935), 76–83.

[We77] C.S. Weisman, *Some congruences for binomial coefficients*. Michigan Math. J. 24 (1977), 141–151.

[Wi06] R.M. Wilson, *A lemma on polynomials modulo p^m and applications to coding theory*. Discrete Math. 306 (2006), 3154–3165.