Manila Journal of Science 17 (2) 2025, pp. 15-35

Word Length and Automorphism-Based Element Classification in the Discrete and Finite Heisenberg Group

Rodman F. Manalang[1,3,*] and Melvin A. Vidar[2]

[1]Mathematics and Statistics Department, College of Arts and Sciences, University of the East, Manila, Philippines
[2]Office of Curriculum Development and Instruction, University of the East, Manila, Philippines
[3]Mathematics Department, De La Salle University, Manila, Philippines

[*]Email: rodman.manalang@ue.edu.ph

ABSTRACT

The Heisenberg group over $\mathbb{Z}$, $H(\mathbb{Z})$, is the set of all 3×3 upper triangular matrices with integer entries above the diagonal and ones on the diagonal together with matrix multiplication. For a positive integer $n \geq 2$, its finite counterpart, H_n, is defined similarly with entries in $\mathbb{Z}_n$ under modular multiplication. Equivalently, these groups can be described as sets of triples $(a, b; c)$ with a nonabelian group operation involving the standard commutator structure.

This paper investigates the word length, the minimal number of generators required to express a group element in $H(\mathbb{Z})$ and H_p, where p is prime, relative to their standard generating sets. We define and analyze two specific automorphisms, σ and φ, that preserve word length and enable us to limit our analysis to a smaller representative subset of group elements without loss of generality. These automorphisms allow for more efficient classification and the derivation of explicit formulas for computing word length. We use these results to categorize elements in H_p into three types based on their algebraic structure and their behavior under these automorphisms. In addition, we explore identities related to word length.

The results reveal new insights into the combinatorial and algebraic structure of Heisenberg groups, with implications for computational group theory, representation theory, and cryptography. Our approach offers a novel perspective that bridges structural group theory with algorithmic word computation.

Keywords: word length, automorphisms, discrete Heisenberg group, finite Heisenberg group, group theory, combinatorial algebra

INTRODUCTION

The Heisenberg group, a classical object rooted in quantum mechanics, has evolved into a central figure in modern mathematics, particularly in group theory, harmonic analysis, and combinatorial geometry. The discrete and finite Heisenberg groups serve as canonical examples of nilpotent, nonabelian, and polynomially growing groups (de la Harpe, 2000). Due to their rich algebraic structure and well defined geometric properties, these groups have been investigated through various lenses. Studies by Blachère (2003) and Duchin and Mooney (2014) examined the Heisenberg group using sub-Riemannian geometry, especially through the Carnot–Carathéodory metric, focusing on distance metrics and asymptotic behaviors. Later, Duchin and Shapiro (2014) advanced this geometric perspective by investigating the rational growth of the group. Alekseev and Magdiev (2019) added a formal language viewpoint by describing the language of geodesics with respect to standard generating sets. These investigations, while foundational, rely primarily on geometric, topological, or formal-theoretic methods.

However, an important research gap remains in the literature: prior works have not thoroughly explored word length analysis from an algebraic and combinatorial perspective, particularly one that uses automorphisms of the group that preserve word length. No existing study, to the best of our knowledge, has utilized such automorphisms to reduce the computational domain or to classify elements of the discrete or finite Heisenberg group based on word length properties. Moreover, earlier approaches did not provide explicit formulas for word lengths using elementary group operations or swapping techniques among standard generators. While studies such as those by Long and Wu (2017) and Ozawa and Rieffel (2005) contextualize the Heisenberg group in advanced analytic frameworks, specifically in terms of C*-algebras and compact quantum metric spaces, they do not address symbolic or combinatorial computation of word lengths in the group setting.

This paper addresses that aforementioned research gap by constructing and analyzing two automorphisms of the discrete and finite Heisenberg groups that are shown to preserve word length. These automorphisms facilitate partitioning group elements into equivalence classes, allowing the study of word length to be confined to a smaller, representative domain. Furthermore, we derive combinatorial formulas for word length; classify group elements into Type-I, Type-II, and Type-III; and demonstrate the word-length-reducing effects of automorphic mappings. Unlike previous research, our approach is constructive, algorithmic, and grounded in elementary group-theoretic operations. In doing so, we not only provide a fresh perspective on Heisenberg group theory but also contribute tools that have potential applications in computational group theory, cryptography, and representation theory.

THE DISCRETE AND FINITE HEISENBERG GROUP

In this section, we define the discrete and finite Heisenberg groups and discuss some of their fundamental properties.

Definition 2.1. The *discrete Heisenberg group*, $H(\mathbb{Z})$, is the set of matrices

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},$$

where $a, b, c \in \mathbb{Z}$, together with matrix multiplication.

Definition 2.2. Let $n \in \mathbb{Z}, n \geq 2$. The *finite Heisenberg group*, H_n, is the set

$$\left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \middle| \, a, b, c \in \mathbb{Z}_n \right\}$$

together with matrix multiplication reduced modulo n.

Let $A = \begin{pmatrix} 1 & a_1 & c_1 \\ 0 & 1 & b_1 \\ 0 & 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & a_2 & c_2 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{pmatrix}$ where $a_{i's}$, $b_{i's}$, $c_{i's} \in \mathbb{Z}$. Then,

$$AB = \begin{pmatrix} 1 & a_1 + a_2 & c_1 + c_2 + a_1 b_2 \\ 0 & 1 & b_1 + b_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

From this, one finds that the *identity* element, denoted by e, is

$$e = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.1)$$

Moreover, the *inverse* of A, denoted by A^{-1}, is

$$A^{-1} = \begin{pmatrix} 1 & -a_1 & -c_1 + a_1 b_1 \\ 0 & 1 & -b_1 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.2)$$

Note that multiplication of two matrices in $H(\mathbb{Z})$ only modifies the entries above the diagonal. Hence, we have these alternative definitions.

Definition 2.3. The *discrete Heisenberg group*, $H(\mathbb{Z})$, is the set of triples $\{(a, b; c) | a, b, c \in \mathbb{Z}\}$ under the group operation

$$(a_1, b_1; c_1)(a_2, b_2; c_2) = (a_1 + a_2, b_1 + b_2; c_1 + c_2 + a_1 b_2). \quad (2.3)$$

This operation is well defined and corresponds to matrix multiplication of upper triangular 3×3 matrices with integer entries and ones on the diagonal.

Definition 2.4. Let $n \in \mathbb{Z}, n \geq 2$. The *finite Heisenberg group*, H_n, is the set of triples $\{(a, b; c) | a, b, c \in \mathbb{Z}_n\}$ under the group operation

$$(a_1, b_1; c_1)(a_2, b_2; c_2) = (a_1 + a_2, b_1 + b_2;\ c_1 + c_2 + a_1 b_2)$$

where addition and multiplication are reduced modulo n.

This operation is well defined, and the resulting group is a finite quotient of $H(\mathbb{Z})$.

With reference to Equation 2.1, we find $e = (0,0; 0)$, and if $A = (a_1, b_1; c_1)$, we find from Equation 2.2 that

$$A^{-1} = (-a_1, -b_1; -c_1 + a_1 b_1).$$

It is known (Blachère, 2003; Osipov, 2015; Pate, 2006) that $H(\mathbb{Z})$ have the following generators

$X = (1,0; 0)$, $Y = (0,1; 0)$ and $Z = (0,0; 1)$ (2.4)

that satisfy the following fundamental commutative relations:

$$XZ = ZX, YZ = ZY \text{ and } XY = YXZ. \quad (2.5)$$

Moreover, if $A = (a, b; c)$, then

$$A = Y^b X^a Z^c. \quad (2.6)$$

Lemma 2.1. (Kahn, 2005, p. 2) *Let α be an integer and $g = (a, b; c) \in H(\mathbb{Z})$ (or H_n). Then,*

$$g^\alpha = \left(\alpha a, \alpha b; \alpha c + ab\frac{\alpha(\alpha - 1)}{2}\right).$$

Proof. We now prove that for any integer α, the power of a group element $g = (a, b; c) \in H(\mathbb{Z})$ (or H_n) satisfies $g^\alpha = \left(\alpha a, \alpha b; \alpha c + ab\frac{\alpha(\alpha-1)}{2}\right)$. Let $g = (a, b; c) \in H(\mathbb{Z})$ (or H_n).

The result is straightforward for $\alpha = 0$, $g^0 = (0,0; 0)$; the identity element of the group and the right-hand side also evaluate the same, that is,

$$\left(0,0; 0 + ab\frac{0(0 - 1)}{2}\right) = (0,0; 0),$$

so the lemma holds for $\alpha = 0$.

We now proceed by induction for positive integers and then extend to negative integers. For $\alpha = 1$, clearly, the formula holds. We have $g^1 = g = (a, b; c)$. Assume it holds for some $\alpha = k$, that is,

$$g^k = \left(ka, kb; kc + ab\frac{k(k - 1)}{2}\right).$$

Then, using the group operation (Equation 2.3), we compute

$$\begin{aligned} g^{k+1} &= g^k \cdot g \\ &= \left(ka, kb; kc + ab\frac{k(k-1)}{2}\right)(a, b; c) \\ &= \left(ka + a, kb + b;\ kc + ab\frac{k(k-1)}{2} + c + ka \cdot b\right) \\ &= \left((k+1)a, (k+1)b;\ (k+1)c + ab\left(\frac{k(k-1)}{2} + k\right)\right) \\ &= \left((k+1)a, (k+1)b;\ (k+1)c + ab\left(\frac{k(k-1) + 2k}{2}\right)\right) \\ &= \left((k+1)a, (k+1)b;\ (k+1)c + ab\left(\frac{k(k+1)}{2}\right)\right). \end{aligned}$$

Hence, the lemma holds for $\alpha = k+1$. Thus, the formula holds for all positive integers α.

Finally, we prove the remaining case as follows. Consider the identity

$$g^{-\alpha} = (g^{-1})^{\alpha} = (-a, -b; -c + ab)^{\alpha}.$$

Since α is positive, we can apply the above formula and obtain the desired result.

This matches the original formula when $\alpha < 0$, so the result holds for negative integers as well. Therefore, the lemma also holds for all integers α, whether positive, negative, or zero. □

Lemma 2.2. *Let X, Y be the elements of $H(\mathbb{Z})$ (or H_n) in Equation 2.4. If $a, b \in Z$, then*

$$X^a Y^b = Y^b X^a Z^{ab}.$$

Proof. Let a, b be integers. Applying Lemma 2.1 to the powers of the generators, we find $X^a = (a, 0; 0)$, $Y^b = (0, b; 0)$and $Z^{ab} = (0,0; ab)$.

Now, computing their product using the group operation in $H(\mathbb{Z})$, we have

$$X^a Y^b = (a, 0; 0)(0, b; 0) = (a, b; ab).$$

For the other side of the equation,

$$Y^b X^a = (0, b; 0)(a, 0; 0) = (a, b; 0).$$

Note that $Z^{ab} = (0,0; ab)$; thus,

$$Y^b X^a Z^{ab} = (a, b; 0)(0,0; ab) = (a, b; ab).$$

Hence, the identity holds for all integers a and b. □

The next two lemmas will be used in the development of several results on word length, including Theorem 4.4, Theorem 4.5, Lemma 5.4, and Theorem 6.2.

Lemma 2.3. *Let X and Y be the generators in Equation 2.4. Let n be a positive integer and let $a_i, b_i \in \mathbb{Z}$ $(1 \le i \le n)$. Then,*

$$X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_n} Y^{b_n} = \left(\sum_{i=1}^{n} a_i , \sum_{j=1}^{n} b_j ; \sum_{\substack{i,j=1 \\ i \le j}}^{n} a_i b_j \right).$$

Proof. We prove the result by induction on n.

For the base case $n = 1$, we compute $X^{a_1} Y^{b_1} = (a_1, 0; 0)(0, b_1; 0) = (a_1, b_1; a_1 b_1)$, and by Lemma 2.2, $X^{a_1} Y^{b_1} = Y^{b_1} X^{a_1} Z^{a_1 b_1}$, and the formula holds as

$$\sum_{i=1}^{1} a_i = a_1, \ \sum_{i=1}^{1} b_i = b_1,$$

and

$$\sum_{\substack{i,j=1 \\ i \le j}}^{1} a_1 b_1.$$

Assume now that the lemma holds for some $n = k$, that is,

$$X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_k} Y^{b_k} = \left(\sum_{i=1}^{k} a_i , \sum_{j=1}^{k} b_j ; \sum_{\substack{i,j=1 \\ i \le j}}^{k} a_i b_j \right).$$

We want to show that the result also holds for $n = k+1$. Using the inductive hypothesis, we have

$$X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_k} Y^{b_k} X^{a_{k+1}} Y^{b_{k+1}} = \left(\sum_{i=1}^{k} a_i , \sum_{j=1}^{k} b_j ; \sum_{\substack{i,j=1 \\ i \le j}}^{k} a_i b_j \right) (a_{k+1}, 0; 0)(0, b_{k+1}; 0).$$

Applying the group operation and using Lemma 2.2, this becomes

$$\left(\sum_{i=1}^{k} a_i , \sum_{j=1}^{k} b_j ; \sum_{\substack{i,j=1 \\ i \le j}}^{k} a_i b_j \right) (a_{k+1}, 0; 0)(0, b_{k+1}; 0)$$

$$= \left(\sum_{i=1}^{k} a_i + a_{k+1}, \sum_{j=1}^{k} b_j ; \sum_{\substack{i,j=1 \\ i \le j}}^{k} a_i b_j \right) (0, b_{k+1}; 0)$$

$$= \left(\sum_{i=1}^{k} a_i + a_{k+1}, \sum_{j=1}^{k} b_j + b_{k+1} ; \sum_{\substack{i,j=1 \\ i \le j}}^{k} a_i b_j + \left(\sum_{i=1}^{k} a_i + a_{k+1} \right) b_{k+1} \right)$$

$$= \left(\sum_{i=1}^{k} a_i + a_{k+1}, \sum_{j=1}^{k} b_j + b_{k+1} ; \sum_{\substack{i,j=1 \\ i \le j}}^{k} a_i b_j + \sum_{i=1}^{k} a_i b_{k+1} + a_{k+1} b_{k+1} \right)$$

$$= \left(\sum_{i=1}^{k} a_i + a_{k+1}, \sum_{j=1}^{k} b_j + b_{k+1} ; \sum_{\substack{i,j=1 \\ i \le j}}^{k+1} a_i b_j \right)$$

$$= \left(\sum_{i=1}^{k+1} a_i , \sum_{j=1}^{k+1} b_j ; \sum_{\substack{i,j=1 \\ i \le j}}^{k+1} a_i b_j \right).$$

Hence, the lemma holds for $n = k + 1$, and by the principle of mathematical induction, the formula holds for all positive integers n. □

Lemma 2.4. *Let X and Y be the generators in Equation 2.4. Let n be a positive integer and let $a_i, b_i \in \mathbb{Z}$ $(1 \le i \le n)$. Then,*

$$Y^{b_n}X^{a_n}Y^{b_{n-1}}X^{b_{n-1}} \dots Y^{b_2}X^{a_2}Y^{b_1}X^{a_1} = \left(\sum_{i=1}^{n} a_i, \sum_{j=1}^{n} b_j; \sum_{\substack{i,j=1 \\ i>j}}^{n} a_i b_j\right).$$

Proof. We prove the result by induction on n, similar to the approach used in Lemma 2.3.

For the base case $n = 1$, we compute $Y^{b_1}X^{a_1} = (0, b_1; 0)(a_1, 0; 0) = (a_1, b_1; 0)$, which matches the desired formula with $\sum_{i=1}^{1} a_i = a_1$, $\sum_{i=1}^{1} b_i = b_1$ and $\sum_{\substack{i,j=1 \\ i>j}}^{1} a_i b_j = 0$.
Thus, the base case holds.

Assume the statement holds for $n = k$, that is,

$$Y^{b_k}X^{a_k}Y^{b_{k-1}}X^{b_{k-1}} \dots Y^{b_2}X^{a_2}Y^{b_1}X^{a_1} = \left(\sum_{i=1}^{k} a_i, \sum_{j=1}^{k} b_j; \sum_{\substack{i,j=1 \\ i>j}}^{k} a_i b_j\right).$$

To show the formula holds for $n = k + 1$, we prepend $Y^{b_{k+1}}X^{a_{k+1}}$ to the expression $Y^{b_k}X^{a_k} \dots Y^{b_1}X^{a_1}$ to obtain

$$Y^{b_{k+1}}X^{a_{k+1}}Y^{b_k}X^{a_k}Y^{b_{k-1}}X^{b_{k-1}} \dots Y^{b_2}X^{a_2}Y^{b_1}X^{a_1}.$$

By the inductive hypothesis, the product evaluates as $(0, b_{k+1}; 0)(a_{k+1}, 0; 0)(a', b'; c')$ where $a' = \sum_{i=1}^{k} a_i$, $b' = \sum_{j=1}^{k} b_j$, and

$$c' = \sum_{\substack{i,j=1 \\ i>j}}^{k} a_i b_j.$$

Using the group operation,

$$(0, b_{k+1}; 0)(a_{k+1}, 0; 0)(a', b'; c') = (a' + a_{k+1}, b' + b_{k+1}; c' + a_{k+1}b').$$

Thus, the third coordinate becomes

$$c' + a_{k+1}b' = \sum_{\substack{i,j=1 \\ i>j}}^{k} a_i b_j + a_{k+1}\sum_{j=1}^{k} b_j = \sum_{\substack{i,j=1 \\ i>j}}^{k+1} a_i b_j.$$

Hence, the formula holds for $n = k + 1$, completing the induction. □

It is known (Osipov, 2015; Robinson, 1996; Suzuki, 1986) that the center of $H(\mathbb{Z})$ and H_n is the subgroup $\langle Z \rangle$ where Z is the element $(0,0; 1)$ in Equation 2.4. Here, $\langle Z \rangle$ denotes the cyclic subgroup generated by the element Z.

The next lemmas concern the structure of the center of $H(\mathbb{Z})$ and H_n. All results follow from Lemmas 2.3 and 2.4 and can be more compactly expressed using commutators.

Corollary 2.5. *With reference to the generators X, Y, and Z in Equation 2.4, each of the following expressions is equal to Z.*

i. $XYX^{-1}Y^{-1}$
ii. $X^{-1}Y^{-1}XY$
iii. $YX^{-1}Y^{-1}X$
iv. $Y^{-1}XYX^{-1}$

Corollary 2.6. *For any integer k, the following expressions are all equal to Z^k.*

i. $XY^kX^{-1}Y^{-k}$
ii. $X^{-1}Y^{-k}XY^k$
iii. $X^kYX^{-k}Y^{-1}$
iv. $X^{-k}Y^{-1}X^kY$
v. $YX^{-k}Y^{-1}X^k$
vi. $Y^{-1}X^kYX^{-k}$
vii. $Y^kX^{-1}Y^{-k}X$
viii. $Y^{-k}XY^kX^{-1}$

The next lemma further expresses the elements of $\langle Z \rangle$ using the standard generators.

Corollary 2.7. *For any integers s, t, the following expressions are all equal to Z^{st}.*

i. $Y^tX^{-s}Y^{-t}X^s$
ii. $Y^{-t}X^sY^tX^{-s}$
iii. $Y^sX^{-t}Y^{-s}X^t$
iv. $Y^{-s}X^tY^sX^{-t}$
v. $X^{-s}Y^{-t}X^sY^t$
vi. $X^sY^tX^{-s}Y^{-t}$
vii. $X^{-t}Y^{-s}X^tY^s$
viii. $X^tY^sX^{-t}Y^{-s}$

The next lemma will be used in enumerating the words with minimal length corresponding to the elements of the center.

Corollary 2.8. *For any integers s,t,j, the following expressions are all equal to Z^{st}.*

i. $Y^{t-j}X^{-s}Y^{-t}X^{s}Y^{j}$, $Y^{-t+j}X^{s}Y^{t}X^{-s}Y^{-j}$, $X^{j}Y^{s}X^{-t}Y^{-s}X^{t-j}$, $X^{-j}Y^{-s}X^{t}Y^{s}X^{-t+j}$, $Y^{t}X^{-s}Y^{-t}X^{s}Y^{t-j}$, $Y^{-j}X^{s}Y^{t}X^{-s}Y^{-t+j}$, $X^{-t+j}Y^{-s}X^{t}Y^{s}X^{-j}$, $X^{t-j}Y^{s}X^{-t}Y^{-s}X^{j}$ $(0 \le j \le t)$.

ii. $X^{j}Y^{t}X^{-s}Y^{-t}X^{s-j}$, $X^{-j}Y^{-t+j}X^{s}Y^{t}X^{-s+j}$, $Y^{s-j}X^{-t}Y^{-s}X^{t}Y^{j}$, $Y^{-s+j}X^{t}Y^{s}X^{-t}Y^{-j}$, $X^{-s+j}Y^{-t}X^{s}Y^{t}X^{-j}$, $X^{s-j}Y^{t}X^{-s}Y^{-t}X^{j}$, $Y^{j}X^{-t}Y^{-s}X^{t}Y^{s-j}$, $Y^{-j}X^{t}Y^{s}X^{-t}Y^{-s+j}$ $(0 \le j \le s)$.

Remark. The expressions in Corollary 2.7 are obtained by setting $j = 0$ in Corollary 2.8.

WORD LENGTH AND WORD-LENGTH-PRESERVING AUTOMORPHISMS

In this section, we introduce the notion of the word length of an element of a group and apply it to $H(\mathbb{Z})$ and H_p, where p is prime. Then, we discuss two word-length-preserving automorphisms on $H(\mathbb{Z})$ and H_p that help us prove some identities and relations on word length.

Definition 3.1. Let G be a group and let M be a non-empty subset of G. Then, M is a *generating set* for G if $\forall g \in G$,

$$g = m_1^{\pm 1} m_2^{\pm 1} \dots m_k^{\pm 1}$$

where $m_i \in M$, $1 \le i \le k$.

Remark. If there exists a generating set M for G such that $|M| < \infty$, then G is said to be *finitely generated*.

Definition 3.2. Let M be a generating set for G. A *word* W in M is defined as a finite sequence of elements of the form $m_1^{\pm 1} m_2^{\pm 1} \dots m_k^{\pm 1}$ where $m_i \in M$ $(1 \le i \le k)$ and $k \ge 0$.

Remark. If $k = 0$, W is called the *empty word* and $W = e$, where e is the *identity element* in G.

Definition 3.3. Let W be a word in M. The *length* of the word $W = m_1^{\pm 1} m_2^{\pm 1} \dots m_k^{\pm 1}$ in M is the value of the integer k. The word length of W, denoted by $\ell(W)$ in M, is the nonnegative integer, $\ell := \ell(W)$, defined by

$$\ell = min\,\{k \mid W = m_1^{\pm 1} m_2^{\pm 1} \dots m_k^{\pm 1} \text{ where } m_i \in M, 1 \le i \le k\}.$$

Since M is a generating set of G, every word in M has at least one such representation; thus, $\ell(W)$ is well defined.

Definition 3.4. A word W in M is called *reduced* if it contains no pair of consecutive symbols of the form mm^{-1} or $m^{-1}m, m \in M$.

Example 3.1. Let $G = H(\mathbb{Z})$. Then, by Equation 2.4, $M_1 = \{X, Y, Z\}$ is a generating set for G. Moreover, by Corollary 2.5, $M_2 = \{X, Y\}$ is also a generating set for G.

Definition 3.3 defines the word length of a word with respect to a generating set, M. Henceforth, we will fix $M = \{X, Y\}$, where X, Y are the triples in Equation 2.4. In the case of the finite Heisenberg group, we will focus our attention on H_p where p is prime.

Example 3.2. In $H(\mathbb{Z})$, using Equation 2.6 and Corollary 2.5 *(i)*,

$$\begin{aligned}(1,1;1) = YXZ &= YX(XYX^{-1}Y^{-1}) \\ &= YX^2YX^{-1}Y^{-1}.\end{aligned}$$

One also verifies that $(1,1;1) = XY$.

Since the only words of length 1 are $X^{\pm 1}$ and $Y^{\pm 1}$ (corresponding to $(\pm 1,0;0)$ and $(0,\pm 1;0)$, respectively), we find that any other nonidentity element $g \in H(\mathbb{Z})$ has a word length at least 2. It follows that XY is one of the shortest words equal to $(1,1;1)$. Thus, the word length of $(1,1;1)$ is $\ell(XY) = 2$.

Remark. We call M the *standard generating set* and X and Y the standard generators for $H(\mathbb{Z})$.

The next result establishes an interesting relationship between the word length of an element of $H(\mathbb{Z})$ or H_p and its inverse.

Theorem 3.1. *Let $g \in H(\mathbb{Z})$ or H_p. If $\ell(g) = k$, then $\ell(g^{-1}) = k$.*

Proof. Suppose that $l(g) = k$. Then, there exists $m_1, m_2, \dots m_{k-1}, m_k \in M$ and $i_1, i_2, \dots, i_{k-1}, i_k \in \{1, -1\}$ such that

$$g = m_1^{i_1} m_2^{i_2} \dots m_{k-1}^{i_{k-1}} m_k^{i_k} \quad (3.1)$$

is a reduced word and k is minimal.

Now, $g^{-1} = \left(m_1^{i_1} m_2^{i_2} \dots m_{k-1}^{i_{k-1}} m_k^{i_k}\right)^{-1}$

$$= m_k^{-i_k} m_{k-1}^{-i_{k-1}} \dots m_2^{-i_2} m_1^{-i_1}. \quad (3.2)$$

Assume, for contradiction, that $\ell(g^{-1}) < k$.

Observe that g^{-1} is a reduced word. Otherwise, if there exists a pair mm^{-1} or $m^{-1}m$ in Equation 3.2, the pair $m^{-1}m$ or mm^{-1} also appears in Equation 3.1.

Moreover, if there exists an expression for g^{-1} of word length less than k, then taking the inverse gives a reduced word for g and has a word length less than k, a contradiction to the minimality of k. □

Now, we construct two automorphisms of $H(\mathbb{Z})$ and H_p that preserve word length. These maps help us derive some word length identities that help reduce the domain in finding the word length of the elements of $H(\mathbb{Z})$ and H_p.

Theorem 3.2. *The map $\sigma: H(\mathbb{Z}) \to H(\mathbb{Z})$ (respectively $\sigma: H_p \to H_p$) given by*

$$\sigma\big((a, b; c)\big) = (b, -a; c - ab)$$

is an automorphism of $H(\mathbb{Z})$ (respectively H_p).

Proof. Let $g_1 = (a_1, b_1;\ c_1)$, $g_2 = (a_2, b_2; c_2) \in H(\mathbb{Z})$ (or H_p). We verify the σ is a homomorphism by direct computation using Equation 2.3. For $g_1 g_2 = (a_1 + a_2, b_1 + b_2; c_1 + c_2 + a_1 b_2)$, we compute,

$$\sigma(g_1 g_2) = \big(b_1 + b_2, -(a_1 + a_2); (c_1 + c_2 + a_1 b_2) - (a_1 + a_2)(b_1 + b_2)\big).$$

On the other hand, applying σ to each element and then multiplying,

$$\sigma(g_1)\sigma(g_2)$$
$$= (b_1, -a_1; c_1 - a_1 b_1)(b_2, -a_2; c_2 - a_2 b_2)$$
$$= \big(b_1 + b_2, -a_1 - a_2; (c_1 - a_1 b_1) + (c_2 - a_2 b_2) + b_1(-a_2)\big).$$

Note that $(c_1 + c_2 + a_1 b_2) - (a_1 + a_2)(b_1 + b_2) = (c_1 - a_1 b_1) + (c_2 - a_2 b_2) + b_1(-a_2)$.
So $\sigma(g_1 g_2) = \sigma(g_1)\sigma(g_2)$ and σ is a homomorphism.

Also σ is injective since,

$$Ker\ \sigma = \{g = (a, b; c) | \sigma(g) = e\}$$
$$= \{g | (b, -a; c - ab) = (0,0; 0)\}$$
$$= \{(0,0; 0)\}.$$

Moreover, σ is surjective since for a given $g = (a, b; c) \in H(\mathbb{Z})$, we have

$$\sigma\big((-b, a; c - ab)\big) = \big(a, b; (c - ab) - (-b)(a)\big)$$
$$= (a, b; c)$$
$$= g.$$

Thus, σ is established as an automorphism of $H(\mathbb{Z})$ (respectively H_p). □

Corollary 3.3. *Let X, Y, Z be the generators of $H(\mathbb{Z})$ (respectively H_p) in Equation 2.4. Then, the following hold:*

i. $\sigma(X) = Y^{-1}$,
ii. $\sigma(Y) = X$, *and*
iii. $\sigma(Z) = Z$.

Proof. Apply Theorem 3.2 to *X, Y,* and *Z* as follows:

i. $\sigma(X) = \sigma\big((1,0; 0)\big) = (0, -1; 0) = Y^{-1}$,
ii. $\sigma(Y) = \sigma\big((0,1; 0)\big) = (1,0; 0) = X$, and lastly
iii. $\sigma(Z) = \sigma\big((0,0; 1)\big) = (0,0; 1) = Z$.

Theorem 3.4. *The map σ in Theorem 3.2 preserves word length.*

Proof. Let $g = A_1^{i_1} A_2^{i_2} \dots A_k^{i_k}$ be a reduced word where $A_i \in \{X, Y\}$ and $i_1, i_2, \dots, i_k \in \{1, -1\}$, such that $\ell(g) = k$ is a word that has minimal length among all the expressions representing g.

By Theorem 3.2 and Corollary 3.3, the automorphism σ maps each generator to another single generator: $\sigma(X) = Y^{-1}$ and $\sigma(Y) = X$. Thus,

$$\sigma(g) = \sigma\left(A_1^{i_1} A_2^{i_2} \dots A_k^{i_k}\right)$$
$$= \sigma\left(A_1^{i_1}\right)\sigma\left(A_2^{i_2}\right) \dots \sigma\left(A_k^{i_k}\right)$$
$$= B_1 B_2 \dots B_k \text{ where } B_j = \begin{cases} X^{i_j}, \text{if } A_j = Y \\ Y^{-i_j}, \text{if } A_j = X \end{cases}.$$

We note that the image $\sigma(g) = B_1B_2 \dots B_k$ is an expression of length k.

Assume, for contradiction, a shorter word representing $\sigma(g)$, say $\sigma(g) = B_1'B_2' \dots B_l'$, with $l < k$ exists. Applying σ^{-1} of both sides, we find $g = \sigma^{-1}(B_1')\sigma^{-1}(B_2') \dots \sigma^{-1}(B_l')$, an expression for g of length at most l, so that $\ell(g) \leq l < k$, leads directly to a contradiction with the minimality assumption. Therefore, $\ell\big(\sigma(g)\big) = k = \ell(g)$. □

The proof of the next three results is like the proofs of the previous three results.

Theorem 3.5. *The map $\phi: H(\mathbb{Z}) \to H(\mathbb{Z})$ (respectively $\phi: H_p \to H_p$) given by*

$$\phi\big((a, b; c)\big) = (-a, b; -c)$$

is an automorphism of $H(\mathbb{Z})$ (respectively H_p).

Corollary 3.6. *Let X, Y, Z be the generators of $H(\mathbb{Z})$ (or H_p) in Equation 2.4. Then, the following hold:*

i. $\phi(X) = X^{-1}$,
ii. $\phi(Y) = Y$, *and*
iii. $\phi(Z) = Z^{-1}$.

Theorem 3.7. *The map ϕ in Theorem 3.5 preserves word length.*

The automorphisms σ and ϕ satisfy σ^4 is the identity and ϕ^2 is the identity. While they do not generate all word-length-preserving automorphisms of $H(\mathbb{Z})$, they are sufficient to produce distinct but equivalent representatives that preserve word length.

By applying these automorphisms, elements of $H(\mathbb{Z})$ can be grouped into sets of elements that are mapped into each other while preserving word length. This allows us to focus on a single representative from each such set when analyzing minimal-length expressions.

Instead of analyzing the entire set of triples $(a, b; c) \in H(\mathbb{Z})$, we may therefore restrict attention to a smaller subset—such as those satisfying $a \geq 0, b \geq 0$ or lying in a specific region of the integer grid (e.g., where both a and b are nonnegative)—without loss of generality. This is what we mean by saying that these automorphisms "reduce the domain" of the triples: they allow us to focus on a smaller, manageable set of representative elements, making it possible to study word length systematically without considering every case.

Remark. By convention, whenever we write a word W as

$$W = X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_{r-1}} Y^{b_{r-1}} X^{a_r} Y^{b_r},$$

the exponents a_1 and b_r may be zero but $a_{i+1} \neq 0$ and $b_i \neq 0$ for $1 \leq i \leq r-1$.

Definition 3.6. Let

$$W = X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_{r-1}} Y^{b_{r-1}} X^{a_r} Y^{b_r}$$

be a word of finite length. The *reverse* of W, denoted by W^R, is

$$Y^{b_r} X^{a_r} Y^{b_{r-1}} X^{a_{r-1}} \dots Y^{b_2} X^{a_2} Y^{b_1} X^{a_1}.$$

Lemma 3.8. *Let W be a word of minimal length corresponding to $g = (a, b; c) \in H(\mathbb{Z})$ (or H_p). Then, W^R is a word of minimal length corresponding to $g' = (a, b; -c + ab)$. In particular, $\ell(W) = \ell(W^R)$.*

Proof. Let $g = (a, b; c) \in H(\mathbb{Z})$ and let $W = X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_{r-1}} Y^{b_{r-1}} X^{a_r} Y^{b_r}$ be a word with minimal length representing g, that is, $g = W$.

Recall that from Theorem 3.1, any nontrivial element and its inverse have the same word length.

Moreover, the mapping σ in Theorem 3.2 preserves word length. Applying σ^2 to W and taking its inverse, we have,

$$\begin{aligned}
&\big(\sigma^2(W)\big)^{-1} \\
&= \big(\sigma^2(X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \dots X^{a_{r-1}} Y^{b_r - 1} X^{a_r} Y^{b_r})\big)^{-1} \\
&= (X^{-a_1} Y^{-b_1} X^{-a_2} Y^{-b_2} \dots X^{-a_{r-1}} Y^{-b_{r-1}} X^{-a_r} Y^{-b_r})^{-1} \\
&= Y^{b_r} X^{a_r} Y^{b_{r-1}} X^{a_{r-1}} \dots Y^{b_2} X^{a_2} Y^{b_1} X^{a_1} \\
&= W^R.
\end{aligned}$$

Doing the same to g, we find

$$\big(\sigma^2(g)\big)^{-1} = (-a, -b; c)^{-1} = (a, b; -c + ab) = g'.$$

Thus, $g' = W^R$ and $\ell(W^R) = \ell(W)$. □

Lemma 3.9. *Let* $g = (a, b; c) \in H(\mathbb{Z})$ $(or\ H_p)$. *Then,* $\ell((a, b; c)) = \ell((b, a; c))$.

Proof. Let $g = (a, b; c) \in H(\mathbb{Z})$ *or* H_p. Then,

$$\begin{aligned}\left(\phi(\sigma(g))\right)^{-1} &= \left(\phi(\sigma(b, -a; c - ab))\right)^{-1} \\ &= \left((-b, -a; -c + ab)\right)^{-1} \\ &= (b, a; c).\end{aligned}$$

Since word length is preserved by σ, ϕ, and inverses (from Theorems 3.1, 3.4, and 3.7), we have $\ell((a, b; c)) = \ell((b, a; c))$. □

WORD LENGTH IN THE DISCRETE HEISENBERG GROUP

In this section, we provide combinatorial formulas for the word length of elements in the discrete Heisenberg group $H(\mathbb{Z})$. The word-length-preserving automorphisms introduced earlier help us reduce the domain of $H(\mathbb{Z})$, allowing us to derive formulas for the word length of elements more systematically and efficiently.

We begin by reducing the domain of the discrete Heisenberg group.

Let $g = (a, b; c) \in H(\mathbb{Z})$, where $a, b, c \geq 0$. Observe that by applying σ^2, ϕ, and $\sigma^2\phi$ to g, we have the following:

$$\begin{aligned} g &= (a, b; c), \\ \sigma^2(g) &= (-a, -b; c), \\ \phi(g) &= (-a, b; -c) \text{ and} \\ \sigma^2(\phi(g)) &= (a, -b; -c). \end{aligned}$$

Similarly, if $a, b, c \geq 0$, we have

$$\begin{aligned} g' &= (a, b; -c), \\ \sigma^2(g') &= (-a, -b; -c), \\ \phi(g') &= (-a, b; c) \text{ and} \\ \sigma^2(\phi(g')) &= (a, -b; c). \end{aligned}$$

This shows that for every triple $(a, b; c)$, there exists a representative (through automorphisms) with $a \geq 0$ and $b \geq 0$. Therefore, with respect to the word length in $H(\mathbb{Z})$, we can limit the values of $a, b,$ and c to $a \geq 0$ and $b \geq 0$.

To further reduce the domain, we now show that c may also be assumed to be nonnegative. Let $a, b \geq 0$ and $c < 0$. We have

$$\sigma^2((a, b; c)^{-1}) = \sigma^2((-a, -b; -c + ab)) = (a, b; -c + ab).$$

But σ preserves word length and $\ell(g) = \ell(g^{-1})$. Thus, it follows that

$$\ell((a, b; c)) = \ell((a, b; -c + ab)) \qquad (4.1)$$

with $-c + ab > 0$. This means that if $c < 0$, there exists an element $g' = (a', b'; c')$ with $c' > 0$ such that $\ell(g) = \ell(g')$. Therefore, we may assume $c \geq 0$ as well.

Therefore, due to the symmetry of the discrete Heisenberg group $H(\mathbb{Z})$ induced by the automorphisms σ and ϕ, we can restrict analysis without loss of generality to elements with nonnegative coordinates.

It can be verified rigorously that none of the words of length equal to $1, 2,$ or 3 are equivalent to Z. However, one verifies that among the words of length 4, four expressions are equal to Z. These are the same expressions we enumerated in Corollary 2.5. Hence, $\ell(Z) = 4$.

Let $g \in H(\mathbb{Z})$ be an element of the center of $H(\mathbb{Z})$. Then, $g = Z^k$ for some $k \in \mathbb{Z}$. While there are infinitely many words that represent Z^k, we illustrate four standard forms that arise from Lemma 2.2 and Corollary 2.5 as follows:

$$\begin{aligned} Z^k &= (0,0; k) \\ &= (XYX^{-1}Y^{-1})^k = (XYX^{-1}Y^{-1})(XYX^{-1}Y^{-1}) \dots (XYX^{-1}Y^{-1})\ [k\ factors] \quad (4.2) \\ &= (Y^{-1}XYX^{-1})^k = (Y^{-1}XYX^{-1})(Y^{-1}XYX^{-1}) \dots (Y^{-1}XYX^{-1})\ [k\ factors] \quad (4.3) \\ &= (X^{-1}Y^{-1}XY)^k = (X^{-1}Y^{-1}XY)(X^{-1}Y^{-1}XY) \dots (X^{-1}Y^{-1}XY)\ [k\ factors] \quad (4.4) \\ &= (YX^{-1}Y^{-1}X)^k = (YX^{-1}Y^{-1}X)(YX^{-1}Y^{-1}X) \dots (YX^{-1}Y^{-1}X)\ [k\ factors] \quad (4.5) \end{aligned}$$

The expressions in Equations 4.2–4.5 do not represent all possible words for Z^k. Other equivalent words may be formed by inserting identities such as ZZ^{-1} or group relations involving commutators that preserve the group element.

We begin by considering four forms for representing powers of Z, based on juxtaposing the factors, say $XYX^{-1}Y^{-1}$. Letting $Z^k = (0,0; k) \in H(\mathbb{Z})$, the expression $Z^k = (XYX^{-1}Y^{-1})^k$ has length $4k$, as each factor contributes four letters.

Corollary 2.6 provides reduced expressions for powers of Z^k, constructed through careful reordering of generators and guided cancellations that significantly shorten lengths. These expressions have length $2(k+1)$, which is significantly shorter than the naive concatenation of k copies of the commutator form $XYX^{-1}Y^{-1}$, having length $4k$. For any nonzero k, we explicitly have the inequality $2(k+1) \leq 4k$, confirming explicitly that expressions from Corollary 2.6 indeed provide strictly shorter lengths compared to the naive expansion.

However, while these expressions are systematically constructed and preserve word length under the automorphisms described earlier, they are not always minimal. For instance, Lemma 2.2 shows that the expression $X^{-2}Y^{-3}X^2Y^3$, which represents Z^6, has length 10. This is shorter than the bound $2(6+1) = 14$ suggested by Corollary 2.6. Thus, while such bounds serve as useful benchmarks, they do not guarantee minimality in all cases.

This phenomenon is also observed when c is a prime number**.** Consider Z^7, which is represented by the word $X^{-2}Y^{-2}X^3Y^3X^{-1}Y^{-1}$, of length 12. This is strictly shorter than the bound $2(7+1) = 16$, again indicating that the expression in Corollary 2.6 only provides an upper bound and does not characterize the minimal length.

In general, for any positive integer k, we may express Z^k using a commutator of the form $W = X^sY^tX^{-s}Y^{-t}$, where $s, t \in \mathbb{Z}^+$ and $st = k$. This yields a word of length $2(s+t)$, providing the upper bound $\ell(Z^k) \leq 2(s+t)$. Now consider the case when k is a perfect square, say $k = m^2$. Choosing $s = t = m$ yields the expression $W = X^mY^mX^{-m}Y^{-m}$, which represents Z^{m^2} and has length $4m$. Therefore, $\ell(Z^{n^2}) \leq 4m$.

We now single out the perfect-square case, which will be useful in the discussion that follows.

Lemma 4.1. *If $k = n^2$ and n is a positive integer, then $\ell(Z^k) = 4n$.*

Proof. Let $k = n^2$, where n is a positive integer. From Corollary 2.7, we have, $Z^{n^2} = X^nY^nX^{-n}Y^{-n}$. This word contains n copies each of X, Y, X^{-1}, and Y^{-1}, and hence has length $4n$. Therefore, if $k = n^2$, then $\ell(Z^k) \leq 4n$.

Now, let $W = X^{a_1}Y^{b_1}X^{a_2}Y^{b_2} \cdots X^{a_r}Y^{b_r}$ be any reduced word representing Z^k. Let s be the total of all positive X-exponents and t the total of all positive Y-exponents. By Lemma 2.3, the exponent of Z contributed by the word is $\sum_{i \leq j} a_i b_j = n^2$.

Since at most s units of positive X can pair with at most t units of positive Y, this sum is at most st. Hence $st \geq n^2$. This pairing bound is justified rigorously in Appendix A *(see Lemma A.1).*

Assume for contradiction that $\ell(W) < 4n$. Then, $2(s+t) < 4n \Rightarrow s + t < 2n$.

But arithmetic mean–geometric mean (AM-GM) inequality gives $s + t \geq 2\sqrt{st} \geq 2n$, a contradiction.

Hence, no shorter word exists. The word $X^nY^nX^{-n}Y^{-n}$ of length $4n$ is minimal, and therefore, $\ell(Z^{n^2}) = 4n$. □

We can also see that the word length of any central element is always even.

In any reduced word for a central element, the total power of X is zero and the total power of Y is zero. This means the number of X letters is even and the number of Y letters is even, so their sum, which is the word length, is also even. We give a formal proof of this property below.

Lemma 4.2. *For any $c \in \mathbb{Z}^+$, the word length $\ell(Z^c)$ is always even.*

Proof. Let $W = X^{a_1}Y^{b_1}X^{a_2}Y^{b_2}\cdots X^{a_m}Y^{b_m}$ be a reduced word representing $Z^c = (0,0,c)$, where each $a_i, b_i \in \mathbb{Z}$, possibly allowing the first or last exponent to be zero. Then, the word length is $\ell(W) = 2r$, for some positive integer r.

Suppose, for contradiction, that $\ell(W)$ is odd. Then, we may write $\ell(W) = 2q + 1$ for some integer $q \geq 0$. This implies that the total number of generator terms (each being a power of X, Y, or their inverses) is odd. Thus, either the number of X-terms or the number of Y-terms is odd.

However, since W represents $Z^c = (0,0,c)$, it must satisfy $\sum a_i = 0$ and $\sum b_i = 0$, which is only possible when both the number of X-terms and the number of Y-terms are even. This contradiction shows that $\ell(W)$ cannot be odd.

Therefore, $\ell(Z^c) = \ell(W) = 2r$, for some $r \in \mathbb{Z}^+$. Hence, $\ell(Z^c)$ is always even. □

These observations serve as the foundation for the formulas we now present in Theorem 4.3, which describe the formulas for the word length of the elements of the center of $H(\mathbb{Z})$.

Theorem 4.3. *Let $c > 0$ and $n \in N$. Then, the word length for Z^c is given by*

$$\ell(Z^c) = \begin{cases} 4n, & if\ c = n^2 \\ 2(2n+1), & if\ c = n^2 + k, where\ 0 < k \leq n \\ 4(n+1), & if\ c = n(n+1) + k, where\ 0 < k \leq n \end{cases}.$$

Proof. To establish the minimal length of Z^c for a given positive integer c, we consider all factorizations $c = st$, where $s, t \in \mathbb{Z}^+$, and aim to minimize the sum $s + t$. This sum directly governs the word length $\ell(Z^c)$, since the commutator expression $X^sY^tX^{-s}Y^{-t}$ yields $\ell(Z^k) \leq 2(s + t)$.

(i) Suppose $c = n^2$. By Lemma 4.1, we have
$$\ell(Z^{n^2}) = 4n.$$

(ii) Now suppose $c = n^2 + k$, where $0 < k \leq n$. From *(i)*, we have $\ell(Z^{n^2}) = 4n$.

We begin when $k = 1$. Observe that $Z^{n^2+1} = Z^{n^2}Z$. By concatenating the minimal word of Z^{n^2} with any expression of Z from Corollary 2.5, we get a word of at least $4n + 4$. But since Z is in the center of $H(\mathbb{Z})$, carefully choosing and inserting an expression of Z from Corollary 2.5 to the expansion $Z^{n^2} = X^nY^nX^{-n}Y^{-n}$ can just add two units to the word length of Z^{n^2}. That is,

$$\begin{aligned} Z^{n^2+1} &= Z^{n^2}Z \\ &= X^nY^nX^{-n}Y^{-n}(X^{-1}Y^{-1}XY) \\ &= X^n(X^{-1}Y^{-1}XY)Y^nX^{-n}Y^{-n} \\ &= X^{n-1}Y^{-1}XY^{n+1}X^{-n}Y^{-n}. \end{aligned}$$

From this, $\ell(Z^{n^2+1}) \leq 4n + 2$.

When $k = n$, we have $c = n^2 + n = n(n + 1)$. The sum $s + t$ is minimized when $s = n$ and $t = n + 1$ or $s = n + 1$ and $t = n$. Without loss of generality, taking $s = n$ and $t = n + 1$, the explicit minimal expression from Corollary 2.8 *(i)* is

$$Z^{n^2+n} = Z^{n(n+1)} = X^{n+1-j}Y^nX^{-(n+1)}Y^{-n}X^j,\ (0 \leq j \leq n). \quad (4.6)$$

In either case, $s + t = 2n + 1$ and by Corollary 2.8 yields
$$\ell(Z^{n(n+1)}) \leq 4n + 2 = 2(2n + 1).$$

Since Z lies in the center of $H(\mathbb{Z})$, inserting or moving Z does not alter the group element it represents. Performing l swaps of Y with one adjacent X, each swap reduces the third coordinate by exactly ℓ. Hence,

$$Z^{n^2+n-l} = X^{n-j}Y^lXY^{n-l}X^{-(n+1)}Y^{-n}X^j = X^{n+1-j}Y^nX^{-n}Y^{-l}X^{-1}Y^{-(n-l)}X^j,\ (1 \leq l < n). \quad (4.7)$$

Though $n^2 + n - l \neq n(n + 1)$, the pairs $s = n$ and $t = n + 1$ (or vice versa) minimize $s + t$ ensuring minimality. Thus, combining Equations 4.6 and 4.7, we find that for $0 < k \leq n$,
$$\ell\left(Z^{n^2+k}\right) \leq 4n + 2 = 2(2n + 1).$$

We now show $\ell\left(Z^{n^2+k}\right) = 4n + 2 = 2(2n + 1)$.

First, we will show that $\ell\left(Z^{n^2+k}\right) > 4n$.
Let $c = n^2 + k$ with $1 \leq k \leq n$. Suppose, for contradiction, that there exists a reduced

word $W = X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \cdots X^{a_m} Y^{b_m}$ for Z^c with total length $A + B \leq 4n$, where

$$A = \sum_{i=1}^{m} |a_i|, \quad B = \sum_{i=1}^{m} |b_i|.$$

Since Z^c is central, we have $\sum a_i = 0$ and $\sum b_i = 0$, so exactly $\frac{A}{2}$ of the X-letters and $\frac{B}{2}$ of the Y-letters are positive. Only the positive X-letters and positive Y-letters can contribute to the largest possible value of the third coordinate. By Lemma 2.3, the third coordinate is $c = \sum_{i \leq j} a_i b_j$.

Placing all positive X-letters before all positive Y-letters ensures each positive X pairs with each positive Y, giving the largest possible value in Lemma 2.3. This gives

$$c \leq \frac{A}{2} \cdot \frac{B}{2} = \frac{AB}{4}.$$

From the AM-GM inequality,

$$AB \leq \frac{(A+B)^2}{4} \leq \frac{(4n)^2}{4} = 4n^2,$$

so $c \leq \frac{AB}{4} \leq n^2$, contradicting $c = n^2 + k > n^2$.

Thus, no reduced word for Z^{n^2+k} can have length at most $4n$, and so $\ell(Z^{n^2+k}) > 4n$.

By Lemma 4.2, every central element has even word length. Therefore, since $\ell(Z^{n^2+k})$ is even and strictly greater than $4n$, the smallest possible value is $\ell(Z^{n^2+k}) = 4n + 2 = 2(2n+1)$.

(iii) For the remaining case, suppose

$$c = n(n+1) + k \text{ with } 0 < k \leq n.$$

Minimizing $s + t$ is achieved by setting $s = n$ and $t = n + 2$ or $s = n + 2$ and $t = n$. In either case, $s + t = 2n + 2$ and by Corollary 2.8,

$$\ell(Z^{n(n+2)}) \leq 2(2n+2) = 4(n+1).$$

Without loss of generality, choose $s = n$ and $t = n + 2$ and from Corollary 2.8 *(i)*, the explicit minimal form is

$$Z^{n^2+2n} = Z^{n(n+2)} = X^{n+2-j} Y^n X^{-(n+2)} Y^{-n} X^j, \ (0 \leq j \leq n) \quad (4.8)$$

which indeed has word length $4(n+1)$.

We now show how to lower the exponent without changing the length. Inside the word in Equation 4.8, there is an adjacent pair XY. Using $XY = YXZ$, we may swap this pair, which introduces one central Z.

$$\begin{aligned} Z^{n^2+2n} &= X^{n+1-j}(XY)Y^{n-1}X^{-(n+2)}Y^{-n}X^j \\ &= X^{n+1-j}(YXZ)Y^{n-1}X^{-(n+2)}Y^{-n}X^j \\ &= X^{n+1-j}YXY^{n-1}X^{-(n+2)}Y^{-n}X^jZ. \end{aligned}$$

Multiplying both sides on the right by Z^{-1} yields a word of the same length representing $Z^{n^2+2n-1} = X^{n+1-j}YXY^{n-1}X^{-(n+2)}Y^{-n}X^j$ and therefore decreases the central exponent by exactly 1 while leaving the number of generator letters unchanged.

Repeating the same computation once more produces $Z^{n^2+2n-2} = X^{n+1-j}Y^2XY^{n-2}X^{-(n+2)}Y^{-n}X^j$, still of word length $4(n+1)$.

Continuing in this manner, after l swaps we obtain

$$Z^{n^2+2n-l} = X^{n+1-j}Y^lXY^{n-l}X^{-(n+2)}Y^{-n}X^j, 0 \leq l < n \quad (4.9)$$

which again has word length $4(n+1)$.

Thus, every exponent in the range $n^2 + n + 1 \leq n^2 + 2n - l \leq n^2 + 2n$ can be obtained without increasing the word length.

Now, write $c = n^2 + n + k$. We solve $n^2 + 2n - l = n^2 + n + k$, which gives $l = n - k$. Since $0 < k \leq n$, we have $0 \leq l < n$, so an appropriate number of swaps exists. Substituting $l = n - k$ into Equation 4.9 yields a reduced word representing Z^c of length $4(n+1)$. Thus $\ell(Z^c) \leq 4(n+1)$.

Now, we show $\ell(Z^c) = 4(n+1)$.

Note that by Lemma 4.2, the word length of any central element is even; thus, the inequality $\ell(Z^c) > 4n + 2$ forces $\ell(Z^c) \geq 4n + 4$. Together with the upper bound $\ell(Z^c) \leq 4(n+1)$, the only possible value is $\ell(Z^c) = 4(n+1)$.

Let $c = n(n+1) + k$ with $0 < k \leq n$. Suppose, for contradiction, that there exists a reduced word $W = X^{a_1} Y^{b_1} X^{a_2} Y^{b_2} \cdots X^{a_m} Y^{b_m}$ for Z^c with total length $A + B \leq 4n + 2$, where

$$A = \sum_{i=1}^{m} |a_i|, \quad B = \sum_{i=1}^{m} |b_i|.$$

Since Z^c is central, we have $\sum a_i = 0$ and $\sum b_i = 0$, so exactly $\frac{A}{2}$ of the X-letters and $\frac{B}{2}$ of the Y-letters are positive. By Lemma 2.3, the third coordinate is $c = \sum_{i<j} a_i b_j$.

The sum is the largest when all positive X-letters appear before all positive Y-letters, giving, $c \leq \frac{A}{2} \cdot \frac{B}{2} = \frac{AB}{4}$.

From the AM-GM inequality,

$$AB \leq \frac{(A+B)^2}{4} \leq \frac{(4n+2)^2}{4} = 4n^2 + 4n + 1,$$

so $c \leq \frac{AB}{4} \leq n^2 + n + \frac{1}{4}$, contradicting $c = n(n+1) + k > n^2 + n + \frac{1}{4}$. Thus, $\ell(Z^c) > 4n + 2$.

Thus, no shorter word exists. Since a word of length $4(n+1)$ exists by construction, and since the word length of any central element must be even (Lemma 4.2), the strict inequality $\ell(Z^c) > 4n + 2$ forces $\ell(Z^c) \geq 4n + 4$.

Combined with the upper bound $\ell(Z^c) \leq 4(n+1)$, therefore, for $c = n(n+1) + k$ with $0 < k \leq n$, we conclude that $\ell(Z^c) = 4(n+1)$. □

Hence, Theorem 4.3 classifies all word length of Z^c for any positive integer c. Each case achieves minimality based on the arithmetic structure of c.

In the next results, we give a formula for $\ell\big((a,b;c)\big)$ for all nonnegative integers a, b, and c. Specifically, Theorem 4.4 tackles the case when $0 \leq c \leq ab$ while Theorem 4.5 discusses the case when $c > ab$.

Theorem 4.4. *Let $a, b, c \in \mathbb{Z}_{\geq 0}$ such that $c \leq ab$. Then, $\ell\big((a,b;c)\big) = a + b$.*

Proof. Consider the word $W = X^a Y^b$. By Lemma 2.3, this word represents the group element $(a,b;ab)$ and has length $a + b$. Thus, $\ell\big((a,b;ab)\big) \leq a + b$.

Now in Equation 2.5, we have the relation $XY = YXZ$. Swapping one X past a Y decreases the last coordinate by one. Specifically, we have

$$\begin{aligned}(a,b;\ ab-1) &= (a,b;\ ab)(0,0;-1)\\ &= X^a Y^b Z^{-1}\\ &= X^a Y^b (X^{-1}YXY^{-1}),\end{aligned}$$

which simplifies to $(a,b;ab-1) = X^{a-1}YXY^{b-1}$.

Notice explicitly that after inserting Z^{-1}, four letters initially appear, but due to the fundamental commutative relation (Equation 2.5), these letters precisely cancel. This scenario achieves optimal cancellation, minimizing the final word length. Repeated application of such swaps allows us to construct a word of the same length, $a + b$, that represents the element $(a,b;ab-k)$ for any $0 \leq k \leq ab$. Therefore, for every integer c with $0 \leq c \leq ab$, there exists a word of length exactly $a + b$ representing $(a,b;c)$. Thus, $\ell\big((a,b;c)\big) \leq a + b$.

More generally, starting from $(a,b;ab) = X^a Y^b$, we systematically swap using $XY = YXZ$ and $Z = XYX^{-1}Y^{-1}$ yielding

$$\begin{aligned}(a,b;ab) &= X^a Y^b\\ (a,b;ab-1) &= X^{a-1}YXY^{b-1}\\ (a,b;ab-2) &= X^{a-2}YX^2Y^{b-1} = X^{a-1}Y^2XY^{b-2}\\ &\vdots\\ (a,b;ab-a) &= X^{a-a}YX^aY^{b-1} = YX^aY^{b-1}. \quad (4.10)\end{aligned}$$

Thus, in general, if $c = ab - (ka + l)$ where $0 \leq k \leq b-1, 0 \leq l \leq a$, then $(a,b;c) = (a,b;ab-(ka+l)) = Y^k X^{a-l} Y X^l Y^{b-k-1}$.

Now suppose, for contradiction, that there exists a shorter word W' with $\ell(W') < a + b$ representing $(a,b;c)$. But then the sum of the X- and Y-exponents (and its inverses) used in W' must be fewer than $a + b$. But by Lemma 2.3, the first two coordinates of any element represented by a word are equal to the total exponents of X and Y, respectively. Hence, using fewer than a Xs or fewer than b Ys would make it impossible to reach the required first two coordinates. This contradiction shows that no such shorter word exists. Thus, $\ell((a,b;c) = a + b$ for all $c \leq ab$. □

Theorem 4.5. *Let $a, b, c \in \mathbb{Z}_{\geq 0}$ and suppose $c > ab$. Then, there exist nonnegative integers s and t such that $(a+s)(b+t) \geq c$ and $s+t$ is minimal. Then,*

$$\ell\left((a, b;\ c)\right) = a + b + 2(s+t).$$

Proof. Let $(a, b; c) \in H(\mathbb{Z})$ with $c > ab$. From Lemma 2.3, any word consisting only of a copies of X and b copies of Y has a third coordinate at most ab. Hence, it is impossible to represent $(a, b; c)$ using $a + b$ generators when $c > ab$. Thus, additional copies of X and Y are necessary to reach c.

To address this, we explicitly select integers $s, t > 0$ such that $(a+s)(b+t) > c$, ensuring that the sum $s+t$ is minimal among all feasible solutions. The minimality of $s+t$ ensures that we are adding the smallest number of additional generators needed.

Starting from

$$(a+s, b+t; (a+s)(b+t)) = X^{a+s}Y^{b+t},$$

we apply the optimal swapping operations as described previously *(analogous to Equation 4.10)*, yielding a reduced word of the form

$$(a+s, b+t; c) = X^{a_1}Y^{b_1}X^{a_2}Y^{b_2}\ldots X^{a_r}Y^{b_r}, \quad (4.11)$$

where $\sum_{i=1}^{r} a_i = a+s$, $\sum_{j=1}^{r} b_j = b+t$.

Because the word in Equation 4.11 uses exactly $s+t$ added generators, and we choose $s+t$ to be minimal subject to the constraint $(a+s)(b+t) \geq c$, by the structure established in Theorem 4.4, we have

$$\begin{aligned}\ell\left((a+s, b+t;\ c)\right) &= (a+s) + (b+t)\\ &= (a+b) + (s+t).\end{aligned}$$

Premultiplying the right-hand side of Equation 4.11 by Y^{-t} and postmultiplying it by X^{-s} yields

$$(a, b; c) = Y^{-t}\left(X^{a_1}Y^{b_1}X^{a_2}Y^{b_2}\ldots X^{a_r}Y^{b_r}\right)X^{-s}.$$

Note that by the group operation in Equation 2.3, left multiplication by $Y^{-t} = (0, -t; 0)$ changes only the second coordinate, since the added $a_1 b_2$ term is $0 \cdot b = 0$.

Similarly, right multiplication by $X^{-s} = (-s, 0; 0)$ changes only the first coordinate, since the added term $a\, b_2 = a \cdot 0 = 0$. Thus, the third coordinate c remains unchanged. Therefore, the word constructed above has minimal length since $\ell((a, b; c)) \leq a + b + 2(s+t)$.

Suppose, for contradiction, that there exists another word W representing $(a, b; c)$ with length strictly less than $a + b + 2(s+t)$. Then, it must use fewer than s additional Xs or fewer than t additional Ys.

However, having fewer X- or Y-terms than required would decrease the third coordinate below c. Therefore, any word fewer generators cannot represent $(a, b; c)$.

This contradicts the assumption that W represents $(a, b; c)$.

Since the chosen values of s and t already minimize the added terms while still reaching the required c, the constructed word has the shortest possible length. Thus, $\ell((a, b; c)) = a + b + 2(s+t)$. □

WORD LENGTH IN THE FINITE HEISENBERG GROUP

We start this section by reducing the domain of the finite Heisenberg group. Applying the word-length-preserving automorphisms σ and ϕ to H_p allows us to limit the values of the elements of the field $\mathbb{Z}_p$ in the set $H_p = \{(a, b; c) | a, b, c \in \mathbb{Z}_p\}$ as follows:

$$\begin{aligned} g &= (a, b; c),\\ \sigma^2(g) &= (-a, -b; c),\\ \phi(g) &= (-a, b; -c), \text{and}\\ \sigma^2\left(\phi(g)\right) &= (a, -b; -c).\end{aligned}$$

Thus, if we set $\mathbb{Z}_p = \left\{-\left\lfloor\frac{p}{2}\right\rfloor, \ldots, -2, -1, 0, 1, 2, \ldots, \left\lfloor\frac{p}{2}\right\rfloor\right\}$, then we may limit the values of a, b to $0 \leq a, b \leq \left\lfloor\frac{p}{2}\right\rfloor$ and $c \in \mathbb{Z}_p$. That is, σ and ϕ reduces the domain of $H_p = \{(a, b; c) | a, b, c \in \mathbb{Z}_p\}$ to about a quarter of its total number of elements.

Theorem 5.1. *Let* $0 < c < \left\lfloor \frac{p}{2} \right\rfloor$ *and* $n \in \mathbb{N}$. *Then*

$$\ell(Z^c) = \begin{cases} 4n, & if\ c = n^2 \\ 2(2n+1), & if\ c = n^2 + k, where\ 0 < k \leq n \\ 4(n+1), & if\ c = n(n+1) + k, where\ 0 < k \leq n \end{cases}.$$

Proof. In the finite Heisenberg group H_p, the third coordinate is taken modulo p. The condition $0 < c < \left\lfloor \frac{p}{2} \right\rfloor$ ensures that the central element $Z^c = (0,0,c)$ does not undergo any reduction modulo p. Within this range, the third coordinate of Z^c in H_p agrees exactly with its third coordinate in the infinite group $H(\mathbb{Z})$. Consequently, the word length of Z^c in the finite case is identical to its word length in the infinite case. Thus, Theorem 4.3 applies directly and yields the stated formula for $\ell(Z^c)$.

Remark. In Theorem 5.1, the expressions $4n$, $2(2n+1)$, and $4(n+1)$ describe the actual word lengths of the elements Z^c. These values are nonnegative integers and are not taken modulo p.

Remark. When $c > \left\lfloor \frac{p}{2} \right\rfloor$, $c \in \left\{-1, -2, \ldots, -\left\lfloor \frac{p}{2} \right\rfloor\right\}$, so that $\ell(Z^c) = \ell(Z^{-c})$, where $0 < -c < \left\lfloor \frac{p}{2} \right\rfloor$. By Theorem 3.1, which states that taking inverses does not change word length, $\ell(Z^c) = \ell(Z^{-c})$. Thus, $\ell(Z^c)$ can be obtained from $\ell(Z^{-c})$ and Theorem 5.1.

Theorem 5.2. *Let* $a, b, c \in \mathbb{Z}_p$ *such that* $0 < a, b < \left\lfloor \frac{p}{2} \right\rfloor$. *If* $0 \leq c \leq ab$ *(ab is not reduced mod p), then* $\ell((a,b;c)) = a + b$.

Proof. The condition $0 < a, b < \left\lfloor \frac{p}{2} \right\rfloor$ comes from the reduction used in Theorem 5.1: the automorphisms σ and ϕ map any element of H_p to another representative with the same word length but whose first two coordinates lie in this range. Thus, we may assume a and b satisfy this bound without loss of generality.

Under this assumption, the third coordinate c is not reduced modulo p, so the word length of $(a,b;c)$ in H_p is the same as in the discrete Heisenberg group. The conclusion now follows directly from Theorem 4.4.

Corollary 5.3. *Let* $p \geq 11$. *If* $0 < a, b < \left\lfloor \frac{p}{2} \right\rfloor$ *and* $ab \geq p - 1$ *(ab is not reduced modulo p), then* $\ell\left((a,b;\ c)\right) = a + b,\ \forall c\ \in \mathbb{Z}_p$.

Remark. The condition $ab \geq p - 1$ with $0 < a, b < \left\lfloor \frac{p}{2} \right\rfloor$ can occur only when $p \geq 11$, since $\left(\left\lfloor \frac{p}{2} \right\rfloor - 1\right)^2 \geq p - 1$ first holds for $p = 11$.

Lemma 5.4. *Let* $(a,b;c) \in H_p$ *with* $0 \leq a, b \leq \lceil \sqrt{p}\ \rceil$ *and* $\left\lceil \frac{ab}{2} \right\rceil \leq c \leq ab$. *If* $ab < p$ *and* W *is a word with word length* ℓ_W *representing* $(a,b;c)$, *then*

i. $(a,b;ab-c) = W^R$ *and*
ii. $\ell\big((a,b;ab-c)\big) = \ell_W$.

Proof.
i. Let $(a,b;c) \in H_p$ and
$W = X^{a_1}Y^{b_1}\,X^{a_2}Y^{b_2}\ldots X^{a_{k-1}}Y^{b_{k-1}}\,X^{a_k}Y^{b_k}$ with $0 \leq a, b \leq \lceil \sqrt{p}\ \rceil$ and $\left\lceil \frac{ab}{2} \right\rceil \leq c \leq ab$ and $\ell(W) = \ell_W$.

If $ab < p$, then by Lemma 2.3,

$$X^{a_1}Y^{b_1}\,X^{a_2}Y^{b_2}\ldots X^{a_{k-1}}Y^{b_{k-1}}\,X^{a_k}Y^{b_k} = \left(\sum_{i=1}^{k} a_i\,, \sum_{j=1}^{k} b_j\,; \sum_{\substack{i,j=1 \\ i \leq j}}^{k} a_i\,b_j\right).$$

So we have

$$\begin{aligned} ab - c &= \sum_{i=1}^{k} a_i \sum_{j=1}^{k} b_j - \sum_{\substack{i,j=1 \\ i \leq j}}^{k} a_i\,b_j \\ &= \sum_{i,j=1}^{k} a_i b_j - \sum_{\substack{i,j=1 \\ i \leq j}}^{k} a_i\,b_j \\ &= \sum_{\substack{i,j=1 \\ i > j}}^{k} a_i\,b_j \end{aligned}$$

and by Lemma 2.4,

$$\begin{aligned} (a,b;ab-c) &= \left(\sum_{i=1}^{k} a_i\,, \sum_{j=1}^{k} b_j\,; \sum_{\substack{i,j=1 \\ i > j}}^{k} a_i\,b_j\right) \\ &= Y^{b_k}X^{a_k}Y^{b_{k-1}}X^{a_{k-1}}\ldots Y^{b_2}X^{a_2}Y^{b_1}X^{a_1} \\ &= W^R. \end{aligned}$$

ii. Immediate from *(i)* above and the assumption that W is of word length ℓ_W. □

Remark. The assumptions $0 \le a, b \le \lfloor\sqrt{p}\rfloor$ ensure that $ab < p$, so the third coordinate of $(a, b; c)$ in H_p is not reduced modulo p. While the condition $\left\lceil\frac{ab}{2}\right\rceil \le c \le ab$ guarantees that the third coordinate appearing in Lemma 2.3 remains valid in H_p, the integer value of the third coordinate coincides with its value in $H(\mathbb{Z})$ and is unaffected by reduction modulo p. With these two assumptions, the computation of the third coordinate proceeds exactly as in $H(\mathbb{Z})$, allowing Lemma 2.3 to be applied without modification.

Lemma 5.5. *Let* $(a, b;\ c) \in H_p$ *with* $0 <\ a, b \le \left\lfloor\frac{p}{2}\right\rfloor$. *If* $ab > p$, *then there exists an* $s, 0 < s < p$, *such that* $ab \equiv s \bmod p$, *so that* $(a, b;\ ab) = (a, b; s) \in H_p$.
Moreover, for $\left\lceil\frac{s}{2}\right\rceil \le c \le s$, *let* $(a, b; c) = W'$, *where* W' *is a word with word length* $\ell_{W'}$, *then*

i. $(a, b; s - c) = W'^R$, *and*
ii. $\ell\big((a, b; s - c)\big) = \ell_{W'}$.

Proof. Since $ab > p, ab \equiv s \bmod p,\ 0 < s < p$. Thus, $(a, b; ab) = (a, b; s)$. Let $\left\lceil\frac{s}{2}\right\rceil \le c \le\ s$ and let $(a, b; c) = W'$. Since $ab > p,\ p > s,\ s \ge c$, we have $ab > c$. By Theorem 5.2, $\ell\ ((a, b;\ c)) = a + b$.

Thus, $W' = X^{a_1}Y^{b_1}\ X^{a_2}Y^{b_2}\ldots X^{a_k}Y^{b_k}$ with $a = \sum_{i=1}^{k} a_i$, $b = \sum_{j=1}^{k} b_j$, and $c = \sum_{\substack{i,j=1 \\ i\le j}}^{k} a_i\, b_j$.

So, we have $\ell_{W'} = \sum_{i=1}^{k} a_i + \sum_{j=1}^{k} b_j = a + b$.

Now, from Lemma 5.4 *(i)*,

$$\begin{aligned} W'^R &= Y^{b_k}X^{a_k} \ldots Y^{b_2}X^{a_2}Y^{b_1}X^{a_1} \\ &= \left(\sum_{i=1}^{k} a_i\,, \sum_{j=1}^{k} b_j\,; \sum_{\substack{i,j=1 \\ i>j}}^{k} a_i\, b_j\right) \\ &= (a, b; ab - c) \\ &= (a, b; s - c) \end{aligned}$$

since $ab \equiv s \bmod p$. This proves part (i).

Part *(ii)* immediately follows since $ab > s - c$. □

Lemma 5.6. *Let* $(a, b; c) \in H_p$ *with* $0 < a, b < \left\lfloor\frac{p}{2}\right\rfloor$. *If* $ab < c$, *then there exists an integer* $\gamma > 0$ *such that* $(a, b; c) = (a, b; ab + \gamma)$. *Moreover, if* $(a, b; ab + \gamma) = W$, *where* W *is a word of word length* ℓ_W, *then*

i. $(a, b;\ -\gamma) = W^R$ *and*
ii. $\ell\ ((a, b;\ -\gamma)) = \ell_W$.

Proof. Let $(a, b; c) \in H_p$ with $0 < a, b < \left\lfloor\frac{p}{2}\right\rfloor$.

If $ab < c$, then by division algorithm $c = ab + \gamma$, where $0 < \gamma \le p - ab$ and $(a, b; c) = (a, b; ab + \gamma)$.

Now to prove *(i)*, let $W = (a, b; ab + \gamma)$. By Lemma 3.8,

$$W^R = (a, b; -(ab + \gamma) + ab) = (a, b; -\gamma).$$

Finally, *(ii)* follows from second statement of Lemma 3.8. □

Lemma 5.7. *Let* $(a, b;\ c) \in H_p$ *with* $0 < a, b < \left\lfloor\frac{p}{2}\right\rfloor$ *such that* $ab < c$. *Then, there exists a* $\delta > 0$ *such that*

$$(a, b; ab + \delta) = \begin{cases} (a, b; -\delta), & \text{if } ab \text{ is odd} \\ (a, b; -\delta - 1), & \text{if } ab \text{ is even} \end{cases}.$$

Proof. Set $\delta = \left\lfloor\frac{p-ab}{2}\right\rfloor$.

If ab is odd, then $(p - ab)$ is even. Thus, $\delta = \frac{p-ab}{2}$ or equivalently $ab = p - 2\delta$, and we have $(a, b; ab + \delta) = (a, b; (p - 2\delta) + \delta) = (a, b; -\delta)$.

Now, suppose ab is even. Then, $(p - ab)$ is odd, and $\delta = \frac{p-ab-1}{2}$ or equivalently $ab = p - 2\delta - 1$.

Hence,

$$\begin{aligned} (a, b; ab + \delta) &= (a, b; (p - 2\delta - 1) + \delta\ ) \\ &= (a, b; -\delta - 1). \end{aligned}$$ □

Remark. Let $(a, b;\ c) \in H_p$ with $ab < c$. Lemmas 5.6 and 5.7 indicate that with

respect to word length, we can further limit the value of c to $ab < c \leq ab + \left\lfloor \frac{p-ab}{2} \right\rfloor$.

Theorem 5.8. *Let $a, b, c \in \mathbb{Z}_p$ such that $0 < a, b < \left\lfloor \frac{p}{2} \right\rfloor$. If $ab < c \leq ab + \left\lfloor \frac{p-ab}{2} \right\rfloor$, then*

$$\ell\left((a, b;\ c)\right) = a + b + 2(s + t)$$

for some nonnegative integers s and t such that $(a + s)(b + t) \geq c$ and $s + t$ is minimal.

Proof. Follows from Theorem 4.5. In this range of c, the third coordinate never reaches p, so no reduction modulo p occurs. Hence the computation is the same as in $H(\mathbb{Z})$, and the formula from Theorem 4.5 applies without change. □

When $a = \left\lfloor \frac{p}{2} \right\rfloor, b = 1$ and $c \in \mathbb{Z}_p = \{0,1,2, \dots, p-1\}$, we get this interesting result.

Theorem 5.9. *In H_p, if $c > \left\lfloor \frac{p}{2} \right\rfloor$, then*

$$\ell\left(\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; c\right)\right) = \left\lfloor \frac{p}{2} \right\rfloor + 2.$$

Proof. If $c > \left\lfloor \frac{p}{2} \right\rfloor$, and write $c \equiv m \bmod p$ where $m \in \left\{-1, -2, \dots, -\left\lfloor \frac{p}{2} \right\rfloor\right\}$. Let $i = m + \left\lfloor \frac{p}{2} \right\rfloor$, so that $i \in \left\{0,1, \dots, \left\lfloor \frac{p}{2} \right\rfloor - 1\right\}$. We start with the expression $X^{-\left\lfloor \frac{p}{2} \right\rfloor} Y = \left(-\left\lfloor \frac{p}{2} \right\rfloor, 1; -\left\lfloor \frac{p}{2} \right\rfloor\right)$ and perform the swapping operation in Equation 4.10.

Observe that swapping one copy of X^{-1} with a copy of Y to its right, that is, $X^{-1}Y \to YX^{-1}$, increases the third coordinate of the triple by one unit. Thus,

$$\begin{aligned} X^{-\left\lfloor \frac{p}{2} \right\rfloor} Y &= \left(-\left\lfloor \frac{p}{2} \right\rfloor, 1; -\left\lfloor \frac{p}{2} \right\rfloor\right) \\ X^{-\left\lfloor \frac{p}{2} \right\rfloor + 1} Y X^{-1} &= \left(-\left\lfloor \frac{p}{2} \right\rfloor, 1; -\left\lfloor \frac{p}{2} \right\rfloor + 1\right) \\ X^{-\left\lfloor \frac{p}{2} \right\rfloor + 2} Y X^{-2} &= \left(-\left\lfloor \frac{p}{2} \right\rfloor, 1; -\left\lfloor \frac{p}{2} \right\rfloor + 2\right) \\ &\vdots \\ X^{-1} Y X^{-\left\lfloor \frac{p}{2} \right\rfloor + 1} &= \left(-\left\lfloor \frac{p}{2} \right\rfloor, 1; -1\right). \end{aligned}$$

In general,

$$X^{-\left\lfloor \frac{p}{2} \right\rfloor + i} Y X^{-i} = \left(-\left\lfloor \frac{p}{2} \right\rfloor, 1; -\left\lfloor \frac{p}{2} \right\rfloor + i\right), \text{ where } 0 \leq i \leq \left\lfloor \frac{p}{2} \right\rfloor - 1. \quad (5.1)$$

Now, postmultiplying both sides of Equation 5.1 by X^{-1} yields

$$X^{-\left\lfloor \frac{p}{2} \right\rfloor + i} Y X^{-(i+1)} = \left(-\left\lfloor \frac{p}{2} \right\rfloor - 1, 1; -\left\lfloor \frac{p}{2} \right\rfloor + i\right) = \left(\left\lfloor \frac{p}{2} \right\rfloor, 1; m\right).$$

Since $m = -\left\lfloor \frac{p}{2} \right\rfloor + i$, this choice of i gives the required representative of $\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; m\right)$. Consequently, we get

$$\ell\left(\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; m\right)\right) = \ell\left(\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; c\right)\right) \leq \left(\left\lfloor \frac{p}{2} \right\rfloor - i + 1\right) + (i + 1) = \left\lfloor \frac{p}{2} \right\rfloor + 2.$$

It remains to show that no shorter word represents $\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; m\right)$. Indeed, any reduced word of length at most $\left\lfloor \frac{p}{2} \right\rfloor + 1$ with second coordinate 1 must contain exactly one letter $Y^{\pm 1}$ and at most $\left\lfloor \frac{p}{2} \right\rfloor$ letters $X^{\pm 1}$. In that case, the third coordinate is determined by the position of this single $Y^{\pm 1}$ among the $X^{\pm 1}$s, and successive swaps change the third coordinate by exactly one.

Starting from $X^{-\left\lfloor \frac{p}{2} \right\rfloor} Y$, these swaps produce precisely the values $-\left\lfloor \frac{p}{2} \right\rfloor, -\left\lfloor \frac{p}{2} \right\rfloor + 1, \dots, -1$; hence, m cannot be obtained with length less than or equal to $\left\lfloor \frac{p}{2} \right\rfloor + 1$ once the first two coordinates are fixed. Therefore,

$$\ell\left(\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; m\right)\right) \geq \left\lfloor \frac{p}{2} \right\rfloor + 2.$$

Combining both inequalities gives

$$\ell\left(\left(\left\lfloor \frac{p}{2} \right\rfloor, 1; c\right)\right) = \left\lfloor \frac{p}{2} \right\rfloor + 2,$$

as desired. □

Remark. Compared to Theorem 5.8, the above results state that when $a = \left\lfloor \frac{p}{2} \right\rfloor, b = 1$ and $ab < c$, then $\ell\left((a, b;\ c)\right) = a + b + 1$.

Corollary 5.10. *In H_p, if $c > \left\lfloor \frac{p}{2} \right\rfloor$, then*

$$\ell\left(\left(1, \left\lfloor \frac{p}{2} \right\rfloor; c\right)\right) = \left\lfloor \frac{p}{2} \right\rfloor + 2.$$

Proof. This result directly follows from Theorem 5.9, Lemma 3.9, and the remark after the proof of Theorem 5.9.

THE ELEMENTS OF H_p

In this section, we focus on the finite Heisenberg group H_p being a finite p-group and study some of its algebraic structures. It turns out that the automorphism σ can be used to classify the elements of H_p. Further investigations on some algebraic structures in H_p reveal connection to the word length of the elements of H_p.

This illustrates that the automorphism σ in Theorem 3.4 also induces other algebraic structures in H_p. We first present a classification of the elements of the finite Heisenberg group using this automorphism. Then, we discuss how this classification induces some structures on the elements of H_p.

Definition 6.1. Let $g \in H_p$. Then

i. g is a Type-I element if g is an element of the center of H_p,
ii. g is a Type-II element if $\sigma^2(g) = g^{-1}$, and
iii. g is a Type-III element if g is neither Type-I nor Type-II.

Note: The identity element $e = (0,0;0)$ satisfies both conditions for Type-I and Type-II and thus belongs to both categories. To avoid ambiguity in classification counts, we explicitly recognize this overlap.

Lemma 6.1. *The word length of a Type-I element is even.*

Proof. Immediate from Definition 6.1 (i) and Lemma 4.2.

Theorem 6.2. *Let $g = (a,b;c) \in H_p$. The following are equivalent:*

i. g is a Type-II element.
ii. $c = \frac{ab}{2}$.
iii. If $g = X^{a_1}Y^{b_1}\,X^{a_2}Y^{b_2}\dots X^{a_n}Y^{b_n}$, then $g = W^R$.

Proof. Let $g = (a,b;c) \in H_p$.

We first prove $(i) \Rightarrow (ii)$. If g is a Type-II element, then $\sigma^2(g) = g^{-1}$.

But $\sigma^2(g) = (-a,-b;c)$ and $g^{-1} = (-a,-b;-c+ab)$. Now, $\sigma^2(g) = g^{-1}$ or equivalently $c = -c + ab$. That is $c = \frac{ab}{2}$.

To show $(ii) \Rightarrow (iii)$, observe first that from Lemma 2.3, if $g = X^{a_1}Y^{b_1}\,X^{a_2}Y^{b_2}\dots X^{a_n}Y^{b_n}$, then

$$g = \left(\textstyle\sum_{i=1}^{k} a_i\,, \sum_{j=1}^{k} b_j\,; \sum_{\substack{i,j=1\\ i\le j}}^{k} a_i\,b_j\right). \quad (6.1)$$

Suppose (ii) holds. Then,

$$c = \frac{ab}{2}, \text{and } 2c = ab. \quad (6.2)$$

Combining Equations 6.1 and 6.2, we find

$$2\left(\sum_{\substack{i,j=1\\ i\le j}}^{n} a_i\,b_j\right) = \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n} b_i\right)$$

$$2\left(\sum_{\substack{i,j=1\\ i\le j}}^{n} a_i\,b_j\right) = \sum_{i,j=1}^{n} a_i\,b_j$$

$$2\left(\sum_{\substack{i,j=1\\ i\le j}}^{n} a_i\,b_j\right) = \sum_{\substack{i,j=1\\ i\le j}}^{n} a_i\,b_j + \sum_{\substack{i,j=1\\ i>j}}^{n} a_i\,b_j$$

$$\sum_{\substack{i,j=1\\ i\le j}}^{n} a_i\,b_j = \sum_{\substack{i,j=1\\ i>j}}^{n} a_i\,b_j$$

$$c = \sum_{\substack{i,j=1\\ i>j}}^{n} a_i\,b_j.$$

This implies that

$$g = \left(\sum_{i=1}^{n} a_i\,, \sum_{b=1}^{n} b_{j;}\, \sum_{\substack{i,j=1\\ i>j}}^{n} a_i\,b_j\right)$$

and from Lemma 2.4, $g = W^R$.

Lastly, we prove $(iii) \Rightarrow (i)$. By hypothesis, if

$$g = X^{a_1}Y^{b_1}\,X^{a_2}Y^{b_2}\dots X^{a_{n-1}}Y^{b_{n-1}}X^{a_n}Y^{b_n}, \quad (6.3)$$

then $g = Y^{b_n}X^{a_n}Y^{b_{n-1}}X^{a_{n-1}}\dots Y^{b_2}X^{a_2}Y^{b_1}X^{a_1}$ (6.4)

But from Equation 6.4 and Corollary 3.3,

$$\sigma^2(g) = Y^{-b_n}X^{-a_n}Y^{-b_{n-1}}X^{-a_{n-1}}\dots Y^{-b_2}X^{-a_2}Y^{-b_1}X^{-a_1} \quad (6.5)$$

Comparing Equations 6.3 and 6.5, we find $\sigma^2(g) = g^{-1}$ and g is a Type-II element. □

Remark. Theorem 6.2 (i) and (iii) indicate that a Type-II element is a sort of a generalized palindrome.

Example 6.1. In many cases, a word representing a Type-II element and its reverse represent the same group element. However, they need not be the same word. For instance, the word XY^2X^3Y and its reverse YX^3Y^2X both represent the element $(4, 3; 6)$, but they are not the same word—highlighting the nonuniqueness of word representations even for symmetric elements.

Lemma 6.3. *Let* $g = (a, b; c) \in H_p$ *where* $0 \le a, b \le \left\lfloor \frac{p}{2} \right\rfloor$. *If* g *is a Type-II element and* ab *is even (*ab *is not reduced mod* p*), then*

$$\ell(g) = a + b.$$

Proof. Let $g = (a, b; c) \in H_p$ where $0 \le a, b \le \left\lfloor \frac{p}{2} \right\rfloor$.

If g is a Type-II element, then by Theorem 6.2 *(ii)*, $c = \frac{ab}{2}$ so that $g = \left(a, b; \frac{ab}{2}\right)$.

Moreover, since ab is even and not reduced $mod\ p$, $ab > \frac{ab}{2}$ so that by Theorem 5.2,

$$\ell(g) = a + b.$$ □

Theorem 6.4. *Let* $g = (a, b; c) \in H_p$ *and* $a \neq 0$. *Then,* g *is a Type-II element if and only if* $g \in \langle\left(1, b'; \frac{b'}{2}\right)\rangle$ *for some* $b' \in \mathbb{Z}_p$.

Proof. Let $g = (a, b;\ c) \in H_p$ with $a \neq 0$.

We first prove sufficiency. If g is a Type-II element, then $c = \frac{ab}{2}$. Direct computation using Lemma 2.1 reveals

$$g = \left(a, b; \frac{ab}{2}\right) = \left(1, a^{-1}b; \frac{a^{-1}b}{2}\right)^a.$$

Thus, $g \in \langle\left(1, b'; \frac{b'}{2}\right)\rangle$, where $b' = a^{-1}b$.

For the other direction, if $g \in \langle\left(1, b'; \frac{b'}{2}\right)\rangle$, then

$$\begin{aligned} g &= \left(1, b'; \frac{b'}{2}\right)^k \text{ for some } k \in \mathbb{Z}_p \\ &= \left(k, k\,b'; \frac{kb'}{2} + \frac{b'k(k-1)}{2}\right) \\ &= \left(k, k\,b'; \frac{k^2 b'}{2}\right). \end{aligned}$$

By Theorem 6.2 *(ii)*, g is a Type-II element. □

Remark. If $b = 0$, then $g \in \langle X\rangle$, and g is a Type-II element.

Theorem 6.5. *Let* $g = (a, b;\ c) \in H_p$ *and* $b \neq 0$. *Then,* g *is a Type-II element if and only if* $g \in \langle\left(a', 1; \frac{a'}{2}\right)\rangle$ *for some* $a' \in \mathbb{Z}_p$.

The proof of Theorem 6.5 follows the same structure as Theorem 6.4 by interchanging the roles of a and b.

In Theorem 6.4, we express the Type-II element as a power of $\left(1, b'; \frac{b'}{2}\right)$, while in Theorem 6.5, the generator is $\left(a', 1; \frac{a'}{2}\right)$, where $a' = ab^{-1}$. Both proofs utilize Lemma 2.2 and the equivalence from Theorem 6.2 *(ii)* but differ only in which coordinate is assumed nonzero.

Remark. If $a = 0$, then $g \in \langle Y\rangle$, g is a Type-II element.

Theorem 6.6. *Let* $g = (a, b;\ c) \in H_p$, *where* $a \neq 0$ *and* $b \neq 0$. *The following are equivalent:*

i. g *is a Type-II word element.*

ii. $g \in \langle\left(1, b'; \frac{b'}{2}\right)\rangle$ *for some* $b' \in \mathbb{Z}_p \backslash \{0\}$.

iii. $g \in \langle\left(a', 1; \frac{a'}{2}\right)\rangle$ *for some* $a' \in \mathbb{Z}_p \backslash \{0\}$.

Proof. Follows from Theorems 6.4 and 6.5.

Corollary 6.7. *With reference to Definition 6.1, the cardinality of each type of elements is as follows:*

i. $\left|\{g \in H_p \mid g \text{ is a } Type - I \text{ element}\}\right| = p$,

ii. $\left|\{g \in H_p \mid g \text{ is a } Type - II \text{ element}\}\right| = p^2$, *and*

iii. $\left|\{g \in H_p | g \text{ is a } Type - III \text{ element}\}\right| = p^3 - p^2 - p + 1$.

Proof.

i. Since the center of H_p is $\langle Z \rangle$, it consists of the elements Z^k for $k \in \mathbb{Z}_p$. The center has exactly p elements; thus, there are exactly p Type-I elements.

ii. From Theorem 6.2, the Type-II elements are exactly those satisfying $\sigma^2(g) = g^{-1}$. These are precisely the elements of the form $\left(a, b; \frac{ab}{2}\right)$ which total p^2 in number. The number of nonidentity Type-II elements is $p^2 - 1$.

iii. Since H_p has p^3 total elements and the identity is shared by Type-I and Type-II, the number of Type-III elements is
$$p^3 - (p + p^2 - 1) = p^3 - p^2 - p + 1. \quad \square$$

The product of two Type-II elements may not always be a Type-II element. The next result presents a necessary and sufficient condition for this product to be a Type-II element.

Theorem 6.8. *Let $g = (a, b; c)$, $h = (a', b'; c')$ with $a, b, a', b' \neq 0$ be Type-II elements. Then, gh is a Type-II element if and only if g and h belong to the same cyclic subgroup in Theorem 6.6.*

Proof. The necessity follows directly from the subgroup structure established in Theorem 6.6.

Now, we prove sufficiency. Let $g = (a, b; c)$, $h = (a', b'; c')$ with $a, b, a', b' \neq 0$ be Type-II elements so that by Theorem 6.2 *(ii)*, $c = \frac{ab}{2}$, $c' = \frac{a'b'}{2}$ and $gh = (a + a', b + b'; c + c' + ab')$.

From Theorem 6.6,
$$g = \left(a, b; \frac{ab}{2}\right) = \left(1, a^{-1}b; \frac{a^{-1}b}{2}\right)^a.$$

If gh is a Type-II, then by Theorem 6.2 *(ii)*,
$$c + c' + ab' = \frac{(a + a')(b + b')}{2}.$$

Thus, $ab' = a'b$ and we have $a' = ab^{-1}b'$ or $b' = a^{-1}a'b$. Say $b' = a^{-1}a'b$, then
$$h = \left(a', a^{-1}a'b; \frac{(a')^2 a^{-1}b}{2}\right) = \left(1, a^{-1}b; \frac{a^{-1}b}{2}\right)^{a'}.$$

A similar argument can be used if $a' = ab^{-1}b'$ and the result follows. $\square$

CONCLUSION

This paper has developed an algebraic and combinatorial framework for analyzing word length in both the discrete and finite Heisenberg groups.

By constructing two word-length-preserving automorphisms, we demonstrated how these mappings can effectively reduce the domain of group elements, enabling precise classification and explicit word length computation. Through these, we derived exact combinatorial formulas and introduced a meaningful classification of elements into Types I, II, and III, based on their structural properties and their behavior under the word-length-preserving automorphisms.

What sets this work apart is its constructive approach. Instead of relying on complex geometric tools, we focused on elementary group-theoretic methods using generators, relations, and identities. This made it possible to analyze and simplify word representations systematically. Our results contribute both theoretical insights and practical tools that can support further work in group theory, representation theory, and computational mathematics.

Moving forward, we recommend expanding this framework to cover more general Heisenberg groups, and we envision applying these ideas to fields such as algorithm design and cryptography. Additionally, the methods developed here may aid in creating efficient solutions for word problems and optimizing representations in nonabelian settings. The potential for this research to impact both pure and applied mathematics is significant, offering opportunities for further advancements.

ACKNOWLEDGEMENT

The first author gratefully acknowledges the University of the East and the Tan Yan Kee Foundation, Inc. for the scholarship support that made this research possible. The authors also thank the anonymous referees, the associate editor, the managing editor, and the copy editor for their careful review and constructive comments, which significantly improved the quality and presentation of the paper.

REFERENCES

Alekseev, I., & Magdiev, R. (2019). *The language of geodesics for the discrete Heisenberg group*. https://arxiv.org/abs/1905.03226

Blachère, S. (2003). Word distance on the discrete Heisenberg group. *Colloquium Mathematicum, 95*, 21–36.

De la Harpe, P. (2000). *Topics in geometric group theory*. University of Chicago Press.

Duchin, M., & Mooney, C. (2014). Fine asymptotic geometry in the Heisenberg group. *Indiana University Mathematics Journal, 63*(3), 885–916.

Duchin, M., & Shapiro, M. (2014). *Rational growth in the Heisenberg group*. https://arxiv.org/abs/1411.4201

Kahn, P. J. (2005). *Automorphisms of the discrete Heisenberg groups*. http://www.mathlab.cornell.edu/m/sites/default/files/imported/People/Faculty/Heisen.pdf

Long, B., & Wu, W. (2017). Twisted group C*-algebras as compact quantum metric spaces. *Results in Mathematics, 71*, 57–72. https://doi.org/10.1007/s00025-016-0562-7

Osipov, D. V. (2015). Discrete Heisenberg groups and its automorphism group. *Matematicheskie Zametki, 98*(1), 152–155.

Ozawa, N., & Rieffel, M. A. (2005). Hyperbolic group C*-algebras and free-product C*-algebras as compact quantum metric spaces. *Canadian Journal of Mathematics, 57*(5), 1056–1079.

Pate, J. (2006). *Geometric aspects of the Heisenberg groups*. http://math.arizona.edu/ura-reports/061/Pate.John/Final.pdf

Robinson, D. (1996). *A course in the theory of groups* (2nd ed.). Springer-Verlag.

Suzuki, M. (1986). *Group theory: Volume II*. Springer-Verlag.

Appendix A.

Lemma A.1 Let $W = X^{a_1}Y^{b_1}X^{a_2}Y^{b_2}\cdots X^{a_r}Y^{b_r}$ be a reduced word representing a central element. By Lemma 2.3, the third coordinate is $\sum_{i\le j} a_i\, b_j$. Let S be the total of all positive X-exponents and T the total of all positive Y-exponents. Then,

$$\sum_{i\le j} a_i\, b_j \le ST.$$

Proof. Split the sum according to the signs of the exponents: $\sum_{i\le j} a_i\, b_j = PP + NN + NP$, where PP is the sum of all products with $a_i > 0$ and $b_j > 0$, NN is the sum of all products with $a_i < 0$ and $b_j < 0$, and NP is the sum of all remaining mixed-sign products.

Since the total positive X-exponent is S and the total positive Y-exponent is T, the sum of all positive–positive products satisfies $PP \le ST$. If $NN = 0$, then $NP = 0$ and the inequality follows.

Suppose now that $NN > 0$. Then, negative exponents occur in both X and Y. Because the word is central, positive exponents must also occur. Hence, positive and negative exponents coexist in the word, forcing mixed-sign products to appear. These products contribute negatively and reduce the total contribution from potential positive–positive pairings. Therefore, $\sum_{i\le j} a_i\, b_j \le ST$. □

Remark. Negative–negative products do not introduce new positive exponents. The total positive exponents S and T are fixed. When mixed-sign products appear, some positive exponents that could have contributed to positive–positive products are instead paired with negative exponents, yielding negative contributions. Thus, the reduction occurs in the potential positive–positive contribution, not in the positive–positive products themselves. Consequently, the total contribution cannot attain the maximal value ST.