# ON LARGE HALF-FACTORIAL SETS IN ELEMENTARY $p$-GROUPS: MAXIMAL CARDINALITY AND STRUCTURAL CHARACTERIZATION*

BY

ALAIN PLAGNE

*Centre de Mathématiques Laurent Schwartz, UMR 7640 du CNRS*
*École polytechnique, 91128 Palaiseau cedex, France*
*e-mail: plagne@math.polytechnique.fr*

AND

WOLFGANG A. SCHMID

*Institut für Mathematik und Wissenschaftliches Rechnen*
*Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria*
*e-mail: wolfgang.schmid@uni-graz.at*

ABSTRACT

Half-factoriality is a central concept in the theory of non-unique factorization, with applications for instance in algebraic number theory. A subset $G_0$ of an abelian group is called half-factorial if the block monoid over $G_0$, which is the monoid of all zero-sum sequences of elements of $G_0$, is a half-factorial monoid. In this paper we study half-factorial sets with large cardinality in elementary $p$-groups. First, we determine the maximal cardinality of such half-factorial sets, and generalize a result which has been only known for groups of even rank. Second, we characterize the structure of all half-factorial sets with large cardinality (in a sense made precise in the paper). Both results have a direct application in the study of some counting functions related to factorization properties of algebraic integers.

* This work was supported by the Austrian Science Fund FWF (Project P16770-N12) and by the Austrian-French Program 'Amadeus 2003-2004'.
Received March 11, 2004

## 1. Introduction and main results

A monoid (a commutative, cancellative semigroup with unit element) is called **atomic** if each non-unit $a \in H$ has a factorization $a = u_1 \cdot \ldots \cdot u_k$ with **atoms** (i.e. irreducible elements) $u_i \in H$. The integer $k$ is called the **length** of the factorization. An atomic monoid is called **half-factorial** if for each non-unit $a \in H$ the lengths of any two factorizations of $a$ into atoms are equal. Clearly, a domain is half-factorial (respectively, atomic) if and only if its multiplicative monoid is half-factorial (respectively, atomic).

Half-factoriality is a central topic in the theory of non-unique factorization (cf. e.g. [4, 33, 36, 37, 25, 7, 2, 3, 24] and [5] for a survey). If $H$ is a Krull monoid (cf. e.g. Chapters 22 and 23 in [21]), for example the multiplicative monoid of a Krull or a Dedekind domain, then whether $H$ is half-factorial, and sets of lengths of factorizations in general, just depend on the class group $G$ of $H$ and the subset $G_0 \subset G$ of classes containing prime divisors (cf. [12, Proposition 1]), namely $H$ is half-factorial if and only if $G_0 \subset G$ is a half-factorial set.

Let $G$ be an abelian group, additively written, and $G_0 \subset G$. Let $\mathcal{F}(G_0)$ be the free abelian monoid, multiplicatively written, generated by $G_0$ (equivalently, the set of all multi-sets in $G_0$). An element $\prod_{i=1}^{l} g_i \in \mathcal{F}(G_0)$ with $g_i \in G_0$ is called a **zero-sum sequence** (or, a **block**) if $\sum_{i=1}^{l} g_i = 0$. The block monoid $\mathcal{B}(G_0) \subset \mathcal{F}(G_0)$ is the submonoid of all zero-sum sequences in $G_0$. It is an atomic monoid (even a Krull monoid) and its atoms are the minimal zero-sum sequences, the set of which is denoted by $\mathcal{A}(G_0)$. The notion of block monoids was introduced by W. Narkiewicz in [27] and is meanwhile a main tool in the theory of non-unique factorization (cf. [1], in particular the surveys [6, 20]). We shall say that the set $G_0$ is **half-factorial** if the block monoid $\mathcal{B}(G_0)$ is a half-factorial monoid.

A realization result of L. Claborn [8] states that every abelian group is isomorphic to the class group of some Dedekind domain (see also the book [9]). Furthermore, if $G$ is an abelian group and $G_0 \subset G$ a generating (as a semigroup) subset, then there exists a Krull monoid (even a Dedekind domain) with class group isomorphic to $G$ such that $G_0$ corresponds to the set of classes containing prime divisors (cf. [18, 33, 19]). Thus to investigate half-factoriality for Krull monoids is equivalent to investigate half-factorial subsets of abelian groups ([10, 15, 22] are recent articles where this point of view is emphasized).

Apart from their importance in the investigation of half-factorial monoids, results on half-factorial sets can be applied when investigating other problems of non-unique factorization. For example, they are used to investigate differences occurring in sets of lengths that are long arithmetical progressions, namely results on half-factorial sets can be used to obtain results on the so-called $\Delta_1(G)$ set (cf. [13, 11]).

Another reason for the investigation of half-factorial sets and a main reason for introducing the notion of block monoids is the fact that half-factorial sets and related quantities occur when investigating the asymptotic behaviour of a certain counting function, $\mathbf{G}_k(x)$, defined via factorization properties of algebraic integers (cf. below for a definition, some results and references). The investigation of this counting function was initiated by W. Narkiewicz and the following constant was introduced in [34] in this context: For $G$ a finite abelian group, define

$$\mu(G) = \max\{|G_0| \,|\, G_0 \subset G \text{ half-factorial}\}.$$

The problem to determine $\mu(G)$ for arbitrary finite abelian groups is wide open and is open even for cyclic groups (cf. [22] for recent results). Yet, in [17, Theorem 8] A. Geroldinger and J. Kaczorowski could prove that the following chain of inequalities is valid for any elementary $p$-group $G$ of rank $r$:

$$(\dagger) \qquad 1+\left\lfloor\frac{r}{2}\right\rfloor p+2\left(\frac{r}{2}-\left\lfloor\frac{r}{2}\right\rfloor\right) \leq \mu(G) \leq 1+\frac{r}{2}p.$$

As a consequence, they obtained the exact value of $\mu(G)$ when $G$ is an elementary $p$-group of *even* rank, since the lower and the upper bounds in (†) coincide.

In this paper, starting from inequality (†), we first determine $\mu(G)$ when $G$ is an *arbitrary* elementary $p$-group.

THEOREM 1.1: *Let $G$ be an elementary p-group of rank $r$. Then,*

$$\mu(G) = \begin{cases} 2+\frac{r-1}{2}p & \text{if } r \text{ is odd,} \\ 1+\frac{r}{2}p & \text{if } r \text{ is even.} \end{cases}$$

In particular, this theorem shows that, in (†), equality always holds at the lower bound. This was up to now only known for $G$ of even rank and in some special cases (cf. below).

A second aim of this paper is to provide, again in the case where $G$ is an elementary $p$-group, a structural result on half-factorial sets $G_0 \subset G$ for which $|G_0|$ is sufficiently close to $\mu(G)$ (half-factorial sets with "large" cardinality). This result generalizes a result on the structure of half-factorial sets with maximal cardinality (that is, $|G_0| = \mu(G)$) in elementary $p$-groups of even rank obtained in [31, Theorem 3.1]. We shall prove the following theorem.

THEOREM 1.2: *There exists an absolute constant $c > 0$ such that for any elementary $p$-group $G$ of rank $r$, if $G_0 \subset G$ is a half-factorial subset with $|G_0| > \mu(G) - cp$, then there exists a basis $\{e_1, \dots, e_r\} \subset G$, such that*

$$G_0 \subset \bigcup_{i=1}^{\lfloor r/2 \rfloor} \{je_{2i-1} + (p+1-j)e_{2i} | \, j \in [1,p]\} \cup \{e_r, 0\}.$$

*In particular, if $\mu(G) = |G_0|$, then*

$$G_0 = \bigcup_{i=1}^{\lfloor r/2 \rfloor} \{je_{2i-1} + (p+1-j)e_{2i} | \, j \in [1,p]\} \cup \{e_r, 0\}.$$

If we denote by $c_0(p,r)$ the supremum of the values $c$ for which Theorem 1.2 is valid, then our proof will show that

$$c_0(p,r) \geq 1/12$$

for all $p$ and $r$. In Section 5, following the proof of this result, we will discuss possibilities to improve this bound.

For $p \leq 7$ the structure of all half-factorial subsets of elementary $p$-groups is already known (cf. [27, Problem II] for $p = 2$ and [31, Section 6]). In particular, the statement of Theorem 1.1 for $p \leq 7$ is just [31, Theorem 3.2.2] and Theorem 1.2 for $p \leq 7$ follows immediately from the structural characterization of half-factorial sets in [27, 31]. Thus we could assume $p \geq 11$. However, to unify the argument and since it makes it neither longer nor more complicated, we will provide the argument for $p \leq 7$ as well.

Both results (Theorems 1.1 and 1.2) have a direct application to the investigation of the counting function $\mathbf{G}_k(x)$. We recall some definitions and results: Let $H$ be an atomic monoid and $k$ some positive integer. Then the set $\mathcal{G}_k(H) \subset H$ is defined as the set of elements of $H$ with factorizations into atoms of at most $k$ different lengths. Let $K$ be an algebraic number field with class group $G$ and $R$ its ring of integers. Further, let $\mathcal{H}(R)$ denote the monoid of non-zero principal ideals of $R$, which is an atomic monoid. It is a well known result (cf. [4]) that $R$ is half-factorial if and only if $|G| \leq 2$. Suppose $|G| \geq 3$. The counting function is defined as

$$\mathbf{G}_k(x) = |\{I | \, \mathcal{N}(I) \leq x \text{ and } I \in \mathcal{G}_k(\mathcal{H}(R))\}|,$$

i.e. $\mathbf{G}_k(x)$ counts the (non-associated) elements of $R$ with factorizations of at most $k$ different lengths. The investigations on $\mathbf{G}_k(x)$ have been started in [26] and have been continued (respectively generalized to more general settings) by various authors (cf. e.g. [34, 23, 13, 17, 16, 29]).

In order to investigate $\mathbf{G}_k(x)$ one considers a main term, $\mathbf{M}_k(x)$. For $x \geq 1$ let

$$\mathbf{M}_k(x) = \frac{1}{2\pi i}\int_{\mathcal{C}} \zeta(s,\mathcal{G}_k)\frac{x^s}{s}ds,$$

where $\zeta(s,\mathcal{G}_k)$ is defined via $\zeta(s,\mathcal{G}_k) = \sum_{I\in\mathcal{G}_k(\mathcal{H}(R))}\mathcal{N}(I)^{-s}$ for $\Re(s) > 1$ and analytic continuation, and the contour of integration $\mathcal{C}$ goes counterclockwise around the points 1/2 and 1 (cf. [28, 29] and [23] for a similar definition).

For the investigation of the main term as well as for the investigation of the error term $\mathbf{R}_k(x) = \mathbf{G}_k(x) - \mathbf{M}_k(x)$ half-factorial sets with maximal cardinality and related combinatorial problems play an important role.

In particular it is known, in the present form obtained in [13], that

$$\mathbf{G}_k(x) \sim C_k\ x(\log x)^{-1+\mu(G)/|G|}(\log\log x)^{\psi_k(G)}$$

for some constant $C_k$, and where $\psi_k(G)$ is a non-negative integer that depends on the structure of half-factorial sets with cardinality $\mu(G)$. Roughly, $\psi_k(G)$ is equal to the number of elements one can add to a sequence in a half-factorial set with maximal cardinality without obtaining a zero-sum sequence having factorizations of more than $k$ different lengths, i.e. is not an element of $\mathcal{G}_k(\mathcal{B}(G))$. More precisely, $\psi_k(G)$ is defined as the maximal length of a sequence $S$ such that the following holds: There exists some half-factorial set $G_0 \subset G$ with $|G_0| = \mu(G)$ such that $S \in \mathcal{F}(G \setminus G_0)$ and $\emptyset \neq S\cdot\mathcal{F}(G_0)\cap\mathcal{B}(G) \subset \mathcal{G}_k(\mathcal{B}(G))$.

In view of this definition, results on the structure of half-factorial subsets with maximal cardinality of $G$ seem to be necessary in order to evaluate $\psi_k(G)$. It might be interesting to note that using such results, in particular the known special case for even rank of Theorem 1.2, $\psi_k(G)$ has been determined for some types of groups and values of $k$ (cf. [32]).

As mentioned above, problems related to half-factorial sets with maximal cardinality also occur in investigations of the error term $\mathbf{R}_k(x)$. In [29, Theorem 5] it is proved that if $\psi_k(G) > 0$, then the error term is subject to oscillations of positive lower logarithmic frequency and size $x^{1/2-\varepsilon}$. Moreover, it is conjectured there that $\psi_k(G) > 0$ for every finite abelian group $G$ and positive integer $k$. In [30] this conjecture is proved for $k \geq 2$ and arbitrary $G$, and some results related to the positivity of $\psi_1(G)$ are proved. One of these results ([30, Theorem 7.1]) and Theorem 1.1 now implies the conjecture in the case where $G$ is an elementary $p$-group. Thus, all this yields the following:

COROLLARY 1.3: *Let $K$ be an algebraic number field and $k$ a positive integer. If the ideal class group of $K$ is an elementary $p$-group with at least three elements,* *then the error term $\mathbf{R}_k(x)$ is subject to oscillations of positive lower logarithmic frequency and size $x^{1/2-\varepsilon}$.*

The paper is organized as follows: In Section 2 we fix notations and recall some known results. In Section 3 we develop the tools for the proofs of our results, in particular Lemma 3.3 and Proposition 3.4. In Section 4 we prove several preparatory results. Finally, in Section 5, the proofs of Theorem 1.1 and Theorem 1.2 are given. These are simple consequences of the preceding results.

For the sake of completeness and clarity, we have tried to make the present article as self-contained as possible. Hopefully, this approach will draw a unified picture of what is now known in the case of elementary $p$-groups.

## 2. Preliminaries

We denote by $\mathbb{Q}$ the set of rational numbers and by $\mathbb{Z}$ the set of integers. Throughout, $p$ will always denote a prime number and we shall denote by $\mathbb{F}_p$ the field with $p$ elements. For $r, s \in \mathbb{Q}$ let $[r,s] = \{z \in \mathbb{Z} \mid r \le z \le s\} \subset \mathbb{Z}$. It will be necessary to treat integers, and their multiplicative inverses and representatives modulo $p$ in the same equation. To this end we use the following notation: For a prime $p$ and $x \in \mathbb{Z}$ or $x \in \mathbb{F}_p$ we denote by $[x]_p$ its representative modulo $p$ in $[0, p-1]$. For $x \in \mathbb{Z}$ with $p \nmid x$ we denote by $x^{-1}$ its multiplicative inverse in $\mathbb{F}_p$ (although the notation $x^{-1}$ does not indicate with respect to which prime the inverse is taken, the context will always make this clear).

Elementary $p$-groups are in a natural way vector spaces over $\mathbb{F}_p$. Thus we make use of the notions of basis and dimension. However, to be consistent with the usual terminology for groups we will refer to the $\mathbb{F}_p$-dimension of some elementary $p$-group $G$ as the **rank** of $G$ and denote it by $\mathsf{r}(G)$. If $G$ is an elementary $p$-group, $G_1 \subset G$ an independent subset, $g \in \langle G_1 \rangle$, and $e \in G_1$, then we denote by $b_e(g, G_1) \in [0, p-1]$ the $e$-coordinate of $-g$ with respect to $G_1$, that is,

$$g = -\sum_{e \in G_1} b_e(g, G_1)e.$$

To consider negative coordinates will ease the formulation of several results in the sequel.

From the definition of half-factorial sets it follows that a set $G_0 \subset G$ is half-factorial if and only if $G_0 \cup \{0\}$ is half-factorial. Independent sets are half-factorial and $\{0, g\}$ is half-factorial for every $g \in G$.

A key result for the investigation of half-factorial subsets of finite abelian groups is the following characterization of half-factorial sets (cf. [33, 34, 36] and cf. [6] for a proof in the terminology used in this article): $G_0 \subset G$ is half-factorial if and only if

$$(\ddagger) \qquad \sum_{i=1}^{l} \frac{1}{\operatorname{ord}(g_i)} = 1$$

for each atom $\prod_{i=1}^{l} g_i \in \mathcal{A}(G_0)$.

In this paper we will only investigate elementary $p$-groups. From now on, we therefore assume that $G$ is an elementary $p$-group.

In this case, for each $g \in G \setminus \{0\}$, we have $\operatorname{ord}(g) = p$ and it is a convenient simplification to state the condition ($\ddagger$) for a set to be half-factorial as: $l = p$ for each atom $\prod_{i=1}^{l} g_i \in \mathcal{A}(G_0) \setminus \{0\}$.

Also, the following important assertion will be widely used in the sequel: if $a$ is an integer, the set $\{g, ag\}$ is half-factorial if and only if $ag \in \{0, g\}$. In other words, a half-factorial set cannot contain two distinct non-zero collinear elements. This assertion is quite immediate by ($\ddagger$) but can also be proved using Proposition 2.1.1 below.

For any given $G_0 \subset G$, we can determine some independent set $H_0 \subset G$ such that $\langle G_0 \rangle \oplus \langle H_0 \rangle = G$. Then $G_0$ is half-factorial if and only if $G_0 \cup H_0$ is half-factorial. Consequently, whenever it is convenient we will suppose that a given half-factorial set $G_0 \subset G$ generates $G$ (cf. e.g. [10, Lemma 3.1]).

In particular, if $G_0 \subset G$ is half-factorial with maximal cardinality $|G_0| = \mu(G)$, then $0 \in G_0$ and $\langle G_0 \rangle = G$. Note that this does not hold for arbitrary finite abelian groups: it is shown in [10, Corollary 6.5] that it is not the case, among others, for groups of the form $(\mathbb{Z}/p^k\mathbb{Z})^{p+1}$ for any $k \geq 4$ and prime $p$.

In the following propositions we recall some further results on half-factorial sets.

PROPOSITION 2.1: *Let $G$ be an elementary $p$-group of rank $r$ and $\{e_1, \dots, e_r\} \subset G$ a basis. Further, let $g = -\sum_{i=1}^{r} b_i e_i$ with $b_i \in [0, p-1]$ for each $i \in [1, r]$ be a non-zero element of $G$ such that $\{g, e_1, \dots, e_r\}$ is half-factorial, then*

(1) $\sum_{i=1}^{r} b_i = p - 1$,

(2) *if $b_1 \neq 0$, then*

$$\sum_{i=1}^{r} [b_1^{-1} b_i]_p = p - [b_1^{-1}]_p.$$

For ease of notation, Proposition 2.1.2 is just formulated for the first coordinate but clearly applies (and will be applied) to any other one. The same remark applies to the forthcoming Proposition 2.2 and Lemma 3.3.

Proposition 2.1.1 was initially proved in [34, Lemma 1] (cf. also [35]) as an application of Equation (‡) to the minimal zero-sum sequence $g\prod_{i=1}^{r} e_i^{b_i}$. Proposition 2.1.2 follows by exchanging the role of $g$ and $e_j$ using a change of basis formula (see for instance Proposition 2.2.2).

It is important to note that the equalities in Proposition 2.1 (and more generally throughout this paper) are equalities in $\mathbb{Z}$ and not just modulo $p$. In particular, it is necessary to fix a certain set of representatives modulo $p$, in our case $[0, p-1]$. This choice and the, somewhat artificial, representation of elements with negative coordinates, turn out to be an efficient way to express these and related results.

Proposition 2.2: *Let $G$ be an elementary $p$-group of rank $r$ and $\{e_1, \dots, e_r\} \subset G$ a basis. Further, let $g = -\sum_{i=1}^{r} b_i e_i$ and $h = -\sum_{i=1}^{r} b_i' e_i$ with $b_i, b_i' \in [0, p-1]$ for each $i \in [1, r]$ be non-zero elements of $G$ such that $\{g, h, e_1, \dots, e_r\}$ is half-factorial.*

(1) *If $b_1 = b_1' \neq 0$, then $g = h$.*
(2) *If $b_1 \neq 0$, then*

$$\sum_{i=1}^{r} c_i = p - 1$$

*where $c_1 = [-b_1^{-1} b_1']_p \in [0, p-1]$ and $c_i = [c_1 b_i + b_i']_p \in [0, p-1]$ for each $i \in [2, r]$.*

Proposition 2.2.2 was proved in [31, Proposition 4.2]. It follows by observing that $h = -c_1 g - \sum_{i=2}^{r} c_i e_i$, and applying Proposition 2.1.1. Historically, Proposition 2.2.1 was first proved in [17, Lemma 1] but it can also be obtained as a consequence of Proposition 2.2.2, since $b_1 = b_1' \neq 0$ implies $c_1 = p-1$ and thus $c_i = 0$ for each $i \in [2, r]$.

## 3. Basic results

As can be noticed from the preceding propositions we will frequently consider coordinates of some elements in various bases. Thus we introduce some notation concerning these coordinates.

*Definition 3.1:* Let $G$ be an elementary $p$-group.

(1) If $G_1 \subset G$ is an independent set and $g \in \langle G_1 \rangle$, then

$$\mathsf{n}(g, G_1) = |\{e \in G_1 \mid b_e(g, G_1) \neq 0\}|$$

denotes the number of non-zero coordinates of $g$ with respect to $G_1$. Moreover, if $H_1 \subset G_1$, then

$$\mathsf{n}(g, H_1, G_1) = |\{e \in H_1 | \ b_e(g, G_1) \neq 0\}|$$

denotes the number of non-zero coordinates of $g$ among those of $H_1$. (Clearly, $\mathsf{n}(g, G_1, G_1) = \mathsf{n}(g, G_1)$.)

(2) If $G_0 \subset G$ is a nonempty set $\neq \{0\}$, then $\mathsf{N}(G_0)$ denotes the maximal number of non-zero coordinates of the elements $g \in G_0$ with respect to some independent set $G_1 \subset G_0$ that generates $\langle G_0 \rangle$, i.e.

$$\mathsf{N}(G_0) = \max\{\mathsf{n}(g, G_1) | \ g \in G_0, \ G_1 \subset G_0 \text{ a basis of } \langle G_0 \rangle\}.$$

It is obvious that $1 \leq \mathsf{N}(G_0) \leq \mathsf{r}(G)$ for every $G_0 \subset G$. If however $G_0$ is half-factorial, then it follows by Proposition 2.1.1 that also

$$\mathsf{N}(G_0) \leq p - 1.$$

It might be interesting to note that in general not for every $j \in [1, \min\{\mathsf{r}(G), p-1\}]$ there exists some half-factorial set $G_0 \subset G$ with $\mathsf{N}(G_0) = j$ (cf. [31, Proposition 4.7]). However, we will only make use (in the proof of the theorems) of the fact, which can be obtained easily by Proposition 2.1, that for $p = 5$ there does not exist a half-factorial set $G_0 \subset G$ with $\mathsf{N}(G_0) = 3$.

Note that our definitions for $\mathsf{n}(\cdot)$ and $\mathsf{N}(\cdot)$ differ from the ones in [17] and [10]. However the underlying idea in its application is essentially the same: By Proposition 2.2.1, it follows that if $G_0 \subset G$ is half-factorial, $G_1 \subset G_0$ is a basis, $g, h \in G_0$ and $e \in G_1$, then $b_e(g, G_1) = b_e(h, G_1) \neq 0$ implies $g = h$. Formulated in a negative way this means that, in a half-factorial set, two different elements cannot have a common non-zero coordinate. Since the total number of non-zero coordinates is bounded by $\mathsf{r}(G)(p-1)$ and clearly only one element has no non-zero coordinates, this yields an upper bound of $1 + \mathsf{r}(G)(p-1)$ for the cardinality of a half-factorial set in $G$. To improve this upper bound it is necessary to use further results on the coordinates of the elements of $G_0$. Most of these results, clearly, depend on the fact that the set $G_0$ is half-factorial. However, the following result holds in general, even for arbitrary vector spaces. Since we will frequently use it, we state it as a lemma and give a short proof of it.

LEMMA 3.2: *Let $G$ be an elementary $p$-group, $g, h \in G$ and $G_1 \subset G$ be a basis. Suppose $e' \in G_1$ such that $b_{e'}(g, G_1) \neq 0$ and $b_{e'}(h, G_1) \neq 0$. Then the set*

*$G_1' = (G_1 \setminus \{e'\}) \cup \{g\}$ is a basis and*

$$\mathsf{n}(h, G_1') \geq \mathsf{n}(h, G_1) + \mathsf{n}(g, G_1) - 2u + 1$$

*with $u = |\{e \in G_1 | \ b_e(g, G_1) \neq 0 \text{ and } b_e(h, G_1) \neq 0\}|$.*

*Proof:* Since $b_{e'}(g, G_1) \neq 0$, obviously $G_1'$ is a basis. If we let

$$c_g = [-b_{e'}(g, G_1)^{-1} b_{e'}(h, G_1)]_p \quad \text{and} \quad c_e = [c_g b_e(g, G_1) + b_e(h, G_1)]_p$$

for each $e \in G_1' \setminus \{g\}$, then a change of basis formula yields

$$h = - \sum_{e \in G_1'} c_e e.$$

We have $c_g \neq 0$. For $e \in G_1' \setminus \{g\}$ it follows that if $b_e(g, G_1) \neq 0$ and $b_e(h, G_1) = 0$, then $c_e \neq 0$, and the same is true if $b_e(h, G_1) \neq 0$ and $b_e(g, G_1) = 0$. ∎

The following two results will play a key role in the remainder of this paper. As already mentioned, the main idea to obtain upper bounds for the cardinality of half-factorial sets in elementary $p$-groups, which was initially applied in [17], is to exploit the fact that for two different elements no two non-zero coordinates are equal. In the following lemma we obtain a further condition for the coordinates of different elements in a half-factorial set.

LEMMA 3.3: *Let $G$ be an elementary $p$-group of rank $r$ and $G_1 = \{e_1, \dots, e_r\} \subset G$ a basis. Further, let $g = -\sum_{i=1}^{r} b_i e_i$ and $h = -\sum_{i=1}^{r} b_i' e_i \in G \setminus \{0\}$ with $b_i, b_i' \in [0, p-1]$ for each $i \in [1, r]$ and $G_0 = \{g, h, e_1, \dots, e_r\}$. If $G_0$ is half-factorial and $b_1' = [2b_1]_p$, then at least one of the following statements holds:*

(i) $b_1 = b_1' = 0$,
(ii) $h = 2g - e_j$ *for some* $j \in [1, r]$ *and* $\mathsf{n}(g, G_1) \leq 2$,
(iii) $h = 2g - e_j$ *for some* $j \in [1, r]$ *and* $b_j = 0$.

*Proof:* If $p = 2$, then the lemma is clearly true. Therefore we assume $p \neq 2$.

Suppose $G_0$ is half-factorial and $b_1' = [2b_1]_p$. If $b_1 = 0$, then nothing needs to be proved. Thus suppose $b_1 \neq 0$.

By Proposition 2.2.2, we have $\sum_{i=1}^{r} c_i = p - 1$ with

$$c_1 = [-b_1^{-1} b_1']_p = [-2]_p = p - 2 \quad \text{and} \quad c_i = [c_1 b_i + b_i']_p = [-2b_i + b_i']_p$$

for each $i \in [2, r]$. Thus we have $c_j = 1$ for some $j \in [2, r]$ and $c_i = 0$ for each $i \in [2, r] \setminus \{j\}$. Without restriction let $j = 2$. For $i \in [3, r]$ we have

$[-2b_i + b_i']_p = 0$ and thus $b_i' \equiv 2b_i \bmod p$. Since $\sum_{i=1}^{r} b_i = \sum_{i=1}^{r} b_i' = p-1$, it follows that $b_2' \equiv 2b_2 + 1 \bmod p$. Thus we have $h = 2g - e_2$. It remains to show that $\mathsf{n}(g, G_1) \le 2$ or $b_2 = 0$.

Assume $b_2 \ne 0$. Since $b_i = 0$ implies $b_i' = 0$ for $i \in [1,r] \setminus \{2\}$, it follows that $G_0$ is half-factorial if and only if $G_0 \setminus \{e_i \mid i \in [1,r] \text{ and } b_i = 0\}$ is half-factorial. Thus we may assume without restriction that $b_i \ne 0$ for each $i \in [1,r]$, an assumption which implies $\mathsf{n}(g, G_1) = r$.

We use again Proposition 2.2.2, now applied to the second coordinate, and obtain $\sum_{i=1}^{r} c_i' = p - 1$ with $c_2' = [-b_2^{-1}b_2']_p = [-2 - b_2^{-1}]_p$ and

$$c_i' = [c_2' b_i + b_i']_p = [(-2 - b_2^{-1})b_i + 2b_i]_p = [-b_2^{-1}b_i]_p$$

for each $i \in [1,r] \setminus \{2\}$. Note that $[-b_2^{-1}b_i]_p = p - [b_2^{-1}b_i]_p$ and by Proposition 2.1.2, $p - [b_2^{-1}]_p = \sum_{i=1}^{r} [b_2^{-1}b_i]_p$. Thus we have

$$\begin{aligned} p - 1 &= [-2 - b_2^{-1}]_p + \sum_{i=1, i\ne 2}^{r} [-b_2^{-1}b_i]_p \\ &= [-2 - b_2^{-1}]_p + (r-1)p - \sum_{i=1, i\ne 2}^{r} [b_2^{-1}b_i]_p \\ &= [-2 - b_2^{-1}]_p + (r-1)p - (p - [b_2^{-1}]_p - 1) > (r-2)p. \end{aligned}$$

This implies $\mathsf{n}(g, G_1) = r \le 2$, which finishes the proof. ∎

PROPOSITION 3.4: *Let $G$ be an elementary $p$-group of rank $r$ and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) = s \ge 3$. Further, let $G_1 \subset G_0$ be a basis, $H_1 \subset G_1$ and $G_j = \{g \in G_0 \mid \mathsf{n}(g, G_1) = j\}$ for each $j \in [2, s]$. Then $G_0 \setminus \{0\}$ is equal to the disjoint union $\dot{\bigcup}_{l=1}^{s} G_l$ and*

(1) *for each $j \in [3, s]$,*

$$\sum_{g \in G_0} \mathsf{n}(g, H_1, G_1) + \sum_{g \in G_j} \mathsf{n}(g, H_1, G_1) \le |H_1|(p-1),$$

(2) *for each $j \in [3, s]$,*

$$\sum_{l=1}^{s} l|G_l| + j|G_j| \le r(p-1),$$

(3) *the following inequality holds:*

$$5|G_0| - 3|G_2| \le r(p-1) + 4r + 5.$$

*Proof:* Since $p - 1 \ge \mathsf{N}(G_0) \ge 3$, we may assume $p \ge 5$ (see the discussion after Definition 3.1).

By the definitions of $\mathsf{N}(G_0)$ and of $G_j$ for $j \in [2,s]$, in order to prove $G_0 \setminus \{0\} = \dot{\bigcup}_{l=1}^{s} G_l$, it suffices to observe that if $\mathsf{n}(g, G_1) = 1$ for $g \in G_0$, then $g \in G_1$, as stated in Section 2.

(1) By definition of $\mathsf{n}(\cdot)$ we have

$$\sum_{e \in H_1} \sum_{g \in G'} \mathsf{n}(g, \{e\}, G_1) = \sum_{g \in G'} \mathsf{n}(g, H_1, G_1)$$

for every $G' \subset G$. Thus it suffices to prove the statement for $|H_1| = 1$. Let $H_1 = \{e\}$.

We define $G'_0$ to be the subset of $G_0$ composed of the elements $g$ such that $b_e(g, G_1) \neq 0$ and let $G'_l = G_l \cap G'_0$ for each $l \in [1,s]$. Notice that $G'_0 = \dot{\cup}_{l=1}^{s} G'_l$.

Let us consider the two maps $\beta$ and $\gamma$ from $G'_0$ to $[1, p-1]$ defined by $\beta(g) = b_e(g, G_1)$ and $\gamma(g) = [2b_e(g, G_1)]_p$, respectively.

By Proposition 2.2.1 we have that $b_e(g, G_1) = b_e(h, G_1)$ for $g, h \in G'_0$ if and only if $g = h$, i.e., $\beta$ is injective and thus $\gamma$ is injective as well. Moreover, if $g \in G'_j$ for some $j \geq 3$, then $b_e(h, G_1) = [2b_e(g, G_1)]_p$ implies $h = 2g - e'$ for some $e' \in G_1$ such that $b_{e'}(g, G_1) = 0$ (indeed, this follows from Lemma 3.3 in which, by our assumptions, only case (iii) can happen): This, in particular, implies $h \in G'_{j+1}$. Therefore we have

$$\beta(G'_0) \cap \gamma(\cup_{l=j}^{s} G'_l) \subset \beta(\cup_{l=j+1}^{s} G'_l),$$

which implies

$$p - 1 \geq |\beta(G'_0) \cup \gamma(\cup_{l=j}^{s} G'_l)| \geq |\beta(G'_0)| + |\gamma(\cup_{l=j}^{s} G'_l)| - |\beta(\cup_{l=j+1}^{s} G'_l)|.$$

Since the maps $\beta$ and $\gamma$ are injective and the unions are disjoint, the right-hand side in this inequality is exactly $|G'_0| + |G'_j|$. But, using

$$\mathsf{n}(g, \{e\}, G_1) = \begin{cases} 1 & \text{if } b_e(g, G_1) \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

we observe that

$$|G'_0| = \sum_{g \in G'_0} \mathsf{n}(g, \{e\}, G_1) = \sum_{g \in G_0} \mathsf{n}(g, \{e\}, G_1)$$

and, for any $j \in [3,s]$,

$$|G'_j| = \sum_{g \in G'_j} \mathsf{n}(g, \{e\}, G_1) = \sum_{g \in G_j} \mathsf{n}(g, \{e\}, G_1).$$

This implies

$$\sum_{g\in G_0} \mathsf{n}(g,\{e\},G_1) + \sum_{g\in G_j} \mathsf{n}(g,\{e\},G_1) \le p-1,$$

and the result follows.

(2) We note that by definition $\mathsf{n}(g,G_1) = j$ for every $g \in G_j$ and $j \in [1,s]$, and moreover $G_0 \subset \{0\} \cup (\dot\cup_{l=1}^{s} G_l)$. Thus the statement follows immediately by (1) with $H_1 = G_1$.

(3) By (2) with $j = 3$ we obtain (setting $G_l = \emptyset$ for $l > s$)

$$|G_1| + 2|G_2| + 6|G_3| + 4|G_4| + \sum_{l=5}^{s} l|G_l| \le r(p-1)$$

and with $j = 4$ we obtain

$$|G_1| + 2|G_2| + 3|G_3| + 8|G_4| + \sum_{l=5}^{s} l|G_l| \le r(p-1).$$

Adding $5/7$ times the first and $2/7$ times the second inequality then yields

$$|G_1| + 2|G_2| + \frac{36}{7}(|G_3| + |G_4|) + \sum_{l=5}^{s} l|G_l| \le r(p-1).$$

Noting that $|G_0| \le 1 + \sum_{l=1}^{s} |G_l|$, we finally obtain

$$\begin{aligned} 5|G_0| - 3|G_2| &\le 5 + 5|G_1| + 2|G_2| + 5(|G_3| + |G_4|) + \sum_{l=1}^{s} 5|G_l| \\ &\le 5 + 4r + \left( |G_1| + 2|G_2| + \frac{36}{7}(|G_3| + |G_4|) + \sum_{l=5}^{s} l|G_l| \right) \\ &\le 5 + 4r + r(p-1). \quad \blacksquare \end{aligned}$$

*Remark 3.5:* In the proof of Proposition 3.4.3, by considering the inequalities obtained in (2), 35/50 times the one for $j = 3$, 14/50 times the one for $j = 4$, and 1/50 times the one for $j = 5$, one could improve this result to

$$\frac{51}{10}|G_0| - \frac{31}{10}|G_2| \le r(p-1) + \frac{41}{10}r + \frac{51}{10}.$$

However, this improvement would not affect our main results, not even the lower bound for $c_0(p,r)$ that we obtain. Thus we favoured the shorter proof and in particular the nicer constants.

## 4. Preparatory results

In this section we investigate half-factorial sets $G_0 \subset G$ with $\mathsf{N}(G_0) \le 2$, $\mathsf{N}(G_0) = 3$ and $\mathsf{N}(G_0) \ge 4$, respectively.

In the following proposition we describe the structure of half-factorial sets $G_0 \subset G$ with $\mathsf{N}(G_0) \le 2$. It is a reformulation of known results, in a way suitable for our application (cf. in particular [34, Lemma 1], [17, Proof of Theorem 8], [31, Proof of Theorem 3.1]). For the sake of completeness, we provide a proof.

Proposition 4.1: *Let $G$ be an elementary $p$-group of rank $r$, $G_0 \subset G$ and $G_1 \subset G_0$ a basis of $G$.*

(1) *$G_0$ is half-factorial with $\mathsf{N}(G_0) = 1$ if and only if $G_0 \subset \{0\} \cup G_1$.*

(2) *$G_0$ is half-factorial with $\mathsf{N}(G_0) \le 2$ if and only if*

$$G_0 \subset \bigcup_{i=1}^{\lfloor r/2 \rfloor} \{je_{2i-1} + (p+1-j)e_{2i} \mid j \in [1,p]\} \cup \{e_r, 0\},$$

*where $G_1 = \{e_1, \dots, e_r\}$.*

*Proof:* (1) The set on the right-hand side is half-factorial: This follows for instance from Equation (‡), since the set of atoms of $G_0 = \{0\} \cup G_1$ is given by $\mathcal{A}(G_0) = \{e_i^p \mid i \in [1,r]\} \cup \{0\}$ (or simply because independent sets are half-factorial in arbitrary abelian groups, cf. the discussion in Section 2). By definition of $\mathsf{n}(\cdot)$ it follows that $\mathsf{N}(G_0) = 1$ and the 'if'-part follows. The 'only if'-part follows since a half-factorial set cannot contain two distinct non-zero collinear elements.

(2) That the set on the right-hand side is half-factorial is proved in [17, Proof of Theorem 8], which clearly implies the 'if'-part.

We now prove the 'only if'-part. Let $G_0 \subset G$ be half-factorial with $G_1 \subset G_0$ and $\mathsf{N}(G_0) = 2$. By (1) it suffices to consider $\mathsf{N}(G_0) = 2$. Let $g \in G_0$ with $\mathsf{n}(g, G_1) = 2$. Such an element exists since otherwise it would follow that $G_0 \subset \{0\} \cup G_1$ and therefore $\mathsf{N}(G_0) = 1$.

Let $\{e_1, e_2\} \subset G_1$ such that $\mathsf{n}(g, \{e_1, e_2\}, G_1) = 2$, i.e.

$$g = -b_1e_1 - b_2e_2 \text{ with } b_1, b_2 \in [1, p-1].$$

By Proposition 2.1.1 we have that $b_1 + b_2 = p-1$, thus there exist some $j \in [2, p-1]$ such that $g = -(p-j)e_1 - (j-1)e_2 = je_1 + (p+1-j)e_2$. It remains to prove that if $h \in G_0 \setminus \{e_1, e_2\}$ with $b_{e_i}(h, G_1) \ne 0$ for some $i \in [1,2]$, then $\mathsf{n}(h, G_1) = 2$ and $h = -b_1'e_1 - b_2'e_2$ with $b_1', b_2' \in [1, p-1]$. Let $h \in G_0 \setminus \{e_1, e_2\}$ and suppose, without loss of generality, that $b_{e_1}(h, G_1) \neq 0$. As above, $\mathsf{n}(h, G_1) > 1$ and thus $\mathsf{n}(h, G_1) = 2$, thus $h = -b_1' e_1 - b_j' e_j'$ with $j \in [2, r]$ and $b_1', b_j' \in [1, p-1]$. We set $G_1' = (G_1 \setminus \{e_1\}) \cup \{g\}$ and obtain by Lemma 3.2 that

$$\mathsf{n}(h, G_1') \geq 5 - 2|\{e \in G_1 \mid b_e(g, G_1) \neq 0 \text{ and } b_e(h, G_1) \neq 0\}|.$$

Since $\mathsf{n}(h, G_1') \leq \mathsf{N}(G_0) = 2$, we obtain

$$|\{e \in G_1 \mid b_e(g, G_1) \neq 0 \text{ and } b_e(h, G_1) \neq 0\}| > 1$$

and the statement follows. ∎

Note that this proposition describes completely the structure of half-factorial sets of elementary 2- and 3-groups, since $\mathsf{N}(G_0) \leq p - 1$ (cf. the discussion following Definition 3.1). In particular, it already proves Theorem 1.1 and Theorem 1.2 for elementary 2- and 3-groups.

From Proposition 4.1, it also follows that $\mu(G) \geq 1 + rp/2$ if $r$ is even and $\mu(G) \geq 2 + (r-1)p/2$ if $r$ is odd, since the sets on the right-hand side of the formula in (2) have these cardinalities — notice that in (2), $e_r$ belongs to the union

$$\bigcup_{i=1}^{\lfloor r/2 \rfloor} \{j e_{2i-1} + (p+1-j) e_{2i} \mid j \in [1, p]\}$$

if and only if $r$ is even.

Finally, Proposition 4.1 implies the following corollary. Note that the condition on the set $G_0$ to be generating, which is made in this and several of the following statements, is of purely technical nature. Obviously, the result can be extended to arbitrary half-factorial sets.

Corollary 4.2: *Let $G$ be an elementary $p$-group of rank $r$ and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) \leq 2$. Then,*

$$|G_0| \leq \begin{cases} 1 + rp/2 & \text{if } r \text{ is even,} \\ 2 + (r-1)p/2 & \text{if } r \text{ is odd.} \end{cases}$$

In particular, if $\mathsf{r}(G) = 1$, then $\mu(G) = 2$ and if $\mathsf{r}(G) = 2$, then $\mu(G) = p + 1$ (cf. e.g. [34, 35, 17]).

Next we investigate half-factorial sets $G_0 \subset G$ with $\mathsf{N}(G_0) = 3$. First, in Lemma 4.3, we treat the case $\mathsf{r}(G) = 3$. In Corollary 4.5 we will generalize this result to elementary $p$-groups of arbitrary rank.

LEMMA 4.3: *Let $G$ be an elementary $p$-group of rank 3 and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) = 3$. Then,*

$$|G_0| \le 3 + 5p/6.$$

*Proof:* Let $G_1 = \{e_1, e_2, e_3\} \subset G_0$ be a basis and $g \in G_0$ with $\mathsf{n}(g, G_1) = 3$, i.e. $g = -b_1e_1 - b_2e_2 - b_3e_3 \in G_0$ with $b_i \in [1, p-1]$ for each $i \in [1,3]$. By Proposition 2.1.1, we have $b_1 + b_2 + b_3 = p - 1$. Further, let

$$G_2 = \{g' \in G_0 |\ \mathsf{n}(g', G_1) = 2\} \quad \text{and} \quad G_3 = \{g' \in G_0 |\ \mathsf{n}(g', G_1) = 3\}.$$

Obviously, the sets $G_2$ and $G_3$ depend on the choice of the basis $G_1$. We will assert that for a suitable basis (in $G_0$) the subset of elements with 3 non-zero coordinates will not be too small relative to $G_0$. Having this at hand we will apply Proposition 3.4 and obtain the upper bound for $|G_0|$.

Let $G_2^{(1)} = \{g' \in G_2 |\ b_{e_1}(g', G_1) \ne 0\}$ and assume without restriction that

$$|G_2^{(1)}| \ge 2|G_2|/3.$$

Further, let $G_1'$ denote the basis $\{g, e_2, e_3\}$. We denote

$$G_2' = \{g' \in G_0 |\ \mathsf{n}(g', G_1') = 2\} \quad \text{and} \quad G_3' = \{g' \in G_0 |\ \mathsf{n}(g', G_1') = 3\}.$$

We assert that

$$|G_3'| \ge |G_2^{(1)}| - 1.$$

Assume for a moment that this assertion is true and let us conclude the proof. By definition $|G_2| + |G_3| = |G_2'| + |G_3'| \ge |G_0| - |G_1| - 1 = |G_0| - 4$ (with equality if and only if $0 \in G_0$) and thus, by the assertion,

$$|G_3'| \ge |G_2^{(1)}| - 1 \ge \frac{2|G_2|}{3} - 1 \ge \frac{2(|G_0| - |G_3| - 4)}{3} - 1.$$

This implies $5\max\{|G_3|, |G_3'|\} \ge 3|G_3'| + 2|G_3| \ge 2|G_0| - 11$. Without restriction suppose $|G_3| \ge (2|G_0| - 11)/5$. By Proposition 3.4.2 (applied with $j = 3$) we have

$$3(p-1) \ge |G_1| + 2|G_2| + 6|G_3|$$

and therefore, since $|G_1| = 3$ and $|G_2| \ge |G_0| - 4 - |G_3|$,

$$3(p-1) \ge 2|G_0| + 4|G_3| - 5 \ge 2|G_0| + 4\Big(\frac{2|G_0| - 11}{5}\Big) - 5.$$

This gives

$$|G_0| \le (5p + 18)/6,$$

which is the desired bound.

Let us now prove the assertion that $|G_3'| \geq |G_2^{(1)}| - 1$. It is trivial if $G_2^{(1)} = \emptyset$. Therefore we assume $G_2^{(1)} \neq \emptyset$ and consider $h \in G_2^{(1)}$. Then $h = -a_1 e_1 - a_j e_j$ with $j \in [2,3]$, $a_1, a_j \in [1, p-1]$ and $a_1 + a_j = p - 1$. Let $k \in [2,3]$ such that $\{j, k\} = [2,3]$. We may compute

$$h = [a_1 b_1^{-1}]_p g + [a_1 b_1^{-1} b_k]_p e_k + [a_1 b_1^{-1} b_j - a_j]_p e_j,$$

from which it is clear that $b_g(h, G_1') \neq 0$ and $b_{e_k}(h, G_1') \neq 0$.

We show that $b_{e_j}(h, G_1') = 0$ if and only if $a_j = [-b_1^{-1} b_j (1 + b_1^{-1} b_j)^{-1}]_p$. First note that this expression for $a_j$ is well-defined, since $b_1 \neq 0$ and $0 < b_1 + b_j < p$. Since $a_1 + a_j = p - 1$ and thus $a_1 = [-1 - a_j]_p$, we can solve the equation $[-b_1^{-1} a_1 b_j + a_j]_p = 0$ with respect to $a_j$ and obtain the unique solution

$$a_j = [-b_1^{-1} b_j (1 + b_1^{-1} b_j)^{-1}]_p,$$

and we are done.

Therefore we have

$$G_2^{(2)} = \{g' \in G_2^{(1)} \mid b_{e_i}(g', G_1) \neq [-b_1^{-1} b_i (1 + b_1^{-1} b_i)^{-1}]_p \text{ for each } i \in [2,3]\} \subset G_3'.$$

Since $[-b_1^{-1} b_i (1 + b_1^{-1} b_i)^{-1}]_p \neq 0$ for each $i \in [2,3]$, it follows by Proposition 2.2.1 that there exists at most one element $g_i \in G_0$ with

$$b_{e_i}(g_i, G_0) = [-b_1^{-1} b_i (1 + b_1^{-1} b_i)^{-1}]_p$$

for each $i \in [2,3]$. This implies $|G_2^{(2)}| \geq |G_2^{(1)}| - 2$.

Using for instance Lemma 3.2, we observe that $\mathsf{n}(e_1, G_1') = 3$ and thus $e_1 \in G_3'$. Since $e_1 \notin G_2$, we infer

$$|G_3'| \geq |G_2^{(2)}| + 1 \geq |G_2^{(1)}| - 1,$$

which proves the assertion and finishes the proof of the lemma. ∎

We now investigate the case of arbitrary rank. We start with a structural result.

Proposition 4.4: *Let $G$ be an elementary $p$-group and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) = 3$. Then there exist three subgroups $G', G''$ and $H$ of $G$ satisfying $G = G' \oplus G'' \oplus H$ and $\mathsf{r}(G') = 3$, such that*

$$G_0 \subset (G' \oplus G'') \cup H \quad \textit{and} \quad |G_0 \setminus (G' \cup H)| \leq \mathsf{r}(G'')\frac{p-1}{2}.$$

*Proof:* Since $p - 1 \geq N(G_0) = 3$, we may assume $p \geq 5$ and in particular $p$ odd.

Let $G_1 \subset G_0$ be a basis and $g \in G_0$ such that $\mathsf{n}(g, G_1) = 3$. Further, let $H_1 \subset G_1$ with $|H_1| = 3$ and $\mathsf{n}(g, H_1, G_1) = 3$. In other words, $H_1$ is the subset of those basis elements for which the coordinates of $g$ are non-zero. We set $G' = \langle H_1 \rangle$ and observe immediately that $\mathsf{r}(G') = 3$.

We define $H_2 \subset G_1 \setminus H_1$ as the set of those $e \in G_1 \setminus H_1$ for which there exists some $h' \in G_0$ with $\mathsf{n}(h', H_1, G_1) \neq 0$ and $\mathsf{n}(h', \{e\}, G_1) \neq 0$. We set

$$G'' = \langle H_2 \rangle \quad \text{and} \quad H = \langle G_1 \setminus (H_1 \cup H_2) \rangle$$

so that $G = G' \oplus G'' \oplus H$.

If $H_2 = \emptyset$, then $G_0 \subset G' \cup H$ and the result follows. Thus we assume that $H_2 \neq \emptyset$.

Let $e \in H_2$ and $h \in G_0 \setminus \{e\}$ with $\mathsf{n}(h, \{e\}, G_1) \neq 0$. We assert that

$$\mathsf{n}(h, H_1, G_1) = 2.$$

First we consider an element $h' \in G_0 \setminus \{e\}$ with

$$\mathsf{n}(h', H_1, G_1) \neq 0 \quad \text{and} \quad \mathsf{n}(h', \{e\}, G_1) \neq 0.$$

By definition of $H_2$, such an element exists. We need to show that $\mathsf{n}(h', H_1, G_1) = 2$. Indeed, assume to the contrary $\mathsf{n}(h', H_1, G_1) = 1$ and let $e'$ denote the element in $H_1$ for which $h'$ (and clearly $g$) has a non-zero coordinate. We apply Lemma 3.2 and obtain, with $G_1' = (G_1 \setminus \{e'\}) \cup \{g\}$,

$$\mathsf{n}(h', G_1') \geq \mathsf{n}(h', G_1) + \mathsf{n}(g, G_1) - 2 + 1 \geq 4,$$

a contradiction to $\mathsf{N}(G_0) = 3$. Consequently, $\mathsf{n}(h', H_1, G_1) = 2$, as announced. Consider now an arbitrary $h \in G_0 \setminus \{e\}$ with $\mathsf{n}(h, \{e\}, G_1) \neq 0$. Since $G_0$ is half-factorial, we have $\mathsf{n}(h, G_1) > 1$ (otherwise $h$ would be collinear to $e$). Again by Lemma 3.2 (with the basis $(G_1 \setminus \{e\}) \cup \{h'\}$), we get that $e$ cannot be the only basis element for which both $h$ and $h'$ have a non-zero coordinate. Since all basis elements except $e$, for which $h'$ has a non-zero coordinate are elements of $H_1$, it follows that there exists some element in $H_1$ for which $h$ has a non-zero coordinate, i.e. $\mathsf{n}(h, H_1, G_1) \neq 0$. Thus the reasoning made for $h'$ applies to $h$ as well and we obtain $\mathsf{n}(h, H_1, G_1) = 2$.

This assertion implies that if $h \in G_0 \setminus \{e\}$ with $\mathsf{n}(h, \{e\}, G_1) \neq 0$ for some $e \in H_2$, then $\mathsf{n}(h, G_1) = 3$ and $h \in G' \oplus G''$. Thus it follows from the definition of $H_2$ that $G_0 \subset (G' \oplus G'') \cup H$ and it remains to consider $|G_0 \setminus (G' \cup H)|$.

For $e \in H_2$ let $G_e = \{h \in G_0 |\ \mathsf{n}(h, \{e\}, G_1) \neq 0\}$. Since

$$G_0 \setminus (G' \cup H) = \bigcup_{e \in H_2} G_e,$$

it suffices to show that $|G_e| \le (p-1)/2$ for each $e \in H_2$. Since $\mathsf{n}(h, G_1) = 3$ for every $h \in G_e \setminus \{e\}$, we have by Proposition 3.4.1, for $j = 3$ and $G_3$ as defined there, that

$$1 + 2(|G_e| - 1) = \sum_{g' \in G_0} \mathsf{n}(g', \{e\}, G_1) + \sum_{g' \in G_3} \mathsf{n}(g', \{e\}, G_1) \le p - 1.$$

The statement follows, since $|G_e|$ is an integer and $p$ is odd. ∎

From this proposition, we now derive a result on the cardinality of half-factorial sets with $\mathsf{N}(G_0) = 3$.

Corollary 4.5: *Let $G$ be an elementary $p$-group of rank $r$ and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) = 3$. Then,*

$$|G_0| \le 3 + \Big(\frac{r}{2} - \frac{2}{3}\Big)p.$$

*Proof:* Let us apply the previous proposition (we keep the same notations). We have

$$\begin{aligned} |G_0| &= |G_0 \setminus (G' \cup H)| + |G_0 \cap (G' \cup H)| \\ &\le |G_0 \setminus (G' \cup H)| + |G_0 \cap G'| + |G_0 \cap H| - 1, \end{aligned}$$

since $G' \cap H = \{0\}$. Proposition 4.4 therefore implies

$$\begin{aligned} |G_0| &\le \mathsf{r}(G'')\frac{p-1}{2} + |G_0 \cap G'| + |G_0 \cap H| - 1 \\ &\le \mathsf{r}(G'')\frac{p-1}{2} + \Big(3 + \frac{5p}{6}\Big) + \mu(H) - 1, \end{aligned}$$

where the second line follows by Lemma 4.3, since $\mathsf{N}(G_0 \cap G') = 3$.

Now since, by (†), $\mu(H) \le 1 + \mathsf{r}(H)p/2$ and $\mathsf{r}(G'') + \mathsf{r}(H) = r - 3$, we get

$$|G_0| \le (r-3)\frac{p}{2} + 3 + \frac{5p}{6} - \frac{\mathsf{r}(G'')}{2} \le 3 + \Big(\frac{r}{2} - \frac{2}{3}\Big)p. \qquad ∎$$

In the following proposition we investigate half-factorial sets with $\mathsf{N}(G_0) \ge 4$.

PROPOSITION 4.6: *Let $G$ be an elementary $p$-group of rank $r$ and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) = s \geq 4$. Then,*

$$|G_0| \leq 1 + \frac{r}{2}p - \frac{3s}{10+2s}(p-1).$$

*Proof:* We may assume $0 \in G_0$ (otherwise, we may consider the half-factorial set $G_0 \cup \{0\}$). Let $G_1 \subset G_0$ be a basis and $g \in G_0$ such that $\mathsf{n}(g, G_1) = s$ and let $H_1 \subset G_1$ with $|H_1| = s$ such that $\mathsf{n}(g, H_1, G_1) = s$. Further, let

$$G_2 = \{g' \in G_0 |\, \mathsf{n}(g', G_1) = 2\} \quad \text{and} \quad G_{\geq 2} = \{g' \in G_0 |\, \mathsf{n}(g', G_1) \geq 2\}.$$

By Proposition 4.1.1 we have $|G_{\geq 2}| = |G_0| - |G_1| - 1 = |G_0| - (r+1)$ and by Proposition 3.4.3 we have

$$5|G_{\geq 2}| - 3|G_2| \leq r(p-2).$$

We note that $G_2 \subset \langle H_1 \rangle \cup \langle G_1 \setminus H_1 \rangle$. Otherwise, there exists some $h \in G_2$ with $\mathsf{n}(h, H_1, G_1) = 1$. Then by Lemma 3.2, considering the basis

$$G_1' = (G_1 \setminus \{e'\}) \cup \{g\}$$

where $e' \in H_1$ is the element for which the coordinate of $h$ is non-zero, we obtain

$$\mathsf{n}(h, G_1') \geq \mathsf{n}(h, G_1) + \mathsf{n}(g, G_1) - 2 + 1 = s + 1,$$

which contradicts $\mathsf{N}(G_0) = s$.

By (†) (or Proposition 3.4.1) we have

$$|G_2 \cap \langle G_1 \setminus H_1 \rangle| \leq 1 + \frac{r-s}{2}p - (1 + (r-s)) = \frac{r-s}{2}(p-2)$$

and thus, using $G_2 \subset \langle H_1 \rangle \cup \langle G_1 \setminus H_1 \rangle$,

$$|G_2 \cap \langle H_1 \rangle| \geq |G_2| - \frac{r-s}{2}(p-2).$$

Therefore there exists some $e \in H_1$ such that for

$$G_e = \{h \in G_2 |\, b_e(h, G_1) \neq 0\}$$

we have

$$|G_e| \geq \frac{2}{s}|G_2 \cap \langle H_1 \rangle| \geq \frac{2}{s}\Big(|G_2| - \frac{r-s}{2}(p-2)\Big)$$

and therefore

$$\begin{aligned}|G_e| &\geq \frac{2}{s}\Big(\frac{5|G_{\geq 2}| - r(p-2)}{3} - \frac{r-s}{2}(p-2)\Big)\\ &= \frac{10|G_{\geq 2}| - (5r-3s)(p-2)}{3s}.\end{aligned}$$

Let $G_1^* = (G_1 \setminus \{e\}) \cup \{g\}$. We set

$$G_2^* = \{g' \in G_0 \mid \mathsf{n}(g', G_1^*) = 2\} \quad \text{and} \quad G_{\geq 2}^* = \{g' \in G_0 \mid \mathsf{n}(g', G_1^*) \geq 2\}.$$

Since, by Lemma 3.2, $\mathsf{n}(e, G_1^*) = s$ and $\mathsf{n}(h, G_1^*) \geq s - 1$ for each $h \in G_e$, we have

$$|G_{\geq 2}^*| - |G_2^*| \geq |G_e| + 1$$

and again by Proposition 3.4

$$5|G_{\geq 2}^*| - 3|G_2^*| \leq r(p-2).$$

Since $|G_{\geq 2}^*| = |G_{\geq 2}|$, the three last inequalities give

$$\begin{aligned} r(p-2) &\geq 2|G_{\geq 2}^*| + 3(|G_{\geq 2}^*| - |G_2^*|) \\ &\geq 2|G_{\geq 2}| + 3(|G_e| + 1) \\ &\geq 2|G_{\geq 2}| + \frac{10|G_{\geq 2}| - (5r - 3s)(p-2)}{s} + 3 \\ &= \left(2 + \frac{10}{s}\right)|G_{\geq 2}| + 3 + \left(\frac{3s - 5r}{s}\right)(p-2). \end{aligned}$$

This implies

$$|G_{\geq 2}| \leq \frac{r}{2}(p-2) - \frac{3s}{10 + 2s}(p-1)$$

and, finally,

$$|G_0| \leq |G_{\geq 2}| + r + 1 \leq 1 + \frac{r}{2}p - \frac{3s}{10 + 2s}(p-1). \qquad \blacksquare$$

*Remark 4.7:* Let all notations be as in Proposition 4.6. It is immediate from the proof of the proposition that $\mathsf{n}(g', G_1^*) \in \{s-1, s\}$ for every $g' \in G_e$. Thus we could apply Proposition 3.4.2 with $j = s - 1$ and $j = s$ to obtain another bound for $|G_0|$, namely

$$|G_0| \leq 1 + \frac{r}{2}p - \frac{3s^2 - 7s}{22s - 42}(p-2).$$

This bound is better for $s > 7$, but for $s = 4$ it is not sufficiently good for our purpose, namely the proofs of our main results.

Since the alternative bound would not improve our main results, not even the lower bound for $c_0(p, r)$ that we obtain, and since we believe that even the alternative bound, for large $s$, is far from best possible, we favoured the shorter argument.

We are now able to deduce the needed corollary.

COROLLARY 4.8: *Let $G$ be an elementary $p$-group of rank $r$ and $G_0 \subset G$ a half-factorial, generating set with $\mathsf{N}(G_0) \geq 4$. Then,*

$$|G_0| \leq \frac{5}{3} + \Big(\frac{r}{2} - \frac{2}{3}\Big)p.$$

*Proof:* If $\mathsf{N}(G_0) = s \geq 4$, then Proposition 4.6 gives

$$|G_0| \leq 1 + \frac{r}{2}p + \Big(\frac{15}{10+2s} - \frac{3}{2}\Big)(p-1) \leq 1 + \frac{r}{2}p - \frac{2}{3}(p-1),$$

and the result follows. ∎

## 5. Proofs of the main results

In this section we give the proofs of our two main results.

*Proof of Theorem 1.1:* For even $r$ the statement was proved in [17, Theorem 8]. Thus suppose $r$ is odd. It is known (this is (†) and cf. also the discussion following Proposition 4.1) that $\mu(G) \geq 2 + (r-1)p/2$, thus it suffices to prove

$$\mu(G) \leq 2 + \frac{r-1}{2}p.$$

By Proposition 4.1 the result holds for $p \leq 3$ and we may therefore suppose $p \geq 5$.

Let $G_0 \subset G$ be half-factorial with cardinality $\mu(G)$ and in particular $\langle G_0 \rangle = G$. We distinguish 3 cases:

1. $\mathsf{N}(G_0) \leq 2$. We have $|G_0| \leq 2 + (r-1)p/2$ by Corollary 4.2.

2. $\mathsf{N}(G_0) = 3$. By Proposition 2.1, $p$ cannot be equal to 5 (cf. the discussion following Definition 3.1) and for $p \geq 7$ we have by Corollary 4.5

$$|G_0| \leq 3 + \Big(\frac{r}{2} - \frac{2}{3}\Big)p < 2 + \frac{r-1}{2}p.$$

3. $\mathsf{N}(G_0) \geq 4$. We have by Corollary 4.8

$$|G_0| \leq \frac{5}{3} + \Big(\frac{r}{2} - \frac{2}{3}\Big)p < 2 + \frac{r-1}{2}p.$$

Thus, in any case, we have $|G_0| \leq 2 + (r-1)p/2$. ∎

*Proof of Theorem 1.2:* By Proposition 4.1 (and the discussion following it) the result holds for $p \leq 3$ and we suppose $p \geq 5$.

Let $G_0 \subset G$ half-factorial. Without restriction we assume that $G_0$ is a generating set. By Theorem 1.1 we have $\mu(G) = 1 + rp/2$ if $r$ is even, respectively $\mu(G) = 2 + (r-1)p/2$ if $r$ is odd. We will establish a lower bound for $\mu(G) - |G_0|$ in case $\mathsf{N}(G_0) > 2$. This will imply the results, since the structure of a half-factorial set with $\mathsf{N}(G_0) \leq 2$ is known by Proposition 4.1.

Suppose $\mathsf{N}(G_0) > 2$. If $\mathsf{N}(G_0) = 3$, then again by Corollary 4.5

$$|G_0| \leq 3 + \Big(\frac{r}{2} - \frac{2}{3}\Big)p$$

and thus $\mu(G) - |G_0| \geq p/6 - 1$, which clearly is positive for $p \geq 7$ and again for $p = 5$ this case cannot occur.

If $\mathsf{N}(G_0) \geq 4$, then again by Corollary 4.8

$$|G_0| \leq \frac{5}{3} + \Big(\frac{r}{2} - \frac{2}{3}\Big)p$$

and consequently $\mu(G) - |G_0| \geq (p-2)/6$.

Thus if $G_0 \subset G$ is half-factorial with $|G_0| = \mu(G)$ it follows that $\mathsf{N}(G_0) \leq 2$ and the 'in particular'-statement follows by Proposition 4.1.

Let $c = 1/12$. Note that we may assume $p \geq 13$, since $|G_0| \geq \mu(G) - cp$ is equivalent to $\mu(G) = |G_0|$ for $p \leq 11$. By the argument given above we know that if $\mathsf{N}(G_0) > 2$, then

$$\mu(G) - |G_0| \geq \frac{p}{6} - 1 > cp.$$

Thus if $|G_0| \geq \mu(G) - cp$, then $\mathsf{N}(G_0) \leq 2$ and the result follows again by Proposition 4.1. ∎

*Remark 5.1:* The proof of Theorem 1.2 gives immediately that the lower bound for $c_0$ could be improved to $1/6 - \varepsilon$ ($\varepsilon > 0$ arbitrarily small), when we restrict our considerations to sufficiently large primes. Moreover, for groups of even rank we can further improve this constant to $2/3 - \varepsilon$.

It might be interesting to note that, seemingly, the most difficult problem towards further improvements on the lower bound we have for $c_0$, in general, is an improvement of Lemma 4.3. It follows immediately from the fact that in general there exist half-factorial sets with $\mathsf{N}(G_0) = 3$, which can be seen easily at least for $p \equiv 1 \bmod 3$ (cf. [31, Lemma 4.9]), that the statement of Theorem 1.2 cannot hold for any $c$ greater than or equal to 1.

## References

[1] D. D. Anderson (ed.), *Factorization in Integral Domains*, Volume 189 of Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, New York, 1997.

[2] D. D. Anderson and D. F. Anderson, *Elasticity of factorizations in integral domains. II*, Houston Journal of Mathematics **20** (1994), 1–15.

[3] D. F. Anderson, S. T. Chapman and W. W. Smith, *On Krull half-factorial domains with infinite cyclic divisor class group*, Houston Journal of Mathematics **20** (1994), 561–570.

[4] L. Carlitz, *A characterization of algebraic number fields with class number two*, Proceedings of the American Mathematical Society **11** (1960), 391–392.

[5] S. T. Chapman and J. Coykendall, *Half-factorial domains, a survey*, in *Non-Noetherian Commutative Ring Theory*, Volume 520 of Mathematics and its Applications, Kluwer Academic, Dordrecht, 2000, pp. 97–115.

[6] S. T. Chapman and A. Geroldinger, *Krull domains and monoids, their sets of lengths and associated combinatorial problems*, in *Factorization in Integral Domains*, Lecture Notes in Pure and Applied Mathematics, Vol. 189, Dekker, New York, 1997, pp. 73–112.

[7] S. T. Chapman and W. W. Smith, *Factorization in Dedekind domains with finite class group*, Israel Journal of Mathematics **71** (1990), 65–95.

[8] L. Claborn, *Every abelian group is a class group*, Pacific Journal of Mathematics **18** (1966), 219–222.

[9] R. Fossum, *The Divisor Class Group of a Krull Domain*, Volume 74 of Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, New York–Heidelberg, 1973.

[10] W. Gao and A. Geroldinger, *Half-factorial domains and half-factorial subsets in abelian groups*, Houston Journal of Mathematics **24** (1998), 593–611.

[11] W. Gao and A. Geroldinger, *Systems of sets of lengths. II*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **70** (2000), 31–49.

[12] A. Geroldinger, *Über nicht-eindeutige Zerlegungen in irreduzible Elemente*, Mathematische Zeitschrift **197** (1988), 505–529.

[13] A. Geroldinger, *Ein quantitatives Resultat über Faktorisierungen verschiedener Länge in algebraischen Zahlkörpern*, Mathematische Zeitschrift **205** (1990), 159–162.

[14] A. Geroldinger, *Systeme von Längenmengen*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **60** (1990), 115–130.

[15] A. Geroldinger and R. Göbel, *Half-factorial subsets in infinite abelian groups*, Houston Journal of Mathematics **29** (2003), 841–858.

[16] A. Geroldinger, F. Halter-Koch and J. Kaczorowski, *Non-unique factorization in orders of global fields,* Journal für die reine und angewandte Mathematik **459** (1995), 89–118.

[17] A. Geroldinger and J. Kaczorowski, *Analytic and arithmetic theory of semigroups with divisor theory,* Séminaire de Théorie des Nombres de Bordeaux (2) **4** (1992), 199–238.

[18] A. P. Grams, *The distribution of prime ideals of a Dedekind domain,* Bulletin of the Australian Mathematical Society **11** (1974), 429–441.

[19] F. Halter-Koch, *Halbgruppen mit Divisorentheorie,* Expositiones Mathematicae **8** (1990), 27–66.

[20] F. Halter-Koch, *Finitely generated monoids, finitely primary monoids and factorization properties of integral domains,* in *Factorization in Integral Domains,* Lecture Notes in Pure and Applied Mathematics, Vol. 189, Dekker, New York, 1997, pp. 31–72.

[21] F. Halter-Koch, *Ideal systems,* Volume 211 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker, New York, 1998.

[22] W. Hassler, *A note on half-factorial subsets of finite cyclic groups,* Far East Journal of Mathematical Sciences **10** (2003), 187–197.

[23] J. Kaczorowski, *Some remarks on factorizations in algebraic number fields,* Acta Arithmetica **43** (1983), 53–68.

[24] F. Kainrath, *On local half-factorial orders,* in *Arithmetical Properties of Commutative Rings and Monoids,* Lecture Notes in Pure and Applied Mathematics, Dekker, New York, to appear.

[25] D. Michel and J. Steffan, *Répartition des idéaux premiers parmi les classes d'idéaux dans un anneau de Dedekind et équidécomposition,* Journal of Algebra **98** (1986), 82–94.

[26] W. Narkiewicz, *On algebraic number fields with non-unique factorization,* Colloquium Mathematicum **12** (1964), 59–68.

[27] W. Narkiewicz, *Finite abelian groups and factorization problems,* Colloquium Mathematicum **17** (1979), 319–330.

[28] M. Radziejewski, *Oscillations of error terms associated with certain arithmetical functions,* Monatshefte für Mathematik, to appear.

[29] M. Radziejewski, *On the distribution of algebraic numbers with prescribed factorization properties,* Acta Arithmetica, to appear.

[30] M. Radziejewski and W. A. Schmid, *On the asymptotic behavior of some counting functions,* submitted.

[31] W. A. Schmid, *Half-factorial sets in elementary p-groups,* Far East Journal of Mathematical Sciences, to appear.

[32] W. A. Schmid, *On the asymptotic behavior of some counting functions, II,* submitted.

[33] L. Skula, *On c-semigroups,* Acta Arithmetica **31** (1976), 247–257.

[34] J. Śliwa, *Factorizations of distinct lengths in algebraic number fields,* Acta Arithmetica **31** (1976), 399–417.

[35] J. Śliwa, *Remarks on factorizations in algebraic number fields,* Colloquium Mathematicum **46** (1982), 123–130.

[36] A. Zaks, *Half factorial domains,* Bulletin of the American Mathematical Society **82** (1976), 721–723.

[37] A. Zaks, *Half-factorial-domains,* Israel Journal of Mathematics **37** (1980), 281–302.