Math. Proc. Camb. Phil. Soc. (2025), **179**, 409–438
doi:10.1017/S0305004125000398
First published online 30 May 2025

Functional degrees and arithmetic applications III: beyond prime exponent

BY PETE L. CLARK
Mathematics Department, University of Georgia Athens, GA 30606*, U.S.A.*
e-mail: plclark@gmail.com

AND UWE SCHAUZ
Department of Pure Mathematics, Xi'an Jiaotong-Liverpool University, Suzhou 215123*, Jiangsu Province, P. R. China.*
e-mail: uwe.schauz@xjtlu.edu.cn

(*Received* 14 *October* 2024*; accepted* 24 *March* 2025)

Abstract

Continuing our work on group-theoretic generalisations of the prime Ax–Katz Theorem, we give a lower bound on the p-adic divisibility of the cardinality of the set of simultaneous zeros $Z(f_1, f_2, \dots, f_r)$ of r maps $f_j : A \to B_j$ between arbitrary finite commutative groups A and B_j in terms of the invariant factors of $A, B_1, B_2, \cdots, B_r$ and the *functional degrees* of the maps $f_1, f_2, \dots, f_r$.

2020 Mathematics Subject Classification: 13M10 (Primary); 20K01, 20D60, 11G25 (Secondary)

1. *Introduction and Main Results*

1·1. *Notation and Terminology*

Throughout this paper, p is a fixed but arbitrary prime number. We denote by ord_p the *p-adic valuation* on $\mathbb{Q}$. For a an integer $q \geq 2$ and a nonzero integer Z, we denote by $\mathrm{ord}_q(Z)$ the largest power of q that divides Z, and we put $\mathrm{ord}_q(0) = \infty$. (When $q = p$ is prime, this is the p-adic valuation.) Also 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 we endow $\tilde{\mathbb{N}}$ with the total ordering that extends the usual ordering on $\mathbb{N}$ so that $-\infty$ is the smallest and ∞ is the largest element.

If $R, R_1, \dots, R_r$ are sets and $f_1 : R \to R_1, \dots, f_r : R \to R_r$ are functions (possibly given as polynomials), such that each of the sets R_j contains a distinguished element that is denoted as $0 = 0_{R_j}$, then we define

$$Z(f_1, \dots, f_r) \;=\; Z_R(f_1, \dots, f_r) := \big\{x \in R \mid f_1(x) = 0\,, \dots, f_r(x) = 0\big\}\,.$$

For arbitrary commutative groups A and B, we denote with B^A the set of all functions $f : A \to B$, and define for each $a \in A$ a *difference operator* $\Delta_a \in \mathrm{End}(B^A)$ by

© The Author(s), 2025. Published by Cambridge University Press on behalf of Cambridge Philosophical Society.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

$$\Delta_a f : x \longmapsto f(x+a) - f(x)\,.$$

Following Aichinger–Moosbauer each $f \in B^A$ has a *functional degree*

$$\mathrm{fdeg}(f) := \sup\left\{n \in \mathbb{N} \mid \exists a_1 \in A, \ldots, \exists a_n \in A\colon \Delta_{a_1} \ldots \Delta_{a_n} f \neq 0\right\} \in \widetilde{\mathbb{N}}\,,$$

where[1] $\sup(\emptyset) := -\infty$. This degree may be infinite but we focus mainly on the subset[2]

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

1·2. *The Story so far*

This paper is a direct continuation of our prior works [4, 5]; in these papers as well as in the present paper, our goal is to synthesize, further develop and apply work of Wilson [19] and Aichinger–Moosbauer [1].

In [1], Aichinger and Moosbauer develop a calculus of finite differences (discrete derivatives) for maps $f : A \to B$ between arbitrary commutative groups A and B, and introduce the functional degree $\mathrm{fdeg}(f)$ based on the simple idea that the functional degree should decrease if a discrete derivative is taken. See also the exposition in our prior work [4, section 2·3], where difference operators and other basics are introduced in more detail. One of the key insights of [1] is that it is often fruitful to view the elements of $\mathcal{F}(A,B)$ as the "polynomial functions from A to B", a point of view that was introduced in [15] already. If P is a polynomial expression in n variables with coefficients in a (not necessarily commutative) rng[3] R, and $E(P) : R^n \to R$ is the corresponding polynomial function, then, by [1, 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 [1, 10] and work of the present authors [5, proposition 2·19 and theorem 4·9], we know how to compute $\mathrm{fdeg}(E(P))$ from the family of coefficients of the monomials of P when R is a commutative integral domain. Equality holds without restrictions, for all polynomials P, if and only if R has characteristic 0. In general, however, the functional degree may even be limited by a constant. For commutative groups A and B, we put

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

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 in [4, theorem 4·9c], [16], and [15, theorem 3·9], using arithmetic results of Weisman [18] and Wilson [19]:

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

[2] $\mathcal{F}(A,B)$ is actually a $\mathbb{Z}[A]$-submodule of the $\mathbb{Z}[A]$-module B^A over the integral group ring $\mathbb{Z}[A]$ of A, where the scalar multiple of $f \in B^A$ by $c = \sum_{a \in A} n_a[a] \in \mathbb{Z}[A]$ is $cf : x \longmapsto \sum_{a \in A} n_a f(x+a)$.

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

THEOREM 1·1. *Let $N, \beta, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$, let $\underline{\alpha} := (\alpha_1, \ldots, \alpha_N)$, and let B be a finite commutative p-group of exponent p^β. Then*

$$\delta\Big(\bigoplus_{i=1}^{N} \mathbb{Z}/p^{\alpha_i}\mathbb{Z}, B\Big) \;=\; \delta_p(\underline{\alpha}, \beta),$$

where

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

For finite commutative p-groups A and B, the quantity $\delta(A, B)$ can be interpreted as the largest possible "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* $\overline{P} \in \mathbb{F}_q[t_1, \ldots, t_n]$ consisting of monomial terms $t_1^{a_1} \cdots t_n^{a_n}$ with $0 \le a_i \le q-1$ and such that $E(P) = E(\overline{P})$, 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 $q := p^N$. Let $f_1, \ldots, f_r \in \mathbb{F}_q[t_1, \ldots, t_n]$ be polynomials of positive degrees. If $\sum_{j=1}^{r} \deg(f_j) < n$, then*:

(*a*) $\operatorname{ord}_p\big(\#Z_{\mathbb{F}_q^n}(f_1, \ldots, f_r)\big) \ge 1$ *(Chevalley–Warning theorem* [**3**, **17**]*)*,

(*b*) $\operatorname{ord}_q\big(\#Z_{\mathbb{F}_q^n}(f_1, \ldots, f_r)\big) \ge \left\lceil \frac{n - \sum_{j=1}^{r} \deg(f_j)}{\max_{j=1}^{r} \deg(f_j)} \right\rceil$ *(Ax–Katz theorem* [**2**, **8**]*)*.

Indeed, Aichinger–Moosbauer used their functional calculus to prove the following result:

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 nonzero functions. If

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

then

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

Proof. This is [**1**, theorem 12·2] with zero functions of degree $-\infty$ excluded.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

If R is a finite rng of prime power order and $P_1, \dots, P_r$ are polynomial expressions over R in N variables, then applying Theorem 1·3 with $A = B = (R, +)$ and with $f_1 = E(P_1), \dots, f_r = E(P_r)$ the associated polynomial functions from R^N to R, one obtains a ring-theoretic generalisation of of Theorem 1·2 a. This gives a "psychological" explanation for the presence of N in Theorem 1·3, but nothing is lost by taking $N = 1$.

The same work [**1**] gave a group-theoretic generalisation of Warning's second theorem [**1**, theorem 14·2], but they left open the problem of applying their calculus to higher p-adic congruences. However, a 2006 work of R. Wilson [**19**] gave a strikingly new and elementary proof of Theorem 1·2(b) over the prime field $\mathbb{F}_p$ using, in particular, the difference operators $\Delta : f \mapsto \Delta f := \big(x \mapsto f(x+1) - f(x)\big)$ from the calculus of finite differences. 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 [**18**]. Moreover, with some further development of the Achinger–Moosbauer calculus – especially that for commutative group B, the elements of $\mathcal{F}(\mathbb{Z}^N, B)$ have series expansions (as recalled in Theorem 3·1) – we were able [**5**, cor. 1·9] to refine Wilson's argument to give the following group-theoretic generalisation 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 $f_1, \dots, f_r \in A^{A^n}$ be nonconstant functions. Then*:

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

We emphasise that like Theorem 1·3, Theorem 1·4 is a purely group-theoretic result. When $A = \mathbb{F}_p$ we recover Theorem 1·2 b 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 [**13**], which however does not imply the full Ax–Katz Theorem over $\mathbb{F}_q$ (cf. [**5**, remark 1·4]).

After seeing a related manuscript of Grynkiewicz [**7**], we noticed that the argument that proves Theorem 1·4 can be adapted to prove a more general result:

THEOREM 1·5 (**5**, theorem 1·7). *Let $N, r, \beta_1, \dots, \beta_r \in \mathbb{Z}^+$, 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*:

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

Let us compare Theorems 1·4 and 1·5. In the former result all the maps $f_1, \dots, f_r$ take values in a fixed finite commutative p-group A that is required to have exponent p but is not required to be cyclic. In the latter result the maps $f_1, \dots, f_r$ take values in varying cyclic p-groups $\mathbb{Z}/p^{\beta_j}\mathbb{Z}$ that are not necessarily of prime exponent. So it may seem that we have lost generality in the requirement that the target groups be cyclic. But there is a *cyclic exchange* phenomenon: as discussed in Section 2·1, in these results we may exchange a map f_j into a finite commutative p-group with K invariant factors for a K-tuple $(f_{j,1}, \dots, f_{j,K})$ of maps into finite cyclic p-groups compatibly with our setup. Thus considering maps with targets in *varying cyclic groups* carries all the content of the general case.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Moreover, there is a Sylow primary decomposition for maps of finite functional degree between arbitrary finite commutative groups [**4**, cor. 3·14c and cor. 3·15]. Using this, one can extend all of these group-theoretic results from finite commutative p-groups to arbitrary finite commutative groups. This was done already in [**5**, rem. 1·8 and cor. 1·9], and we repeat the discussion here in Section 2·2.

1·3. *The Main Theorem*

The main result of this paper is an Ax–Katz type lower bound on $\operatorname{ord}_p(\#Z(f_1,\dots,f_r))$ for maps $f_j : A \to B_j$ between arbitrary finite commutative groups $A, B_1, \dots, B_r$. We may first assume that all groups are p-groups (cf. Section 2·2). By "cyclic exchange" (cf. Section 2·1), we may also assume that each B_j is cyclic. So, in the core result below, we assume

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

and

$$B_1 = \mathbb{Z}/p^{\beta_1}\mathbb{Z}\,,\ \dots\ ,\ B_r = \mathbb{Z}/p^{\beta_r}\mathbb{Z}\,,$$

where $r, \beta_1, \dots, \beta_r, N, \alpha_1, \dots, \alpha_N \in \mathbb{Z}^+$. We also assume, without loss of generality, that each f_j is nonconstant of functional degree at most $d_j \in \mathbb{Z}^+$. In other words, for each $1 \le j \le r$ we have a function

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

We put

$$\mathcal{A} := \sum_{i=1}^{N} \frac{p^{\alpha_i}-1}{p-1} \quad \text{and} \quad \mathcal{B} := \sum_{j=1}^{r} d_j \frac{p^{\beta_j}-1}{p-1}\,,$$

and order the α_i and the f_j so that

$$\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_N \qquad \text{and } d_1 p^{\beta_1} \ge d_2 p^{\beta_2} \ge \cdots \ge d_r p^{\beta_r}.$$

To express our result we also need the *conjugates* $\alpha'_1, \alpha'_2, \dots, \alpha'_{\alpha_1}$ defined by

$$\alpha'_j := \#\big\{1 \le i \le N \mid \alpha_i \ge j\big\}\,,$$

which we discuss in more detail in Section 6. We set

$$\begin{aligned} \alpha &:= \alpha_1 + \alpha_2 + \cdots + \alpha_N \\ &= \alpha'_1 + \alpha'_2 + \cdots + \alpha'_{\alpha_1} \quad \text{(by Example 6·3)} \end{aligned}$$

and

$$\begin{aligned} \breve{\alpha} &:= \breve{\alpha}_1 + \breve{\alpha}_2 + \cdots + \breve{\alpha}_N \\ &= \alpha'_1 + \alpha'_2 + \cdots + \alpha'_{\breve{\alpha}_1} \quad \text{(by Example 6·4)}, \end{aligned}$$

where

$$\breve{\alpha}_i := \min(\alpha_i, L)$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

with

$$L := \beta_1 + \lfloor \log_p(d_1) \rfloor .$$

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

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

We further put

$$\begin{aligned} \breve{\mathcal{A}} &:= \sum_{i=1}^{N} \frac{p^{\breve{\alpha}_i} - 1}{p - 1} \\ &= \alpha'_1 p^0 + \cdots + \alpha'_{\breve{\alpha}_1} p^{\breve{\alpha}_1 - 1} \quad \text{(by Example 6·4)} \\ &= D_1 + \cdots + D_{\breve{\alpha}} \quad (\text{ as } \breve{\alpha} = \alpha'_1 + \alpha'_2 + \cdots + \alpha'_{\breve{\alpha}_1} \text{ by Example 6·4}). \end{aligned}$$

With these parameters and definitions we can prove in the subsequent sections and in particular in Section 3·4, the following main result of our paper, which is a simultaneous generalisation of Theorem 1·3 of Aichinger–Moosbauer and Theorem 1·4 (hence also of Theorem 1·5) of the present authors.

THEOREM 1·6 *With the parameters and settings above,*

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

Note that our lower bound is equal to 0 if $\mathcal{A} \leq \mathcal{B}$. This is because $\breve{\mathcal{A}} \leq \mathcal{A} \leq \mathcal{B}$ (so we are in the second case of Theorem 1·6), and by Example 6·3 also

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

In the case $\mathcal{A} > \mathcal{B}$, however, the lower bound in Theorem 1·6 is positive, since the constraint on the maximum in the second case is not satisfied for $t := \alpha$, as $D_1 + \cdots + D_\alpha = \mathcal{A} \not\leq \mathcal{B}$. We thus obtain the following corollary:

COROLLARY 1·7 *If $\mathcal{A} > \mathcal{B}$ then*

$$\operatorname{ord}_p\big(\#Z_A(f_1, \ldots, f_r)\big) \geq 1 .$$

As we explain in Section 2·1, Corollary 1·7 implies Theorem 1·3: essentially, it is the generalisation of Theorem 1·3 in which the maps $f_1, \ldots, f_r$ are allowed to take values in varying finite commutative p-groups $B_1, \ldots, B_r$.[4]

[4] Aichinger–Moosbauer's proof of Theorem 1·3 can be adapted to prove this generalisation.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Our main theorem takes a somewhat simpler form when $\alpha_1 = \cdots = \alpha_N$. If we set

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

then, as we prove in Section 3·5, we obtain the following corollary:

COROLLARY 1·8 *If* $\alpha_1 = \cdots = \alpha_N$ *then*

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

Theorem 1·5 is the case $\alpha_1 = \cdots = \alpha_N = 1$ of Corollary 1·8 (with $d_j = \operatorname{fdeg}(f_j)$ for $1 \leq j \leq r$). Indeed, in this case, $\breve{\alpha}_1 = 1$, and the lower bounds in Theorem 1·5 and Corollary 1·8 either coincide (whenever $N > \mathcal{B}$) or are both non-positive and vacuous (whenever $N \leq \mathcal{B}$).

1·4 *Pending Sections and Tasks*

In Section 2, we explain how our results can be applied in more general situations and to polynomials over rngs. First in Section 2·1. we look at groups of prime power order, not necessarily cyclic ones. In Section 2·2, we generalise them to commutative groups of finite order. In Section 2·3, polynomials over rngs of prime power order are discussed. In Section 2·4, this is generalised to polynomials over rngs of finite order.

In Sections 3·4 and 3·5, we present the proof of Theorem 1·6 and Corollary 1·8. The proof of Theorem 1·6 actually requires quite some preparation, and the more obvious part of that preparation is given in Sections 3·1 through 3·3.

There are also two less obvious optimisation tasks. The necessity to study these kind of optimisations becomes apparent only during the main proof, after the functions $\nu_p(\underline{\alpha}, \bullet)$ and $\mathcal{N}$ where introduced in Section 3·2 and 3·4. In other words, the main proof motivates and sets up those tasks. We moved those two optimisation tasks into the subsequent Sections 4 and 5, as they can be studied independently. Inside our main proof in Section 3·4, the results of those investigations are then just cited.

Section 4 is about the second lower bound of Theorem 1·6:

$$\alpha - \max\big\{1 \leq t \leq \alpha \mid D_1 + \cdots + D_t \leq \mathcal{B}\big\}\,.$$

It includes in Section 4·2, a supplementary discussion of alternative ways to express this lower bound, which is not needed to understand the main proof and the rest of the paper.

Section 5 deals with the case distinction, and the first lower bound of Theorem 1·6:

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

Section 6 provides some background on conjugate partitions, with a number of lemmas and examples that we conveniently cite throughout the paper. Readers not familiar with conjugate partitions and Ferrers diagrams may want to read through this section first.

2. *Extending the Scope*: *Finite Commutative Groups and Polynomials*

In this section we explain how our results can be applied in more general situations.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

2·1. *Commutative Groups of Prime Power Order*

As mentioned before, our results can still be applied when arbitrary finite commutative p-groups $B_j = \bigoplus_{i=1}^{K_j} (\mathbb{Z}/p^{\beta_{j,i}}\mathbb{Z})$ replace the cyclic p-groups $\mathbb{Z}/p^{\beta_j}\mathbb{Z}$ as codomains. One just has to use the coordinate projections $\pi_k : \bigoplus_{i=1}^{K_j} (\mathbb{Z}/p^{\beta_{j,i}}\mathbb{Z}) \to \mathbb{Z}/p^{\beta_{j,k}}\mathbb{Z}$ first (as already explained in [5, Rem. 1·8]), to define the functions

$$f_{j,k} := \pi_k \circ f_j \quad \text{with} \quad \operatorname{fdeg}(f_{j,k}) \le \operatorname{fdeg}(f_j) \le d_j \,.$$

Then our results can be applied to those $f_{j,k}$ and the degree restrictions $\operatorname{fdeg}(f_{j,k}) \le d_j$, with the outcome interpreted in terms of the f_j.

If, for instance, Corollary 1·4 is applied in this way, then the parameter $\mathcal{B}$ takes the form $\mathcal{B} = \sum_{j=1}^{r} \big(d_j \sum_{k=1}^{K_j} ((p^{\beta_{j,k}} - 1)/(p-1))\big)$, while the parameter $\mathcal{A}$ and the conclusion remain unchanged: still $\mathcal{A} > \mathcal{B}$ implies $\operatorname{ord}_p(\#Z_A(f_1, \ldots, f_r)) \ge 1$. This generalises Theorem 1·3, where $B_1 = \cdots = B_r = \bigoplus_{i=1}^{K} (\mathbb{Z}/p^{\beta_i}\mathbb{Z})$, i.e., $\mathcal{B} = \sum_{j=1}^{r} d_j \sum_{k=1}^{K} ((p^{\beta_k} - 1)/(p-1))$, and $A = \big(\bigoplus_{i=1}^{m} \mathbb{Z}/p^{\alpha_i}\mathbb{Z}\big)^N$, i.e., $\mathcal{A} = N \sum_{i=1}^{m} ((p^{\alpha_i} - 1)/(p-1))$.

2·2. *Commutative Groups of Finite Order*

Let $A, B_1, \ldots, B_r$ be any nontrivial finite commutative groups. We write out the primes dividing $\#\big(A \times \prod_{i=1}^{r} B_i\big)$ as $\ell_1 < \cdots < \ell_s$, and set

$$A[\ell_j^\infty] := \big\{ x \in A \mid \ell_j^k x = 0 \text{ for some } k \in \mathbb{Z}^+ \big\} \,.$$

For each fixed $1 \le j \le r$, we have a canonical $\mathbb{Z}$-module injection

$$\prod_{h=1}^{s} B_j[\ell_h^\infty]^{A[\ell_h^\infty]} \to B_j^A$$

in which we send each vector $(g_{j,1}, \ldots, g_{j,s})$ of functions $g_{j,h} : A[\ell_h^\infty] \to B_j[\ell_h^\infty]$ to the identically named function

$$\begin{aligned} (g_{j,1}, \ldots, g_{j,s}) : A = \prod_{h=1}^{s} A[\ell_h^\infty] &\to B_j = \prod_{h=1}^{s} B_j[\ell_h^\infty] \\ (x_1, \ldots, x_s) &\mapsto \big(g_{j,1}(x_1), \ldots, g_{j,s}(x_s)\big) \,. \end{aligned}$$

By [4, cor. 3·14c and cor. 3·15], upon restriction to functions of finite functional degree, this yields a canonical isomorphism. Via identification, we have

$$\mathcal{F}(A, B_j) = \prod_{h=1}^{s} \mathcal{F}(A[\ell_h^\infty], B_j[\ell_h^\infty])$$

in which moreover $\operatorname{fdeg}((g_{j,1}, \ldots, g_{j,s})) = \max\{\operatorname{fdeg}(g_{j,h}) \mid 1 \le h \le s\}$. In other words, a map $f_j : A \to B_j$ of finite functional degree is determined by its restrictions $g_{j,h} := f_j|_{A[\ell_h^\infty]} \in B_j[\ell_h^\infty]^{A[\ell_h^\infty]}$, and we may write

$$f_j = (g_{j,1}, \ldots, g_{j,s}) \quad \text{and} \quad \operatorname{fdeg}(f_j) = \max\big\{\operatorname{fdeg}(g_{j,h}) \mid 1 \le h \le s\big\} \,.$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

If we now consider r maps $f_j : A \to B_j$, where $1 \le j \le r$, then we get $s \times r$ primary component maps $g_{j,h} : A[\ell_h^\infty] \to B_j[\ell_h^\infty]$, and it is immediate that

$$\#Z(f_1, \ldots, f_r) = \prod_{h=1}^{s} \#Z(g_{1,h}, \ldots, g_{r,h}) .$$

So, with the previous Section 2·1, we obtain for each $1 \le h \le s$ a lower bound on $\mathrm{ord}_{\ell_h}\big(\#Z(f_1, \ldots, f_r)\big)$ in terms of $A, B_1, \ldots, B_r$ and $\mathrm{fdeg}(f_1), \ldots, \mathrm{fdeg}(f_r)$.

2·3. *Polynomials over Rngs of Prime Power Order*

Before [5], 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 [2, 8, 10, 14]. Now let R be a finite rng with order a power of p, so there are $N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$ 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 $\deg(P_j) \le d_j$ for all $1 \le j \le r$. Then Theorem 1·6 with $A := A_1^n$ and the previous Section 2·1 apply to give an Ax–Katz type lower bound on $\mathrm{ord}_p(\#Z(P_1, \ldots, P_r))$. In particular, as $\breve{\mathcal{A}} \ge n$, one sees the following *asymptotic Ax–Katz over a finite rng*: if r and $d_1, \ldots, d_r$ remain fixed, then $\mathrm{ord}_p(\#Z(P_1, \ldots, P_r))$ approaches infinity with n.

This asymptotic Ax–Katz result is also established in a concurrent work by the first author and N. Triantafillou [6, theorem 6·2]. The proof given there uses a new invariant: for nontrivial, finite commutative p-groups A and B the *summation invariant* $\sigma(A, B)$ is the largest $d \in \tilde{\mathbb{N}}$ such that $\int_A f = 0$ for all maps $f : A \to B$ with functional degree at most d. The invariant $\sigma(A, B)$ does not appear explicitly in the work of Aichinger–Moosbauer, but neverthless they give what amounts to a computation of $\sigma(A, B)$ when A and B are p-groups and B has exponent p [1, lemma 12·1], and this is a key ingredient of their proof of Theorem 1·3. In [6] lower bounds are given on $\sigma(A, B)$ in the general case (and exact computations are given in some further special cases). If one takes these results as a "black box," then the proof of asymptotic Ax–Katz over a finite rng given in [6] is much simpler than the proof of our main result. However the lower bound on $\mathrm{ord}_p(\#Z(P_1, \ldots, P_r))$ given by our Theorem 1·6 is better than (or equal to, in certain special cases) the corresponding bound given by the methods of [6]. Thus neither work majorizes the other.

Our present approach implicitly uses a "lifted variant" of $\sigma(A, B)$. A comparison of this lifted variant with $\sigma(A, B)$ is made in [6, 7].

2·4. *Polynomials over Rngs of Finite Order*

The two previous sections can be combined to address the case of polynomial expressions in n variables of degrees $d_1, \ldots, d_r$ over any nontrivial finite rng R. In this case, the asymptotic Ax–Katz phenomenon can be expressed as follows: keeping the number and degrees of the polynomial expressions $f_1, \ldots, f_r$ fixed, we find that $\mathrm{ord}_{\#R}(Z(f_1, \ldots, f_r))$ approaches infinity with n. A slightly different, but equivalent, formulation is given in [6, theorem 6·2].

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

3. *Reduction to Discrete Optimization*

In this section we prove Theorem 1·6 (and afterwards Corollary 1·8) based on some technical results about the minimum values of certain discrete functions. In other words, this main part of the proof reduces us to some discrete optimisation problems. These discrete optimisation problems are stated and solved completely independent from the original problem in later sections, but are cited and used here. We first recall some basics from our earlier work, then introduce some basic number theoretic results, and then start that reductionistic proof.

3·1. *Some Recalled Results*

We provide some basics about series expansions in terms of binomial polynomials $\binom{t}{d} := t(t-1)\dots(t-d+1)/d! \in \mathbb{Q}[t]$, which are integer valued, i.e., $\binom{x}{d} \in \mathbb{Z}$ whenever $x \in \mathbb{Z}$. Again, we write $\underline{n}$ for $(n_1, \dots, n_N)$, and we set $|\underline{n}| := n_1 + \cdots + n_N$.

THEOREM 3·1. *Let B be a commutative group, and let $f \in B^{\mathbb{Z}^N}$.*

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

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

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

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

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

Proof. This is [**5**, theorem 2·8], where it is also mentioned that the sum in the first part is always well-defined, because at each fixed point $\underline{x}$ in $\mathbb{N}^N$ (unlike in $\mathbb{Z}^N \setminus \mathbb{N}^N$) the sum has always only a finite number of nonzero summands, for whatever c.

We now recall some terminology and results concerning proper lifts. Let $\mu : B \to B'$ be a surjective homomorphism of commutative groups, and let $f \in \mathcal{F}(\mathbb{Z}^N, B')$. To define a *proper lift* $\tilde{f} \in \mathcal{F}(\mathbb{Z}^N, B)$ of f (it is not unique), observe first that there exists (by Theorem 3·1 (*b*)) a unique (by Theorem 3·1 (*a*)) function $c : \mathbb{N}^N \to B'$ such that

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

Then a *proper lift* of c to B is a function $\tilde{c} : \mathbb{N}^N \to B$ such that

$$\mu \circ \tilde{c} = c \quad \text{and} \quad \tilde{c}(\underline{n}) = 0 \iff c(\underline{n}) = 0 \quad \text{for all } \underline{n} \in \mathbb{N}^N.$$

Such a proper lift always exists (non-uniquely), and we can use it to define a *proper lift* of f by

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

This sum is actually well defined, because $\tilde{c}:\mathbb{N}^N\to B$ is finitely nonzero (i.e., its support is finite), since our $c:\mathbb{N}^N\to B'$ is finitely nonzero (as in Theorem 3·1 (*b*)). For every proper lift $\tilde{f}$ of f we have

$$f = \mu\circ\tilde{f} \quad\text{and}\quad \mathrm{fdeg}(\tilde{f}) = \mathrm{fdeg}(f)\,.$$

Inside our main proof we work with series expansions of proper lifts of pullbacks, and we need the following corollary to Theorem 3·1, in which the maximal finite degrees

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

of Theorem 1·1 play an important role:

COROLLARY 3·2. *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*)

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

(*b*) *For all $h\in\mathbb{Z}^+$ and all $\underline{n}\in\mathbb{N}^N$ with $|\underline{n}|>\delta_p(\underline{\alpha},h)$,*

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

Proof. This is [**5**, cor. 2·25].

3·2. *The numbers $v_p(\underline{\alpha},\underline{n})$ and $\mathcal{V}_p(\underline{\alpha},D)$*

For $\alpha\in\mathbb{Z}^+$ and $n\in\mathbb{N}$, we put

$$v_p(\alpha,n) := \mathrm{ord}_p\left(\sum_{x=0}^{p^\alpha-1}\binom{x}{n}\right).$$

LEMMA 3·3. *For each $\alpha\in\mathbb{Z}^+$ and $n\in\mathbb{N}$,*

$$v_p(\alpha,n) = \begin{cases}\alpha-\mathrm{ord}_p(n+1) & \textit{if } n\le p^\alpha-1,\\ \infty & \textit{otherwise.}\end{cases}$$

Proof. The case $n=0$ is handled by Proposition 3·4 (*a*), while if $n\ge p^\alpha$ then $\sum_{x=0}^{p^\alpha-1}\binom{x}{n}=0$, so $v_p(\alpha,n)=\infty$. So we may assume that $1\le n\le p^\alpha-1$. Using Pascal's rule $\binom{a}{b}=\binom{a-1}{b}+\binom{a-1}{b-1}$ we see that

$$\sum_{x=0}^{p^\alpha-1}\binom{x}{n} = \binom{n+1}{n+1}+\binom{n+1}{n}+\binom{n+2}{n}+\binom{n+3}{n}+\cdots+\binom{p^\alpha-1}{n}$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

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

Now we can apply Kummer's insight [12] 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 $1000\ldots0$ with α zeros, this number of carries is $\alpha - \mathrm{ord}_p(n+1)$ 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

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

where

$$
[p^{\underline{\alpha}}) := [p^{\alpha_1}) \times [p^{\alpha_2}) \times \cdots \times [p^{\alpha_N}) \quad \text{with} \quad [p^{\alpha_i}) := \{0, 1, \ldots, p^{\alpha_i} - 1\}\,.
$$

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

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

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

PROPOSITION 3·4. *Let $\alpha \in \mathbb{Z}^+$, $\underline{\alpha} \in (\mathbb{Z}^+)^N$, and $D \in \mathbb{N} \cup \{\infty\}$.*

(a)

$$
\nu_p(\alpha, 0) = \alpha\,,
$$

and thus

$$
\nu_p(\underline{\alpha}, \underline{0}) = |\underline{\alpha}|
$$

and

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

(b)

$$
\nu_p(\alpha, p^\alpha - 1) = 0\,,
$$

and thus

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

and

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

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

As we already determined $v_p(\underline{\alpha}, \underline{n}) = \sum_{i=1}^{N} v_p(\alpha_i, n_i)$ in the lemma above, the precise calculation of $\mathcal{V}_p(\underline{\alpha}, D)$ becomes a discrete optimisation problem. We will solve that problem in the next section, in Theorem 4·1. With the definitions of the numbers $\alpha := \alpha_1 + \cdots + \alpha_N$ and $D_1, \ldots, D_\alpha$ used in Theorem 1·6, the result can be stated as follows:

$$\mathcal{V}_p(\underline{\alpha}, D) \;=\; \alpha - \max\left\{0 \leq t \leq \alpha \mid D_1 + \cdots + D_t \leq \tfrac{D}{p-1}\right\}. \tag{3·1}$$

3·3. *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 [**11**], we set

$$\int_S f := \sum_{x \in S} f(x) \;\in\; B \quad \text{and} \quad \int f := \int_A f\,.$$

Here we are mostly interested in the case $A = \mathbb{Z}^N$, $B = \mathbb{Z}$, and $S = [p^{\underline{\alpha}})$. The following results generalise work of Wilson [**19**, lemma 4] about the case $\alpha_1 = \cdots = \alpha_N = 1$:

PROPOSITION 3·5. *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*

$$\mathrm{ord}_p\left(\int_{[p^{\underline{\alpha}})} f\right) \;\geq\; \mathcal{V}_p\big(\underline{\alpha}, D\big)\,.$$

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 3·1, 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:

$$\mathrm{ord}_p\left(\int_{[p^{\underline{\alpha}})} \binom{x_1}{n_1} \cdots \binom{x_N}{n_N}\right) = \prod_{j=1}^{N} \sum_{x_j \in [p^{\alpha_j})} \binom{x_j}{n_j} = v_p\big(\underline{\alpha}, \underline{n}\big) \;\geq\; \mathcal{V}_p\big(\underline{\alpha}, |\underline{n}|\big) \;\geq\; \mathcal{V}_p\big(\underline{\alpha}, D\big)\,.$$

3·4. *The Proof of Theorem 1·6*

Below is the proof of Theorem 1·6, modulo two main discrete optimisation tasks. On one side, our proof 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: the two optimisation tasks that are needed to complete the argument.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Our proof uses previously made definitions and the entire setup of Theorem 1·6 without reintroducing them. Some new definitions are made on the way, as well. As in [**1**, section 6], we also use the tensor product of functions: if $A_1, \dots, A_n$ are commutative groups and R is a rng, then the *tensor product* $\bigotimes_{i=1}^n h_i$ of maps $h_i : A_i \to R$ is the map

$$\bigotimes_{i=1}^n h_i \colon \bigoplus_{i=1}^n A_i \to R\,, \quad (x_1, x_2, \dots, x_n) \mapsto h_1(x_1)h_2(x_2)\dots h_n(x_n)\,.$$

Proof of Theorem 1·6. Let $\beta \in \mathbb{Z}^+$ be fixed given. For each $1 \le j \le r$ define the map $\chi_j : \mathbb{Z} \to \mathbb{Z}/p^\beta\mathbb{Z}$ 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 $\mathbb{Z}$. Using that $\mathbb{Z}/p^\beta\mathbb{Z}$ and $\mathbb{Z}$ are rngs, not just additive groups, set

$$\chi := \bigotimes_{j=1}^r \chi_j \quad \text{and} \quad \widetilde{\chi} := \bigotimes_{j=1}^r \widetilde{\chi}_j\,.$$

If q denotes the quotient map from $\mathbb{Z}^N$ to $A := \bigoplus_{i=1}^N \mathbb{Z}/p^{\alpha_i}\mathbb{Z}$, and $\tilde{F}_j : \mathbb{Z}^N \to \mathbb{Z}$ is a proper lift of the pullback $F_j : \mathbb{Z}^N \to \mathbb{Z}/p^{\beta_j}\mathbb{Z}$ of the function $f_j : A \to \mathbb{Z}/p^{\beta_j}\mathbb{Z}$, then for each $\underline{x} \in [p^{\underline{\alpha}}) := \prod_{i=1}^N \{0, 1, \dots, p^{\alpha_i} - 1\}$,

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

Moreover, when restricted to $[p^{\underline{\alpha}})$, the quotient map q induces a bijection from $[p^{\underline{\alpha}})$ to A. Hence, with the function $\tilde{\chi}(\tilde{F}_1, \dots, \tilde{F}_r) : \underline{x} \mapsto \tilde{\chi}(\tilde{F}_1(\underline{x}), \dots, \tilde{F}_r(\underline{x}))$, we get

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

We may certainly assume that $Z(f_1, \dots, f_r)$ is nonempty, so that $\operatorname{ord}_p\big(\#Z(f_1, \dots, f_r)\big)$ is finite. Hence, after increasing our $\beta \in \mathbb{Z}^+$ if necessary, we may assume

$$\beta > \operatorname{ord}_p\big(\#Z(f_1, \dots, f_r)\big) \in \mathbb{N}\,, \tag{3·2}$$

and with that

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

Now, for each $1 \le j \le r$, Corollary 3·2 provides an integer valued function c_j on the set

$$[\hat{n}_j(\beta)] := \{0, 1, \dots, \hat{n}_j(\beta)\}, \quad \text{where} \quad \hat{n}_j(\beta) := (p^{\beta_j} - 1) + (\beta - 1)p^{\beta_j - 1}(p - 1)\,,$$

such that, for each $x \in \mathbb{Z}$,

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

With

$$\underline{\hat{n}}(\beta) := (\hat{n}_1(\beta), \dots, \hat{n}_r(\beta)) \quad \text{and} \quad [\underline{\hat{n}}(\beta)] := \prod_{j=1}^{r} [\hat{n}_j(\beta)] \subseteq \mathbb{N}^r,$$

for each $\underline{x} \in \mathbb{Z}^N$,

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

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

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

So if we put

$$\mathbf{m} := \min_{\underline{n} \in [\underline{\hat{n}}(\beta)]} \left(\operatorname{ord}_p\big(c_1(n_1)\big) + \dots + \operatorname{ord}_p\big(c_r(n_r)\big) + \operatorname{ord}_p\left(\int_{[p^\alpha)} \binom{\tilde{F}_1}{n_1} \dots \binom{\tilde{F}_r}{n_r} \right) \right),$$

it follows that

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

Thus the matter of it is to give a good lower bound on the quantity $\mathbf{m}$, using that $\operatorname{fdeg}(\tilde{F}_j) = \operatorname{fdeg}(f_j) \leq d_j$ for all $1 \leq j \leq r$ (cf. [**5**, cor. 2·13, section 2·4 and section 2·5]). Part of this can be quickly done in the same way as in [**5**]: Corollary 3·2 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 $n \in [\hat{n}_j(\beta)]$,

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

Taking

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

we have

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

and thus Corollary 3·2 yields

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

where, for real numbers h,

$$\overline{h} := \max(h, 0)\,.$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Moreover, using [**1**, theorem 4·3 and lem. 6·1], we have

$$\operatorname{fdeg}\left(\binom{\tilde{F}_1}{n_1}\dots\binom{\tilde{F}_r}{n_r}\right) \le \sum_{j=1}^{r} d_j n_j\,,$$

and Proposition 3·5 shows that

$$\operatorname{ord}_p\left(\int_{[p^{\underline{\alpha}})}\binom{\tilde{F}_1}{n_1}\dots\binom{\tilde{F}_r}{n_r}\right) \ge \mathcal{V}_p\big(\underline{\alpha}, \textstyle\sum_{j=1}^{r} d_j n_j\big)\,.$$

We deduce that

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

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 the minimum of $\mathcal{N}(\underline{n})$ as $\underline{n}$ runs through $[\hat{\underline{n}}(\beta)]$ is performed through discrete optimisation in Section 5, with the outcome stated in Lemma 5·2. After increasing β if necessary[5], it yields

$$\min_{\underline{n}\in[\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n}) = \begin{cases} \left\lceil \dfrac{\breve{\mathcal{A}} - \mathcal{B}}{d_1 p^{\beta_1-1}} \right\rceil + \alpha - \breve{\alpha} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \alpha - \max\big\{1 \le t \le \alpha \mid D_1 + \dots + D_t \le \mathcal{B}\big\} & \text{if } \breve{\mathcal{A}} \le \mathcal{B}. \end{cases}$$

Notice that the answer obtained is independent of β.

3·5. *The Proof of Corollary* 1·8

As claimed before, the result takes a somewhat simpler form when $\alpha_1 = \dots = \alpha_N$.

Proof of Corollary 1.8. If $\alpha_1 = \dots = \alpha_N$ then

$$\alpha = N\alpha_1\,, \quad \breve{\alpha} = N\breve{\alpha}_1\,, \quad \alpha_1' = \dots = \alpha_{\alpha_1}' = N \quad \text{and} \quad \breve{\mathcal{A}} = N\frac{p^{\breve{\alpha}_1}-1}{p-1}\,,$$

which yields the claimed simplifications in the case $\breve{\mathcal{A}} > \mathcal{B}$. 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 also have

$$Q \le \log_p\big((p-1)\mathcal{B}/N + 1\big) < Q+1$$

[5] We need $\beta > s_0 := \overline{\lceil(\breve{\mathcal{A}} - \mathcal{B})/(d_1 p^{\beta_1-1})\rceil}$ in Lemma 5·2. Within the full proof of Theorem 1·6, however, we assume $\beta > \operatorname{ord}_p(\#Z)$ already in (3·2), and $\operatorname{ord}_p(\#Z) \ge s_0$ by the findings of this paper.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

i.e.,

$$N\frac{p^Q-1}{p-1} \le \mathcal{B} < N\frac{p^{Q+1}-1}{p-1} = N\frac{p^Q-1}{p-1} + Np^Q$$

and

$$R \le \frac{\mathcal{B} - N\frac{p^Q-1}{p-1}}{p^Q} < R+1$$

i.e.,

$$N\frac{p^Q-1}{p-1} + Rp^Q \le \mathcal{B} < N\frac{p^Q-1}{p-1} + (R+1)p^Q\,. \tag{3·3}$$

In particular, $N((p^Q-1)/(p-1)) \le \mathcal{B}$ and $N((p^Q-1)/(p-1)) + Rp^Q \le \mathcal{B} < N((p^Q-1)/(p-1)) + Np^Q$, so that, on one hand,

$$0 \le R < N\,.$$

If we further assume $D_1 + \cdots + D_\alpha > \mathcal{B}$, then $N((p^{\alpha_1}-1)/(p-1)) = D_1 + \cdots + D_\alpha > \mathcal{B} \ge N((p^Q-1)/(p-1))$, so that, on the other hand,

$$0 \le Q < \alpha_1\,.$$

From this follows $QN + R < N\alpha_1 = \alpha$, and we see that $D_{QN+1} = \cdots = D_{QN+R+1} = p^Q$, i.e., $D_1 + \cdots + D_{QN+R} = N((p^Q-1)/(p-1)) + Rp^Q$ and $D_1 + \cdots + D_{QN+R+1} = N((p^Q-1)/(p-1)) + (R+1)p^Q$. Hence, (3·3) can be restated as

$$D_1 + \cdots + D_{QN+R} \le \mathcal{B} < D_1 + \cdots + D_{QN+R+1}\,,$$

which means that

$$\max\{1 \le t \le \alpha \mid D_1 + \cdots + D_t \le \mathcal{B}\} = NQ + R\,.$$

Thus, we can replace the second lower bound $\alpha - \max\{1 \le t \le \alpha \mid D_1 + \cdots + D_t \le \mathcal{B}\}$ with $N(\alpha_1 - Q) - R$. This replacement is also correct in the case $D_1 + \cdots + D_\alpha \le \mathcal{B}$, because then both terms are non-positive. This is clear for $\alpha - \max\{1 \le t \le \alpha \mid D_1 + \cdots + D_t \le \mathcal{B}\}$, but we also see that $N((p^{\alpha_1}-1)/(p-1)) = D_1 + \cdots + D_\alpha \le \mathcal{B} < N((p^{Q+1}-1)/(p-1))$ implies $\alpha_1 < Q+1$, which entails $N(\alpha_1 - Q) - R \le 0$. We obtain

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

4. *Minimisation of* $v_p(\underline{\alpha}, \bullet)$

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

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

where the numbers $N, \alpha_1, \ldots, \alpha_N \in \mathbb{Z}^+$ with $\alpha_1 \ge \cdots \ge \alpha_N$ and $D \in \mathbb{N}$ are fixed given. (The case $D = \infty$ is trivial, as $\mathcal{V}_p(\underline{\alpha}, \infty) = 0$ by Proposition 3·4 (*b*).) In our investigation,

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

the original definition of $v_p(\underline{\alpha}, \bullet)$ does not actually matter. We may view the formula in Lemma 3·3 as the definition. More precisely, for $\underline{n} \in \mathbb{N}^N$,

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

Our final result will be stated in terms of the parameters $\alpha := \alpha_1 + \cdots + \alpha_N$, and $D_1, \ldots, D_\alpha$ of Theorem 1·6, i.e.,

$$\big(D_1, D_2, \ldots, D_\alpha\big) := \big(\underbrace{1, 1, \ldots, 1}_{\alpha_1' \text{ times}}, \underbrace{p, p, \ldots, p}_{\alpha_2' \text{ times}}, \ \ldots, \ \underbrace{p^{\alpha_1-1}, p^{\alpha_1-1}, \ldots, p^{\alpha_1-1}}_{\alpha_{\alpha_1}' \text{ times}} \big) .$$

4·1. *The Minimum Value $\mathcal{V}_p(\underline{\alpha}, D)$ of $v_p(\underline{\alpha}, \bullet)$ over $\mathcal{D}(N, D)$*

THEOREM 4·1. *In the setting above, with $D \in \mathbb{N}$, the function*

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

has minimum value

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

Proof. We may restrict the domain of $v_p(\underline{\alpha}, \bullet)$ from $\mathcal{D}(N, D)$ to $\mathcal{D}(N, D) \cap [p^{\underline{\alpha}})$ with $[p^{\underline{\alpha}}) := \prod_{i=1}^{N} \{0, 1, \ldots, p^{\alpha_i} - 1\}$, because $v_p(\underline{\alpha}, \bullet)$ is finite inside but positive infinite outside of $[p^{\underline{\alpha}})$. Inside $[p^{\underline{\alpha}})$, however,

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

So, we need to find the maximum of the function

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

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

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

where $[\underline{\alpha}] := \prod_{i=1}^{N} \{0, 1, \ldots, \alpha_i\}$ and

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

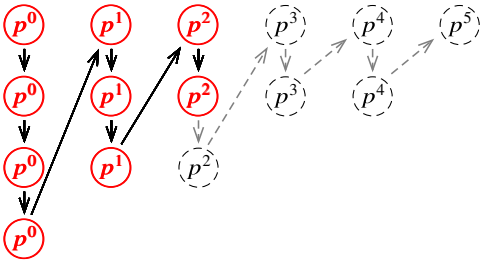


Fig. 1. The minimum weight of 9 dots inside $\underline{\alpha} = (6, 5, 3, 1)$ is $D_1 + D_2 + \cdots + D_9 = 4 + 3p + 2p^2$.

Now, if we draw Ferrers-type diagrams for the potential arguments $\underline{\mu}$ of λ as sub-diagrams of Ferrers' diagram of $\underline{\alpha}$ (representing μ_i by μ_i consecutive dots in row i) then $\lambda(\underline{\mu})$ is the number of dots in the sub-diagrams of $\underline{\mu}$, while $\omega(\underline{\mu})$ gives a weighted count of those dots – a dot in the jth column is counted with weight p^{j-1}, as shown in Figure 1. Hence, to find the maximum of λ, we need to maximize the number of dots in the sub-diagram corresponding to $\underline{\mu}$, while keeping their total weight (total cost) $\omega(\underline{\mu})$ below $D/(p-1)$. So, when selecting the dots in $\underline{\mu}$, we just have to select the cheapest dots first. In our situation of column-wise increasing weights, we have to select the dots column by column, from left to right, starting with the left-most column with dots 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 1. 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 $\underline{\mu}$ with

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

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

$$t \;=\; t(D) := \max\left\{0 \le t \le \alpha \mid D_1 + \cdots + D_t \le \tfrac{D}{p-1}\right\}.$$

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

$$\mathcal{V}_p(\underline{\alpha}, D) \;=\; v_p(\underline{\alpha}, \underline{n}(D)) \;=\; \alpha - \lambda(\underline{\mu}) \;=\; \alpha - \max\left\{0 \le t \le \alpha \mid D_1 + \cdots + D_t \le D/(p-1)\right\}.$$

Using our new formula, we can now show that $D < \sum_{i=1}^N ((p^{\alpha_i} - 1)/(p-1))$ is not just necessary for $\mathcal{V}_p(\underline{\alpha}, D) > 0$, as we already have seen in Proposition 3·4 (b), it is also sufficient:

COROLLARY 4·2. *Maintain the setup of Theorem 4·1, we have*

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Proof. With the last result in Example 6·3, we see that

$$D_1 + \cdots + D_\alpha = \alpha'_1 p^0 + \cdots + \alpha'_N p^{N-1} = \sum_{i=1}^{N} \frac{p^{\alpha_i}-1}{p-1}\,.$$

So, by Theorem 4·1,

$$\begin{aligned} \mathcal{V}_p(\underline{\alpha}, D) > 0 &\iff \max\left\{0 \le t \le \alpha \mid D_1 + \cdots + D_t \le \tfrac{D}{p-1}\right\} < \alpha \\ &\iff D_1 + \cdots + D_\alpha \not\le \tfrac{D}{p-1} \\ &\iff D < \sum_{i=1}^{N} \left(p^{\alpha_i} - 1\right). \end{aligned}$$

4·2. *Alternative Expressions for $\mathcal{V}_p(\underline{\alpha}, D)$ and Special Cases*

In our main theorem (Theorem 1·6), we presented only one formula as the final result, as we did not want to make things any more complicated than necessary. That result can, however, be stated in different forms, by replacing the second lower bound

$$\alpha - \max\left\{0 \le t \le \alpha \mid D_1 + \cdots + D_t \le \tfrac{D}{p-1}\right\} = \mathcal{V}_p(\underline{\alpha}, D)$$

with an alternative expression for $\mathcal{V}_p(\underline{\alpha}, D)$. We present several alternative formulas in the second remark below (and in the subsequent corollary), after extracting an additional insights from our previous calculation of $\mathcal{V}_p(\underline{\alpha}, D)$ in the following first remark:

Remark 4·3. In the proof of Theorem 4·1 we also constructed a minimum point $\underline{n}(D)$ of the function $\mathcal{D}(N, D) \to \mathbb{N} \cup \{\infty\}$, $\underline{n} \mapsto \nu_p(\underline{\alpha}, \underline{n})$. This point may be written as

$$\underline{n}(D) := \left(p^{\mu_i(D)} - 1\right)_{i=1}^{N},$$

where

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

with

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

and

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

Here, we regard sums of the form $\sum_{j=1}^{0}$ as zero and set $\alpha'_{\alpha_1+1} := 0$, so that $\alpha'_{Q(D)+1} = 0$ whenever $Q(D) = \alpha_1$. Within a graphic representation as in Figure 1 or 2, the parameter $\mu_i(D)$ is the number of red dots in the i^{th} row, $Q(D)$ is the number of columns that are

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

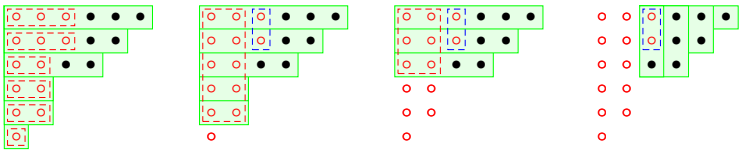


Fig. 2. Four ways to calculate $\mathcal{V}_p(\underline{\alpha}, D)$ (black dots) in Remark 4·4.

completely red, and $R(D)$ is the number of red dots in the next column, if a next column exists. If there actually is a next column, after the last completely red one, this column is not completely red, i.e., $R(D) < \alpha'_{Q(D)+1}$ if (and only if) $Q(D) < \alpha_1$.

Remark 4·4. Using the notations in Remark 4.3, the minimum value $\mathcal{V}_p(\underline{\alpha}, D)$ can also be expressed in the following forms:

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

These four formulas can be derived from the formula $\mathcal{V}_p(\underline{\alpha}, D) = \alpha - t(D) = \alpha - \lambda(\underline{\mu}(D))$ at the end of the proof of Theorem 4·1. The necessary simple rearrangements are illustrated in Ferrers diagrams in Figure 2. In each of the four cases, the minuends of the corresponding formula are shaded green, while the subtrahends are framed in red and blue. The difference $\mathcal{V}_p(\underline{\alpha}, D)$ is represented by black dots. If $\underline{\alpha} = (6, 5, 4, 2, 2, 1)$ and $t(D) = 13$ (red dots), then $\underline{\mu}(D) = (3, 3, 2, 2, 2, 1)$, $Q(D) = 2$, $\alpha'_{Q(D)} = 5$, $\alpha'_{Q(D)+1} = 3$, and $R(D) = 2$.

As in Corollary 1·8, our formula for $\mathcal{V}_p(\underline{\alpha}, D)$ simplifies if $\alpha_1 = \cdots = \alpha_N$. With the parameters

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

and with $\overline{h} := \max(h, 0)$ for real numbers h, we obtain the following corollary:

COROLLARY 4·5. *If* $\alpha_1 = \cdots = \alpha_N$ *then*

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

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

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

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 as in the proof of Corollary 1·8 in Section 3·5 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, using the last formula of Remark 4·4 (to variate the approach in Section 3·5),

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

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 4·2, 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. *Minimisation 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 Theorem 4·1)

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

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 $\alpha := \alpha_1 + \alpha_2 + \cdots + \alpha_N$, that the numbers $D_1, D_2, \ldots, D_\alpha$ are defined by

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

that the components of $\underline{\hat{n}}(\beta) = (\hat{n}_j(\beta))_{j=1}^r$ are given by

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

and that, for every $\hat{\underline{n}} = (\hat{n}_j)_{j=1}^r \in \mathbb{N}^r$,

$$[\hat{\underline{n}}] := [\hat{n}_1] \times [\hat{n}_2] \times \cdots \times [\hat{n}_r] \quad \text{with} \quad [\hat{n}_j] := \{0, 1, \ldots, \hat{n}_j\} .$$

5·1. *A Preparatory Lemma*

It turns out that the minimisation of $\mathcal{N}(\underline{n})$ leads to another optimisation problem that can be stated and solved in more general terms as follows:

LEMMA 5·1. *Assume $D \in \mathbb{N}$, and let $\alpha, \Lambda_1, \Lambda_2, \ldots, \Lambda_\alpha, V_1, V_2, \cdots \in \mathbb{Z}^+$. Suppose that $(\Lambda_t)_{t=1}^\alpha$ is monotone increasing, that $(V_t)_{t\in\mathbb{Z}^+}$ is monotone decreasing, and that $\Lambda_1 \le V_1$. Also assume that $V_t = V_1$ for all $1 \le t \le s_0$, where*

$$s_0 := \overline{\lceil (\Lambda_1 + \cdots + \Lambda_{t_0} - D)/V_1 \rceil} \quad \textit{with} \quad t_0 := \max\{1 \le t \le \alpha \mid \Lambda_t \le V_1\} .$$

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

$$\mathbf{S}(s) := s - \max\big\{0 \le t \le \alpha \mid \Lambda_1 + \cdots + \Lambda_t \le V_1 + V_2 + \cdots + V_s + D\big\}$$

has a minimum at the point s_0, and

$$\mathbf{S}(s_0) = \begin{cases} s_0 - t_0 & \textit{if } s_0 > 0, \\ -\max\big\{0 \le t \le \alpha \mid \Lambda_1 + \cdots + \Lambda_t \le D\big\} & \textit{if } s_0 = 0. \end{cases}$$

Proof. By definition, s_0 is the smallest element of $\mathbb{N}$ with

$$(\Lambda_1 + \cdots + \Lambda_{t_0} - D)/V_1 \le s_0 ,$$

i.e., with

$$\Lambda_1 + \cdots + \Lambda_{t_0} \;\le\; s_0 V_1 + D . \tag{5·1}$$

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 $(\Lambda_t)_{t=1}^\alpha$ to an infinite sequence by setting $\Lambda_{\alpha+1}, \Lambda_{\alpha+2}, \ldots := \infty$. With that extension $t_0 = \max\{t \in \mathbb{Z}^+ \mid \Lambda_t \le V_1\}$.

Case 1: $s_0 > 0$: In this case, by (5·1),

$$\Lambda_1 + \cdots + \Lambda_{t_0} \;\le\; V_1 + \cdots + V_{s_0} + D \tag{5·2}$$

but, by the minimality of s_0 in (5·1), also

$$\Lambda_1 + \cdots + \Lambda_{t_0} \;>\; V_1 + \cdots + V_{s_0-1} + D . \tag{5·3}$$

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 $\Lambda_1 \le \cdots \le \Lambda_{t_0} \le V_1 = \cdots = V_{s_0-1}$. 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 $0 < s \le s_0$,

$$\Lambda_1 + \cdots + \Lambda_{t_0-s+1} \;>\; V_1 + \cdots + V_{s_0-s} + D . \tag{5·4}$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

But, also $\cdots \geq \Lambda_{t_0+2} \geq \Lambda_{t_0+1} > V_1 = V_{s_0} \geq V_{s_0+1} \geq \ldots$. So, we can also add an equal number of subsequent summands on both sides of (5·3). For each $s \in \mathbb{N}$,

$$\Lambda_1 + \cdots + \Lambda_{t_0+s+1} \;>\; V_1 + \cdots + V_{s_0+s} + D\,. \tag{5·5}$$

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 (5·3) and (5·5) with $s = 0$ that

$$\mathbf{S}(s_0) \;=\; s_0 - t_0\,.$$

It follows from (5·4) that, for each $0 < s \leq s_0$,

$$\mathbf{S}(s_0 - s) \;\geq\; s_0 - s - (t_0 - s) \;=\; s_0 - t_0 \;=\; \mathbf{S}(s_0)\,.$$

And, it follows from (5·5) that, for each $s \in \mathbb{N}$,

$$\mathbf{S}(s_0 + s) \;\geq\; s_0 + s - (t_0 + s) \;=\; \mathbf{S}(s_0)\,.$$

We see that $\mathbf{S}$ attains a minimum at s_0 and $\mathbf{S}(s_0) = s_0 - t_0$.

Case 2: $s_0 = 0$: In this case, if we set

$$t(D) := \max\bigl\{0 \leq t \leq \alpha \mid \Lambda_1 + \cdots + \Lambda_t \leq D\bigr\} = \max\bigl\{t \in \mathbb{N} \mid \Lambda_1 + \cdots + \Lambda_t \leq D\bigr\}\,,$$

by the maximality of $t(D)$,

$$\Lambda_1 + \cdots + \Lambda_{t(D)+1} \;>\; D\,. \tag{5·6}$$

Moreover, $\cdots \geq \Lambda_{t(D)+2} \geq \Lambda_{t(D)+1} \geq \Lambda_{t_0+1} > V_1 \geq V_2 \geq \ldots$ since $t(D) \geq t_0$, by (5·1). Hence, we can add summands to (5·6), in the same way as we did it to get (5·5) from (5·3). For each $s \in \mathbb{N}$,

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

and thus

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

So, $\mathbf{S}$ attains a minimum at 0 and $\mathbf{S}(0) = -\max\bigl\{0 \leq t \leq \alpha \mid \Lambda_1 + \cdots + \Lambda_t \leq D\bigr\}$.

5·2. *The Minimum Value of $\mathcal{N}$ over $[\hat{\underline{n}}(\beta)]$*

We are ready to determine the minimum value of $\mathcal{N}(\underline{n})$ when $\underline{n}$ is ranging over $[\hat{\underline{n}}(\beta)]$:

LEMMA 5·2. *In the settings described at the beginning of Section 5, with the derived values $\breve{\alpha}_1, \ldots, \breve{\alpha}_N, \breve{\alpha}, \breve{\mathcal{A}}, \mathcal{B}$ as in Theorem 1·6, and for every integer*

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

we have

$$\min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n}) \;=\; \begin{cases} s_0 + \alpha - \breve{\alpha} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \alpha - \max\bigl\{1 \leq t \leq \alpha \mid D_1 + \cdots + D_t \leq \mathcal{B}\bigr\} & \text{if } \breve{\mathcal{A}} \leq \mathcal{B}. \end{cases}$$

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.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Step 1. If $n_1 \le p^{\beta_1}-1$ then $\overline{\left\lceil (n_1-(p^{\beta_1}-1))/((p-1)p^{\beta_1-1})\right\rceil}=0$. 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 $n_j \ge p^{\beta_j}-1$. In other words, we may write each 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

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

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

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

with updated ranges

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

Step 2. To find a minimum of $\mathcal{U}$ over $[\underline{\hat{u}}(\beta)] = \prod_{j=1}^r [\hat{u}_j(\beta)]$, we can replace the domain $[\hat{u}_j(\beta)] = \{0,1,\dots,\hat{u}_j(\beta)\}$ of each u_j with the smaller domain

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

Indeed, if the jth argument $u_j \in [\hat{u}_j]$ of $\mathcal{U}(u_1,\dots,u_r)$ is replaced with the first multiple of $p^{\beta_j-1}(p-1)$ above or equal to u_j (which still lies inside $[\hat{u}_j(\beta)] = [(\beta-1)p^{\beta_j-1}(p-1)]$), then the summand $\left\lceil u_j/(p^{\beta_j-1}(p-1))\right\rceil$ of $\mathcal{U}(\underline{u})$ stays the same and $\mathcal{U}(\underline{u})$ certainly does not increase. The minimum of $\mathcal{U}$ is already attained at a point $\underline{u}$ of the smaller domain $\prod_{j=1}^r \big(p^{\beta_j-1}(p-1)[\beta-1]\big) \subseteq [\underline{\hat{u}}(\beta)]$. 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

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

Step 3. In our search for the minimum value that $\mathcal{T}(t_1,\dots,t_r)$ may take, 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 term $d_jp^{\beta_j-1}(p-1)t_j$ 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

$$d_1p^{\beta_1-1}\ge d_2p^{\beta_2-1}\ge\cdots\ge d_rp^{\beta_r-1}$$

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

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

The only restriction to such modifications is that all argument t_j must stay within their domains $[\beta-1]$. They cannot increase above $\beta-1$ or go below 0. Through repeated applications of our modification, we can empty some t_j and fill others. This shows that the minimum is attained at a point of the form

$$(t_1,t_2,\dots,t_r)=(\beta-1,\beta-1,\dots,\beta-1,x,0,0,\dots,0)\,.$$

At such points, we have

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

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

$$\big(V_1,V_2,\dots,V_{r(\beta-1)}\big):=\big(\underbrace{d_1p^{\beta_1-1},\dots,d_1p^{\beta_1-1}}_{\beta-1\text{ times}},\ \dots,\ \underbrace{d_rp^{\beta_r-1},\dots,d_rp^{\beta_r-1}}_{\beta-1\text{ times}}\big)\,.$$

Hence, with the function

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

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 $D:=\mathcal{B}$, $\alpha:=\alpha_1+\cdots+\alpha_N$, and $\Lambda_t:=D_t$ for all $1\le t\le\alpha$. We also use the values V_t as defined above for all $1\le t\le r(\beta-1)$, and set $V_t:=V_{r(\beta-1)}$ for all $t>r(\beta-1)$. With the infinite sequence $(V_t)_{t\in\mathbb{Z}^+}$ 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

$$\begin{aligned}&\mathcal{V}_p\Big(\underline{\alpha}\,,(p-1)(V_1+\cdots+V_s+\mathcal{B})\Big)\\&\qquad=\alpha-\max\big\{0\le t\le\alpha\mid\Lambda_1+\cdots+\Lambda_t\le V_1+\cdots+V_s+\mathcal{B}\big\}\end{aligned}$$

makes sense for 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 $s\in\mathbb{N}$,

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

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

We also have $\Lambda_1 \le V_1$ as required in Lemma 5·1. Moreover, as in our situation the sequence (Λ_t) contains repetitions of lengths $\alpha'_1, \alpha'_2, \dots, \alpha'_{\alpha_1}$, the parameter

$$t_0 := \max\{1 \le t \le \alpha \mid \Lambda_t \le V_1\}$$

in Lemma 5·1 can be written as

$$t_0 = \alpha'_1 + \cdots + \alpha'_{i_0} \quad \text{with} \quad i_0 := \max\{1 \le i \le \alpha_1 \mid p^{i-1} \le d_1 p^{\beta_1 - 1}\}.$$

Here, the inequality $p^{i-1} \le d_1 p^{\beta_1 - 1}$ can be written as $i \le \beta_1 + \log_p(d_1)$, and the biggest integer i with this property is $\beta_1 + \lfloor \log_p(d_1) \rfloor$. But, in the definition of i_0 we also have the requirement $i \le \alpha_1$, so that

$$i_0 = \max\{1 \le i \le \alpha_1 \mid p^{i-1} \le d_1 p^{\beta_1 - 1}\} = \min(\alpha_1, \beta_1 + \lfloor \log_p(d_1) \rfloor) = \breve{\alpha}_1 .$$

With that and the second equation of Example 6·4 we get

$$t_0 = \alpha'_1 + \cdots + \alpha'_{i_0} = \alpha'_1 + \cdots + \alpha'_{\breve{\alpha}_1} = \breve{\alpha}_1 + \cdots + \breve{\alpha}_N = \breve{\alpha} .$$

With the third equation of Example 6·4, we further see that

$$\Lambda_1 + \cdots + \Lambda_{t_0} = \Lambda_1 + \cdots + \Lambda_{\breve{\alpha}} = \alpha'_1 p^0 + \cdots + \alpha'_{\breve{\alpha}_1} p^{\breve{\alpha}_1 - 1} = \sum_{i=1}^{N} \frac{p^{\breve{\alpha}_i} - 1}{p - 1} =: \breve{\mathcal{A}} .$$

In particular, the definition of s_0 in Lemma 5·1 coincides with the current one:

$$s_0 = \overline{\lceil (\Lambda_1 + \cdots + \Lambda_{t_0} - \mathcal{B})/V_1 \rceil} = \overline{\left\lceil \frac{\breve{\mathcal{A}} - \mathcal{B}}{d_1 p^{\beta_1 - 1}} \right\rceil} .$$

As assumed, this number is smaller than β, i.e., $s_0 \le \beta - 1$. This shows that $V_t = V_1$ for all $1 \le t \le s_0$, as required in Lemma 5·1. But, $s_0 \le \beta - 1$ 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

$$\begin{aligned} \min_{\underline{n} \in [\hat{\underline{n}}(\beta)]} \mathcal{N}(\underline{n}) &= \min_{s \in [r(\beta-1)]} \mathcal{S}(s) \\ &= \mathbf{S}(s_0) + \alpha \\ &= \begin{cases} s_0 + \alpha - \breve{\alpha} & \text{if } \breve{\mathcal{A}} > \mathcal{B}, \\ \alpha - \max\{1 \le t \le \alpha \mid D_1 + \cdots + D_t \le \mathcal{B}\} & \text{if } \breve{\mathcal{A}} \le \mathcal{B}. \end{cases} \end{aligned}$$

6. *About Conjugate Partitions*

In the previous sections, we repeatedly used results about conjugate sequences, a classic topic in partition theory within combinatorics. In this section, we discuss and prove those results in custom-made lemmas and examples, which we could not find in precisely that form in the literature. Given a sequence $(a_1, a_2, \dots, a_N)$ of integers with $N > 0$ and $a_1 \ge a_2 \ge \cdots \ge a_N > 0$, the *conjugate* numbers $a'_1, a'_2, \dots, a'_{a_1}$ are defined by

$$a'_j := \#\{1 \le i \le N \mid a_i \ge j\} .$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

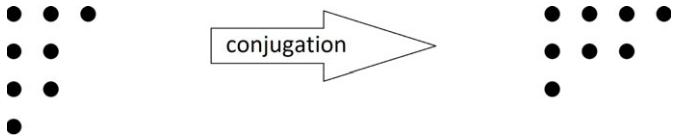


Fig. 3. The conjugate of (3, 2, 2, 1) is (4, 3, 1).

The sequence $\underline{a} := (a_i) \in (\mathbb{Z}^+)^N$ is a *partition* of $a := a_1 + a_2 + \cdots + a_N$, and the finite monotone decreasing sequence $\underline{a}' := (a'_j) \in (\mathbb{Z}^+)^{a_1}$ is called its *conjugate partition*, indicating that $a = a'_1 + a'_2 + \cdots + a'_{a_1}$. The conjugate partition is a dual partition of a in the sense that the conjugate of the conjugate is the original. All this is easy to see if we represent each a_i by a row of a_i dots in a so-called *Ferrers diagram*. The conjugate partition is then obtained by reflecting the corresponding Ferrers diagram about the main diagonal, similar to transposing a matrix:

6·1. *Two General Conjugation Lemmas*

Our first lemma, while certainly not a new deep insight, is formulated in a way that is helpful for our calculations.

LEMMA 6·1. *Let $(a_i) \in (\mathbb{Z}^+)^N$ be monotone decreasing, and $1 \le m \le a_1$. We have the following identity in $\mathbb{Z}[x]$:*

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

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

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

because

$$a_i \ge j \implies a_i \ge m \implies a_1, \ldots, a_i \ge m \implies i \le a'_m,$$

i.e., it is the same as the coefficient of x^j in the left polynomial.

The next lemma is obvious if we consider the minimum as given by the intersection of two Ferrers diagrams, because "intersecting" and "reflecting" commute.

LEMMA 6·2. *If the two sequences $(a_i), (b_i) \in (\mathbb{Z}^+)^N$ are monotone decreasing, then the sequence $(c_i) := \left(\min(a_i, b_i)\right) \in (\mathbb{Z}^+)^N$ is also monotone decreasing. The conjugate sequences (a'_j), (b'_j), and (c'_j) have lengths a_1, b_1, and $c_1 = \min(a_1, b_1)$, respectively; and for all $1 \le j \le c_1$,*

$$c'_j = \min(a'_j, b'_j).$$

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

6·2. *Special Cases*

Example 6·3. If $x = 1$ in Lemma 6·1, we obtain, for all $1 \le m \le a_1$,

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

If $m = 1$ then $a'_m = a'_1 = N$, and then $x = 1$ leads to

$$a'_1 + \cdots + a'_{a_1} = a_1 + \cdots + a_N .$$

If $m = 1$ and $x = p$, then Lemma 6·1 yields

$$a'_1 p^0 + \cdots + a'_{a_1} p^{a_1 - 1} = \sum_{i=1}^{N} (p^0 + p^1 + \cdots + p^{a_i - 1}) = \sum_{i=1}^{N} \frac{p^{a_i} - 1}{p - 1} .$$

Example 6·4. If the sequence (b_i) in Lemma 6·2 is constant equal to a number $b \in \mathbb{Z}^+$, we obtain as the conjugate of the sequence $(c_i) := \big(\min(a_i, b)\big) \in (\mathbb{Z}^+)^N$ the sequence

$$(c'_1, c'_2, \ldots, c'_{c_1}) = (a'_1, a'_2, \ldots, a'_{c_1}) .$$

So, if we apply Lemma 6·1 to (c_i) with $m = 1$ and $x = 1$, we get

$$a'_1 + \cdots + a'_{c_1} = c_1 + \cdots + c_N .$$

If instead $m = 1$ and $x = p$, we get

$$a'_1 p^0 + \cdots + a'_{c_1} p^{c_1 - 1} = \sum_{i=1}^{N} (p^0 + p^1 + \cdots + p^{c_i - 1}) = \sum_{i=1}^{N} \frac{p^{c_i} - 1}{p - 1} .$$

REFERENCES

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

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

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

[4] P.L. Clark and U. Schauz. Functional degrees and arithmetic applications I: the set of functional degrees. *J. Algebra* **608** (2022), 691–718.

[5] P.L. Clark and U. Schauz. Functional degrees and arithmetic applications II: the group-theoretic prime Ax–Katz theorem. Preprint: Arxiv: 2305.01304 (2023).

[6] P.L. Clark and N. Triantafillou. Ax's lemma and Ax's theorem in the Aichinger-Moosbaer calculus. *J. Algebra* **659** (2024), 542–580.

[7] D. Grynkiewicz. A generalisation of the Chevalley–Warning and Ax–Katz theorems with a view towards combinatorial number theory. *Combinatorica* **43** (2023), 1179–1213.

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

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

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

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

[12] E. Kummer. Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen. *J. Reine Angew. Math.* **44** (1852), 93–146.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press

[13] O. MORENO and C.J. MORENO. Improvements of the Chevalley–Warning and the Ax–Katz theorems. *Amer. J. Math.* **117** (1995), 241–244.
[14] M. MARSHALL and G. RAMAGE. Zeros of polynomials over finite principal ideal rings. *Proc. Amer. Math. Soc.* 49 (1975), 35–38.
[15] U. SCHAUZ. Classification of polynomial mappings between commutative groups. *J. Number Theory* **139** (2014), 1–28.
[16] U. SCHAUZ. The largest possible finite degree of functions between commutative groups. Preprint: Arxiv: 2103.16467 (2023).
[17] E. WARNING. Bemerkung zur vorstehenden Arbeit von Herrn Chevalley. *Abh. Math. Sem. Hamburg* **11** (1935), 76–83.
[18] C.S. WEISMAN. Some congruences for binomial coefficients. *Michigan Math. J.* **24** (1977), 141–151.
[19] R.M. WILSON. A lemma on polynomials modulo p^m and applications to coding theory. *Discrete Math.* **306** (2006), 3154–3165.

https://doi.org/10.1017/S0305004125000398 Published online by Cambridge University Press