CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (2) (2026), Pages 2638–2652
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Algebraic Cryptanalysis of an Elliptic-Curve-Based Polyalphabetic Hill Cipher over $\mathbb{Z}/26\mathbb{Z}$

Yanuar Bhakti Wira Tama*, Sindi Lukmaini, and Muhammad Firdhausi Fahmi

Department of Mathematics, Institut Teknologi Kalimantan, Balikpapan, East Kalimantan, Indonesia

Abstract

This study revisits a Hill-type polyalphabetic cipher in which the 26 uppercase letters are represented by points of the elliptic curve $y^2 = x^3 + 9x + 1$ over $\mathbb{F}_{29}$. We prove that the selected curve group contains 26 elements and that $B = (0, 1)$ has order 26; hence the group is cyclic and isomorphic to $\mathbb{Z}_{26}$. Under the generator-coordinate map, the point-valued matrix transformation is therefore conjugate to the classical Hill transformation over $\mathbb{Z}_{26}$. This characterization yields explicit plaintext-based attacks. Under the stated fixed-key and public-mapping assumptions, m chosen plaintext blocks recover an $m \times m$ key exactly. For the reported 4×4 key, a known-plaintext attack also recovers the key from the first four plaintext-ciphertext block pairs because the resulting plaintext matrix is invertible modulo 26. For the Indonesian Constitution benchmark, a reproducible experiment using 100 random invertible keys for each block size $m = 2, 3, 4, 5$ shows that mean ciphertext entropy increases from 4.5592 to 4.6769 bits per symbol and the index of coincidence approaches $1/26$, while chosen-plaintext key recovery remains 100%. Thus, flatter symbol frequencies do not remove the linear weakness of the construction or provide elliptic-curve discrete-logarithm security.

Keywords: Algebraic cryptanalysis; Chosen-plaintext attack; Elliptic-curve group; Hill cipher; Known-plaintext attack

Copyright © 2026 by Authors, Published by CAUCHY Group. This is an open access article under the CC BY-SA License (https://creativecommons.org/licenses/by-sa/4.0)

1. Introduction

Matrix-based cryptography remains a useful meeting point between linear algebra, modular arithmetic, and cryptographic design. The classical Hill cipher introduced by Hill [1] encrypts blocks by an invertible matrix over a residue ring. Its mathematical simplicity makes encryption and decryption transparent, but the same linearity also creates a fundamental weakness: sufficiently informative plaintext-ciphertext relations can be converted into linear equations for the key. Recent studies therefore continue to modify the Hill framework through alternative key-generation mechanisms, recursive matrices, larger alphabets, affine terms, substitution-permutation layers, and other algebraic constructions [2–8].

Elliptic curves provide a different algebraic structure. In modern public-key cryptography, their security role is normally associated with computational problems such as the elliptic-curve discrete logarithm problem [9]. A number of recent encryption schemes combine Hill-type matrix operations with elliptic-curve mechanisms for point encoding, key exchange, key-matrix generation, or image encryption [10–13]. These combinations can produce large key spaces or

*Corresponding author. E-mail: yanuar.bhakti@lecturer.itk.ac.id

Received: September 8, 2026 Revised: October 29, 2026 Accepted: November 6, 2026
DOI: https://doi.org/10.18860/cauchy.v11i2.45911

favorable statistical properties, but the presence of elliptic-curve points alone does not guarantee that the resulting transformation inherits standard ECC security. Lone et al. [10], Wen et al. [14], and Saeed et al. [15], for example, demonstrate that apparently complex matrix-based or hybrid constructions may still admit effective cryptanalysis when deterministic algebraic relations remain exposed.

The distinction between statistical appearance and algebraic security is especially important for deterministic ciphers. Histogram flattening, high symbol entropy, or low correlation can indicate that obvious frequency patterns have been reduced, yet such metrics are not proofs of resistance to known-plaintext or chosen-plaintext attacks. Contemporary cryptanalytic work increasingly evaluates the structural relation between plaintext, ciphertext, and secret parameters rather than relying only on output statistics [14–16]. Recent mathematical constructions using elliptic-curve point-valued blocks likewise emphasize explicit group actions, invertibility, correctness, and limitations of deterministic transformations [17].

Closely related studies by Tama and collaborators provide direct context for the present analysis. Tama and Fahmi [18] constructed a classical cryptosystem by mapping 26 alphabetic symbols to elliptic-curve points through the order structure of a Montgomery-form curve. Lukmaini et al. [19] combined the Hill cipher with a permutation layer and evaluated encryption and decryption computation time, while Tama and Mujahidin [20] applied a modified Hill approach based on elliptic curves and matrices to medical images and assessed entropy together with computational time. More recently, Tama and Mujahidin [21] studied grayscale image encryption through sub-matrix multiplication over $\mathbb{Z}_{256}$ and recovered inverse key matrices by Gauss–Jordan elimination. That work provides a modular matrix-cryptography precursor in which invertibility and ciphertext distribution are central implementation concerns, whereas the present study asks a different structural question: whether the algebraic form of the fixed transformation permits exact key recovery despite apparently favorable output statistics. Taken together, these studies motivate the present examination of whether an elliptic-curve representation or improved statistical behavior changes the underlying algebraic exposure of a Hill-type transformation.

Target scheme, research gap, novelty, and contributions. The target of this cryptanalysis is the specific deterministic point-valued Hill construction completely specified in this manuscript by the curve in Eq. (1), the public alphabet-to-point bijection in Table 1, and the fixed-key point-valued matrix action in Eq. (9). It is characterized by three features: (i) the curve $E/\mathbb{F}_{29}$ has 26 group elements; (ii) a fixed public bijection maps the 26 uppercase letters onto all elements of $E(\mathbb{F}_{29})$; and (iii) a fixed secret matrix $K \in \mathrm{GL}_m(\mathbb{Z}_{26})$ acts on point blocks through elliptic-curve scalar multiplication and point addition. The related constructions in Refs. [18–21] provide methodological context but are not asserted to be the exact target of the present cryptanalysis. No separately published cryptosystem is claimed as the exact target; rather, the construction is fully specified here and the work is positioned as a structural security analysis of this deterministic point-valued Hill transformation. The conclusions of this paper concern this concrete construction and not elliptic-curve/Hill-type schemes in general.

The research gap is the absence of a structural cryptanalysis that connects this point-valued transformation to its generator-coordinate representation and derives the corresponding key-recovery conditions. The novelty is therefore not the standard observation that a classical Hill cipher is linear. In the classical Hill setting, residue coordinates are available from the outset; here, the nontrivial cryptanalytic step is to prove that the public point alphabet exhausts a cyclic group of order 26, construct the resulting public generator-coordinate isomorphism, and translate the point-valued encryption oracle back into explicit linear key-recovery equations on actual letter blocks. We then instantiate those equations for both chosen-plaintext and known-plaintext settings, including recovery of the concrete 4×4 key. Beyond the isomorphism itself, the independent contribution is the derivation and proof of explicit recovery conditions, their concrete instantiation in both attack models, and computational verification across multiple plaintext

datasets. The scientific contributions are fourfold: (i) a proof that the selected curve group is cyclic of order 26 with generator $(0, 1)$; (ii) a correctness result and an algebraic-equivalence theorem for the point-valued matrix action; (iii) explicit sufficient conditions for chosen-plaintext and known-plaintext key recovery, together with a more general invertible-coordinate condition for chosen plaintexts; and (iv) a reproducible 400-key computational verification evaluated on three public plaintext datasets, separating descriptive ciphertext statistics from algebraic security.

2. Methods

The methods combine an explicit elliptic-curve group construction with algebraic attack conditions and computational verification. We first specify the group operation and public symbol mapping, then define the point-valued matrix action, attack models, and reproducible experiment.

2.1. Algebraic setting

Let

$$E : y^2 \equiv x^3 + 9x + 1 \pmod{29} \tag{1}$$

be an elliptic curve over $\mathbb{F}_{29}$. Throughout the manuscript, $\mathbb{Z}_{26}$ denotes the residue ring $\mathbb{Z}/26\mathbb{Z}$, and O denotes the identity point of the elliptic-curve group. Its discriminant is nonzero modulo 29 because $4(9)^3 + 27(1)^2 \not\equiv 0 \pmod{29}$, so Eq. (1) defines a nonsingular elliptic curve. The set $E(\mathbb{F}_{29})$ consists of the affine solutions of Eq. (1) together with the identity point O and forms an abelian group under the usual elliptic-curve point addition.

2.2. Elliptic-curve group operation

For completeness, the group operation used throughout the construction is stated explicitly. Let

$$E_{a,b} : y^2 \equiv x^3 + ax + b \pmod{p}$$

be a nonsingular elliptic curve over the prime field $\mathbb{F}_p$. For distinct points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ with $P \neq -Q$, define

$$\mu \equiv (y_2 - y_1)(x_2 - x_1)^{-1} \pmod{p}. \tag{2}$$

Using the slope defined in Eq. (2), $P + Q = (x_3, y_3)$, where

$$x_3 \equiv \mu^2 - x_1 - x_2 \pmod{p}, \tag{3}$$
$$y_3 \equiv \mu(x_1 - x_3) - y_1 \pmod{p}. \tag{4}$$

For point doubling with $Q = P$ and $y_1 \neq 0$, the slope is

$$\mu \equiv (3x_1^2 + a)(2y_1)^{-1} \pmod{p}, \tag{5}$$

Using the doubling slope in Eq. (5), the coordinates are again obtained from Eq. (3) and Eq. (4). The identity point O (the point at infinity) satisfies $P + O = P$, while $-P = (x_1, -y_1)$ and $P + (-P) = O$. For an integer k, the notation $[k]P$ denotes scalar multiplication by repeated point addition [9].

For the specific curve in Eq. (1) and $B = (0, 1)$, the doubling slope is

$$\mu \equiv 9(2^{-1}) \equiv 9(15) \equiv 19 \pmod{29},$$

Thus

$$x_{2B} \equiv 19^2 \equiv 13 \pmod{29}, \qquad y_{2B} \equiv 19(0 - 13) - 1 \equiv 13 \pmod{29},$$

and hence $2B = (13, 13)$. This is exactly the point assigned to the letter J in Table 1, providing a direct check of the group operation used by the implementation.

2.3. Public symbol representation

The construction analyzed here uses a public bijection

$$\varphi : \Sigma \longrightarrow E(\mathbb{F}_{29}), \qquad \Sigma = \{A, B, \ldots, Z\}, \tag{6}$$

shown in Table 1. The point O is assigned to A, and $(0, 1)$ is assigned to B. The table also lists the generator coordinate $\lambda(P)$ relative to $B = (0, 1)$, obtained after the group structure is established in Subsection 3.1.

Table 1: Public alphabet-to-point mapping and generator coordinates

Letter	λ	Point	Letter	λ	Point
A	0	O	N	7	$(16, 23)$
B	1	$(0, 1)$	O	12	$(19, 10)$
C	25	$(0, 28)$	P	14	$(19, 19)$
D	22	$(7, 1)$	Q	3	$(22, 1)$
E	4	$(7, 28)$	R	23	$(22, 28)$
F	20	$(8, 11)$	S	17	$(24, 11)$
G	6	$(8, 18)$	T	9	$(24, 18)$
H	8	$(9, 12)$	U	15	$(26, 11)$
I	18	$(9, 17)$	V	11	$(26, 18)$
J	2	$(13, 13)$	W	16	$(27, 2)$
K	24	$(13, 16)$	X	10	$(27, 27)$
L	13	$(14, 0)$	Y	5	$(28, 7)$
M	19	$(16, 6)$	Z	21	$(28, 22)$

$$\begin{array}{ccccc} \Sigma & \xrightarrow{\varphi} & E(\mathbb{F}_{29}) & \xrightarrow{\lambda} & \mathbb{Z}_{26} \\ \Sigma & \xleftarrow{\varphi^{-1}} & E(\mathbb{F}_{29}) & \xleftarrow{\psi} & \mathbb{Z}_{26} \end{array}$$

Fig. 1: Public symbol representation and generator-coordinate view. After Proposition 1, $\lambda(P)$ is the unique coordinate satisfying $P = [\lambda(P)]B$, and $\psi = \lambda^{-1}$.

The schematic in Fig. 1 anticipates the structural result proved in Subsection 3.1: the elliptic-curve representation is public, and once B is identified as a generator, every point has a unique coordinate in $\mathbb{Z}_{26}$. This coordinate view is the bridge between the point-valued implementation and the algebraic cryptanalysis developed later.

2.4. Hill-type action on elliptic-curve point blocks

For a block size m, let

$$K = (k_{ij}) \in \mathrm{GL}_m(\mathbb{Z}_{26}), \tag{7}$$

where $\mathrm{GL}_m(\mathbb{Z}_{26})$ denotes the invertible $m \times m$ matrices over $\mathbb{Z}_{26}$. The key-space requirement in Eq. (7) is equivalently characterized by

$$\gcd(\det K, 26) = 1. \tag{8}$$

A plaintext block $P = (P_1, \ldots, P_m)^{\mathsf{T}} \in E(\mathbb{F}_{29})^m$ is transformed by

$$T_K(P)_i = \sum_{j=1}^{m} [k_{ij}]P_j, \qquad i = 1, \ldots, m, \tag{9}$$

where $[k]P$ denotes elliptic-curve scalar multiplication and the summation uses the elliptic-curve group law. A ciphertext point block is converted back to letters using φ^{-1}.

For the 4×4 case, the following invertible key matrix is considered:

$$K_4 = \begin{pmatrix} 3 & 3 & 1 & 13 \\ 1 & 2 & 17 & 15 \\ 23 & 14 & 12 & 0 \\ 10 & 14 & 24 & 5 \end{pmatrix} \pmod{26}. \tag{10}$$

For example, the plaintext block `BAHW` is mapped to points and transformed by Eq. (9), yielding the ciphertext block `VLUD`. The inverse matrix returns `VLUD` to `BAHW`.

2.5. Attack models

Two standard plaintext-based attack models are considered.

Definition 1 (Known-plaintext setting)**.** The adversary knows the public curve, the fixed symbol-point mapping φ, the block size m, and plaintext blocks together with their corresponding ciphertext blocks encrypted under the same fixed secret matrix K.

Definition 2 (Chosen-plaintext setting)**.** The adversary knows the public parameters and may submit plaintext blocks of its choice to an encryption oracle that uses the same fixed secret matrix K, observing the corresponding ciphertext blocks. The goal is to recover K.

Throughout the remainder of the manuscript, CPA and KPA denote the chosen-plaintext and known-plaintext attack models, respectively, while *key recovery* denotes the corresponding recovery procedure under the stated algebraic conditions. These models are particularly relevant to deterministic Hill-type transformations because repeated encryption under a fixed matrix preserves exact algebraic relations. Modern cryptanalysis of Hill-derived systems frequently uses related plaintext relations to expose equivalent keys or keystreams [14, 15].

Assumptions and scope of the attacks. The attacks derived in this study require a fixed deterministic matrix key and the fixed public mapping in Eq. (6). The simple m-query chosen-plaintext recovery in Theorem 2 uses the additional fact that the public alphabet contains symbols representing both the identity point O and the generator B, so that the standard basis vectors can be submitted directly to the encryption oracle. This is a sufficient condition for the particularly simple basis-query attack, not a claim that such symbols are necessary for every possible chosen-plaintext attack. As stated explicitly after Theorem 2, any m chosen plaintext blocks whose generator-coordinate matrix is invertible also recover the key through $K = YX^{-1}$. The known-plaintext recovery does not require oracle access, but it likewise requires m aligned plaintext-ciphertext blocks whose plaintext coordinate matrix is invertible modulo 26; with more known pairs, an attacker may search for an invertible m-block subset. The conclusions therefore apply to the fixed deterministic construction under these stated algebraic conditions. They do not automatically extend to constructions with randomized encryption, nonce-dependent or secret symbol mappings, changing matrix keys, nonlinear layers, or other mechanisms that destroy the fixed public linear relation used in the proofs.

2.6. Computational experiment and reproducibility

To test whether the descriptive ciphertext statistics are specific to one plaintext, the revised experiment uses three public benchmark texts: (i) the Preamble of the 1945 Constitution of Indonesia from JDIH Sekretariat Jenderal DPR RI [22]; (ii) the U.S. Declaration of Independence transcription from the U.S. National Archives [23]; and (iii) the U.S. Bill of Rights transcription from the U.S. National Archives [24]. After spaces, punctuation, and nonalphabetic characters are removed and letters are converted to uppercase, the first 1211 symbols of each source are

retained. This matches the cleaned length of the Indonesian benchmark and controls plaintext length across datasets. The character `X` is used for deterministic padding when the length is not divisible by m.

For each $m \in \{2, 3, 4, 5\}$, 100 matrices are generated by drawing entries independently and uniformly from $\{0, \ldots, 25\}$ and rejecting matrices until Eq. (8) is satisfied. The same 100 matrices at each block size are applied to all three plaintext datasets, so plaintext effects can be compared without changing the sampled keys. The experiment therefore contains 400 distinct invertible keys and $400 \times 3 = 1200$ encryption-decryption evaluations. A fixed pseudorandom seed, 20260902, is used. Each of the 400 keys is also tested once with the chosen-plaintext key-recovery procedure derived in Subsection 3.2. Because attack success follows deterministically from Theorem 2 under its assumptions, these recovery runs provide computational verification of the theorem rather than independent empirical security evidence.

Three single-symbol statistics are calculated for each ciphertext. For symbol frequencies f_i and total length N, Shannon entropy is

$$H = -\sum_{i=1}^{26} p_i \log_2 p_i, \qquad p_i = \frac{f_i}{N}, \tag{11}$$

the Pearson statistic relative to a uniform frequency model is

$$\chi^2 = \sum_{i=1}^{26} \frac{(f_i - N/26)^2}{N/26}, \tag{12}$$

and the index of coincidence is

$$\mathrm{IC} = \frac{\sum_{i=1}^{26} f_i(f_i - 1)}{N(N-1)}. \tag{13}$$

The statistics in Eq. (11), Eq. (12), and Eq. (13) are evaluated for every ciphertext in the experiment. The maximum entropy for a 26-symbol alphabet is $\log_2 26 \approx 4.70044$, while the uniform-distribution reference for the index of coincidence is $1/26 \approx 0.03846$.

The experiment was run with Python 3.13.5 using a standalone script that implements the curve group law, modular matrix inversion, encryption, decryption, random-key generation, statistics, and key-recovery tests without external packages. The reproducibility materials supplied with the manuscript are cross-referenced as follows: Supplementary File S1 contains the standalone Python script; S2 contains all 1200 row-level encryption–decryption results; S3 contains the CPA key-recovery verification for all 400 sampled keys; S4 contains the per-dataset and per-block-size summary statistics; and S5 contains the execution log. The row-level values underlying Table 3 are provided in S2, while the 400/400 CPA verification reported in the Results is documented in S3. The workflow is summarized in Algorithm 1.

3. Results and Discussion

This section presents the structural and computational results in a sequence that follows the cryptanalytic argument. First, the group structure of the selected elliptic curve is established and the point-valued transformation is related to its generator-coordinate representation over $\mathbb{Z}_{26}$. The resulting linear relation is then used to derive chosen-plaintext and known-plaintext key-recovery procedures. The theoretical findings are subsequently verified across three public plaintext datasets, followed by a discussion of computational cost, security interpretation, and the relation of the construction to recent matrix and elliptic-curve approaches. The overall algebraic reduction and recovery mechanism are summarized in Fig. 2.

3.1. Group structure and algebraic equivalence

Direct enumeration of Eq. (1) gives 25 affine points. Including O, the group therefore has order 26. The public mapping in Table 1 uses every group element exactly once.

Algorithm 1 Multi-text random-key evaluation and cryptanalysis workflow

Require: Cleaned texts $\mathcal{M} = \{M_1, M_2, M_3\}$ of 1211 symbols, sizes $\{2, 3, 4, 5\}$, 100 keys per size, seed 20260902

for $m \in \{2, 3, 4, 5\}$ **do**
  for $r = 1$ to 100 **do**
    Sample $K \in \mathbb{Z}_{26}^{m \times m}$ until $\gcd(\det K, 26) = 1$
    Execute the chosen-plaintext recovery of Theorem 2 once for K
    for each $M_d \in \mathcal{M}$ **do**
      Encrypt M_d using the point-valued action T_K
      Decrypt with K^{-1} and verify exact padded-plaintext recovery
      Record H, χ^2, IC, and decryption correctness
    end for
  end for
end for
return Row-level results, key-recovery verification, and summary statistics

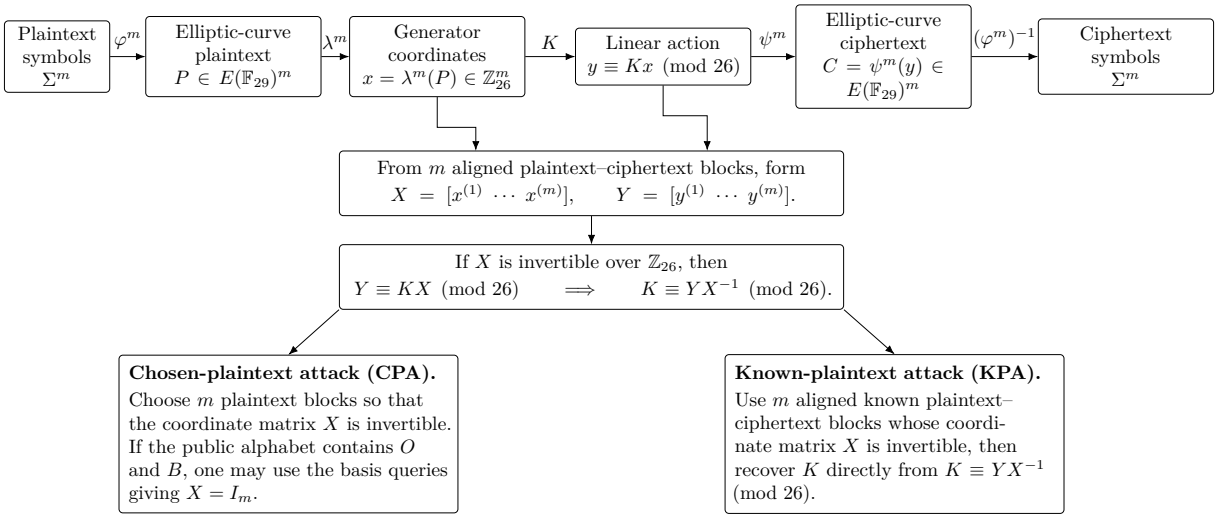


Fig. 2: Schematic of the algebraic cryptanalysis framework. The public elliptic-curve representation is converted to generator coordinates over $\mathbb{Z}_{26}$, where the point-valued transformation becomes the linear relation $y \equiv Kx \pmod{26}$. From m aligned coordinate pairs, the key matrix is recovered as $K \equiv YX^{-1} \pmod{26}$ whenever X is invertible, under either chosen-plaintext or known-plaintext conditions.

Proposition 1. *The point $B = (0, 1)$ has order 26 in $E(\mathbb{F}_{29})$. Consequently, $E(\mathbb{F}_{29})$ is cyclic and*

$$E(\mathbb{F}_{29}) \cong \mathbb{Z}_{26}.$$

Proof. Since $|E(\mathbb{F}_{29})| = 26$, Lagrange's theorem gives $\operatorname{ord}(B) \mid 26$. The possible orders are $1, 2, 13$, and 26. Clearly $B \neq O$. Because the y-coordinate of B is nonzero, $B \neq -B$, so $2B \neq O$. Repeated point addition gives

$$13B = (14, 0) \neq O.$$

Hence the order is neither 1, 2, nor 13, and therefore $\operatorname{ord}(B) = 26$. □

Let

$$\psi : \mathbb{Z}_{26} \longrightarrow E(\mathbb{F}_{29}), \qquad \psi(a) = [a]B, \tag{14}$$

which is a group isomorphism by Proposition 1. Its inverse

$$\lambda = \psi^{-1} : E(\mathbb{F}_{29}) \longrightarrow \mathbb{Z}_{26} \tag{15}$$

assigns to each curve point its generator coordinate. The mutually inverse maps in Eq. (14) and Eq. (15) provide the coordinate conversion used throughout the cryptanalysis. The λ values for the public alphabet mapping appear in Table 1. Because the group has only 26 elements, this coordinate table is trivial to construct by repeated addition; no cryptographically hard discrete-logarithm problem is present in this parameter choice.

Theorem 1 (Algebraic equivalence)**.** *Let $K \in \mathbb{Z}_{26}^{m\times m}$ and let T_K be the point-valued transformation in Eq. (9). For any $P = (P_1, \ldots, P_m)^{\mathsf{T}} \in E(\mathbb{F}_{29})^m$,*

$$\lambda^m(T_K(P)) \equiv K\,\lambda^m(P) \pmod{26}, \tag{16}$$

where λ^m is applied componentwise. Equivalently,

$$T_K = (\psi^m) \circ K \circ (\lambda^m). \tag{17}$$

Thus the point-valued cipher is conjugate, through a public coordinate relabeling, to the classical Hill matrix action over $\mathbb{Z}_{26}$.

Proof. Write $P_j = [x_j]B$ with $x_j = \lambda(P_j)$. For the ith output component,

$$T_K(P)_i = \sum_{j=1}^{m} [k_{ij}]P_j = \sum_{j=1}^{m} [k_{ij}x_j]B = \left[\sum_{j=1}^{m} k_{ij}x_j\right] B.$$

Applying λ gives $\lambda(T_K(P)_i) \equiv \sum_j k_{ij}x_j \pmod{26}$. Collecting the m components yields Eq. (16), and Eq. (17) follows from $\psi = \lambda^{-1}$. □

Corollary 1 (Correctness)**.** *If $K \in \mathrm{GL}_m(\mathbb{Z}_{26})$, then T_K is bijective and*

$$T_{K^{-1}} \circ T_K = \mathrm{id}_{E(\mathbb{F}_{29})^m}.$$

Proof. By Theorem 1, applying T_K corresponds in generator coordinates to multiplication by K. Since K is invertible modulo 26, multiplication by K^{-1} returns every coordinate vector to its original value. Mapping back through ψ^m proves the result. □

Theorem 1 changes the interpretation of the scheme. It is important to distinguish the use of elliptic-curve arithmetic from elliptic-curve cryptographic security. The construction genuinely performs point addition and scalar multiplication, but its security is not reduced to solving an ECDLP in a large cryptographic subgroup. Here, the entire group contains only 26 elements, the generator-coordinate map is public, and its complete coordinate table can be constructed exhaustively by repeated addition. Consequently, the elliptic-curve layer acts as a public algebraic representation of the alphabet; after a 26-entry coordinate lookup, the encryption relation is linear over $\mathbb{Z}_{26}$. Thus the construction does not inherit the computational hardness normally associated with standard ECC merely because elliptic-curve operations appear in the implementation. This distinction is consistent with recent work that separates elliptic-curve point-valued algebraic actions from claims of modern semantic security [17].

3.2. Chosen-plaintext key recovery

The public mapping has a particularly useful pair of symbols: $\varphi(A) = O = [0]B$ and $\varphi(B) = B = [1]B$. This allows standard basis vectors to be submitted directly as letter blocks.

Theorem 2 (Chosen-plaintext recovery)**.** *For an $m \times m$ fixed key matrix K in the construction above, an adversary with an encryption oracle recovers all entries of K using m chosen plaintext blocks, provided the public mapping contains symbols for O and the generator B.*

Proof. For $j = 1, \ldots, m$, choose the plaintext point block $P^{(j)} = (O, \ldots, O, B, O, \ldots, O)^{\mathsf{T}}$ with B in position j. In generator coordinates, $\lambda^m(P^{(j)}) = e_j$, the jth standard basis vector. By Eq. (16),

$$\lambda^m(T_K(P^{(j)})) = Ke_j,$$

which is exactly the jth column of K. Hence one oracle query recovers one complete key column, and m queries recover K. □

General chosen-plaintext condition. The basis-query construction above is the simplest special case because its plaintext coordinate matrix is the identity. More generally, let m chosen plaintext blocks form an invertible coordinate matrix X and let the corresponding ciphertext coordinate matrix be Y. By Theorem 1, $Y = KX$, and therefore

$$K = YX^{-1} \pmod{26}.$$

Thus, direct alphabet symbols for both O and B make the m-query attack immediate, but they are not necessary if the adversary can choose another set of blocks whose public generator-coordinate matrix is invertible. Conversely, if the available chosen blocks cannot produce an invertible coordinate matrix, this particular linear-recovery argument does not by itself establish key recovery.

Table 2 applies the theorem to the examined matrix in Eq. (10). The ciphertext letters are first converted using the public λ table. The resulting vectors are precisely the four columns of K_4.

Table 2: Chosen-plaintext recovery of the examined 4×4 key

Query	Plaintext	Ciphertext	Generator-coordinate ciphertext
1	`BAAA`	`QBRX`	$(3, 1, 23, 10)^{\mathsf{T}}$
2	`ABAA`	`QJPP`	$(3, 2, 14, 14)^{\mathsf{T}}$
3	`AABA`	`BSOK`	$(1, 17, 12, 24)^{\mathsf{T}}$
4	`AAAB`	`LUAY`	$(13, 15, 0, 5)^{\mathsf{T}}$

Reading these four coordinate vectors as columns gives

$$\begin{pmatrix} 3 & 3 & 1 & 13 \\ 1 & 2 & 17 & 15 \\ 23 & 14 & 12 & 0 \\ 10 & 14 & 24 & 5 \end{pmatrix},$$

which is exactly Eq. (10). The attack does not search the nominal key space and does not solve an elliptic-curve discrete logarithm of cryptographic size. It uses four one-block queries and a 26-element public lookup table.

3.3. Known-plaintext key recovery from the test message

Chosen-plaintext access is not necessary when a sufficient set of known plaintext blocks is already available.

Theorem 3 (Known-plaintext recovery). *Let $P^{(1)}, \ldots, P^{(m)} \in E(\mathbb{F}_{29})^m$ be known plaintext blocks and $C^{(1)}, \ldots, C^{(m)}$ their ciphertext blocks under the same fixed key K. Form the coordinate matrices*

$$X = [\lambda^m(P^{(1)}) \ \cdots \ \lambda^m(P^{(m)})], \quad Y = [\lambda^m(C^{(1)}) \ \cdots \ \lambda^m(C^{(m)})].$$

If $X \in \mathrm{GL}_m(\mathbb{Z}_{26})$, then

$$K \equiv YX^{-1} \pmod{26}. \tag{18}$$

Proof. By Theorem 1, each pair satisfies $y^{(j)} = Kx^{(j)}$. Stacking the vectors as columns gives $Y = KX$. If X is invertible modulo 26, right multiplication by X^{-1} yields Eq. (18). □

For the examined 4×4 case, the first four plaintext blocks are `BAHW`, `ASES`, `UNGG`, and `UHNY`, while the corresponding ciphertext blocks are `VLUD`, `WMAQ`, `FLZG`, and `VSZL`. Using Table 1 gives

$$X = \begin{pmatrix} 1 & 0 & 15 & 15 \\ 0 & 17 & 7 & 8 \\ 8 & 4 & 6 & 7 \\ 16 & 17 & 6 & 5 \end{pmatrix}, \qquad Y = \begin{pmatrix} 11 & 16 & 20 & 11 \\ 13 & 19 & 13 & 17 \\ 15 & 0 & 21 & 21 \\ 22 & 3 & 6 & 13 \end{pmatrix}. \tag{19}$$

For the coordinate matrices in Eq. (19), the determinant of X satisfies

$$\det X \equiv 19 \pmod{26}, \qquad \gcd(19, 26) = 1,$$

so X is invertible. Its inverse is

$$X^{-1} \equiv \begin{pmatrix} 23 & 19 & 20 & 13 \\ 6 & 6 & 25 & 5 \\ 14 & 25 & 7 & 7 \\ 14 & 24 & 9 & 6 \end{pmatrix} \pmod{26}.$$

Substitution in Eq. (18) recovers K_4 exactly. Therefore, for this 4×4 example, four consecutive plaintext-ciphertext blocks from the beginning of the test message are sufficient for key recovery under the condition of Theorem 3.

This result is stronger than a generic statement that the Hill cipher is "vulnerable to known plaintext." It shows the actual coordinate matrices generated by the public mapping, verifies invertibility over the composite ring $\mathbb{Z}_{26}$, and reconstructs the specific key used in the implementation.

3.4. Computational verification across three plaintext datasets

The revised experiment evaluates the same 400 invertible keys on three plaintext datasets of equal cleaned length (1211 symbols). The primary Indonesian benchmark retains the numerical values reported in the original experiment, while the two additional public English texts test whether the descriptive trend is peculiar to that single plaintext. Table 3 summarizes the results; the complete row-level observations are provided in Supplementary File S2 and the corresponding summary table is provided in Supplementary File S4.

The three plaintexts begin with different single-symbol distributions: their plaintext entropies are 3.8382, 4.0931, and 4.0926 bits per symbol, respectively, and their plaintext IC values are 0.09202, 0.07057, and 0.07008. Despite these differences, the same qualitative pattern occurs for all three datasets. Moving from $m = 2$ to $m = 4$ or $m = 5$ raises mean ciphertext entropy toward the 26-symbol maximum $\log_2 26 \approx 4.70044$ and moves mean IC toward the uniform reference $1/26 \approx 0.03846$. The precise numerical values depend on the plaintext, which is why the table reports the datasets separately rather than pooling them into a single security statistic.

Table 3: Ciphertext statistics over three plaintext datasets and 100 keys per matrix size

Dataset	m	Entropy (mean ± SD)	χ^2 (mean ± SD)	IC (mean ± SD)	Dec.
Indonesia 1945	2	4.5592 ± 0.0647	253.88 ± 129.68	0.04573 ± 0.00412	100%
	3	4.6599 ± 0.0151	68.66 ± 26.02	0.03985 ± 0.00083	100%
	4	4.6728 ± 0.0084	46.41 ± 14.07	0.03914 ± 0.00045	100%
	5	4.6769 ± 0.0072	39.86 ± 12.49	0.03893 ± 0.00040	100%
U.S. Declaration	2	4.6153 ± 0.0494	149.24 ± 89.61	0.04241 ± 0.00285	100%
	3	4.6715 ± 0.0127	48.65 ± 20.89	0.03921 ± 0.00066	100%
	4	4.6811 ± 0.0070	32.28 ± 11.52	0.03869 ± 0.00037	100%
	5	4.6837 ± 0.0047	28.14 ± 7.77	0.03856 ± 0.00025	100%
U.S. Bill of Rights	2	4.6134 ± 0.0476	147.89 ± 81.65	0.04236 ± 0.00259	100%
	3	4.6714 ± 0.0138	48.68 ± 23.04	0.03921 ± 0.00073	100%
	4	4.6830 ± 0.0041	29.09 ± 6.90	0.03859 ± 0.00022	100%
	5	4.6830 ± 0.0052	29.33 ± 8.76	0.03860 ± 0.00028	100%

All 1200 encryption-decryption evaluations returned the padded plaintext exactly. Separately, all 400 distinct sampled keys were recovered by the chosen-plaintext procedure of Theorem 2. The latter is expected from the theorem and is reported as an implementation verification, not as an estimate of attack probability. The multi-text results strengthen the more limited statistical observation: favorable entropy, χ^2, and IC values are seen for several plaintext distributions, yet the algebraic key-recovery condition is unchanged because it depends on the fixed public linear structure rather than on the plaintext frequency profile. Accordingly, these statistics describe ciphertext appearance but do not establish resistance to known-plaintext or chosen-plaintext cryptanalysis. Similar distinctions between favorable statistical tests and exploitable algebraic relations have motivated recent cryptanalyses of modified Hill-based schemes [10, 14, 15].

3.5. Computational cost and security interpretation

For one point block of length m, Eq. (9) computes m output points, each as a combination of m scalar multiples. Thus a direct implementation requires m^2 scalar multiplications and at most $m(m-1)$ point additions per block. For a message of L symbols, approximately L/m blocks are processed, giving $O(Lm \log 26)$ group additions when double-and-add scalar multiplication is counted explicitly, or $O(Lm)$ for this fixed 26-element group. Decryption has the same per-message cost after K^{-1} has been computed. Matrix inversion can be performed in polynomial time; because $26 = 2 \cdot 13$, it may also be handled through the Chinese remainder decomposition into computations modulo 2 and modulo 13.

Increasing m therefore increases diffusion across more symbols but also increases matrix-action cost. More importantly, it does not change the number of chosen plaintext blocks needed for exact key recovery: the attack cost grows only linearly with m in oracle queries. Within the fixed deterministic construction studied here, no choice among $m = 2, 3, 4, 5$ changes the linear relation established by Theorem 1. This statement is not intended to cover schemes that alter the mapping, key, or encryption rule between blocks.

The role of the elliptic curve must consequently be stated precisely. The points and group law provide an alternative realization of symbol blocks, but for the selected parameters the group is cyclic of composite order 26 and its complete discrete-log table contains only 26 entries. Standard ECC security instead requires carefully selected large subgroups and protocols whose hardness is actually reduced to an ECDLP-type problem [9]. Recent hybrid schemes typically use elliptic curves for key establishment or large-field/subgroup operations rather than merely as a deterministic public symbol representation [12, 13]. The construction studied here should therefore be understood as a mathematical Hill-type transformation over an elliptic-curve group, not as an ECC-secure public-key cryptosystem.

3.6. Relation to recent matrix and elliptic-curve constructions

The present result also clarifies its position relative to newer matrix-based designs. Recursive or structured matrices can enlarge or organize the key-generation space [2–4, 7], while high-order random-matrix generation can improve the practicality of obtaining invertible Hill keys [5]. In related work over a composite modulus, Tama and Mujahidin [21] used sub-matrix multiplication and Gauss–Jordan inversion in $\mathbb{Z}_{256}$ for grayscale image encryption. These mechanisms alter the representation, block organization, or key-generation procedure, but they do not automatically remove linear key-recovery relations if the final encryption equation remains a fixed linear transformation.

Coggins [8] instead embeds multidimensional Hill mixing inside a substitution-permutation network designed to introduce nonlinear and round-dependent behavior. Pandia et al. [25] similarly motivate probabilistic or hybrid mechanisms to address the deterministic weaknesses of classical matrix encryption. These directions illustrate what is missing from the fixed transformation analyzed in this paper: randomness, nonce dependence, nonlinear layers, or a formal modern security construction.

The point-valued formulation of Akinci and Uygun [17] is algebraically close in spirit because scalar matrices act on elliptic-curve point blocks. Their formulation, however, deliberately distinguishes a large elliptic-curve base field from a selected prime-order subgroup and analyzes deterministic limitations. The present paper addresses a different and more specific question: what occurs when an alphabet of 26 symbols exhausts a cyclic group of order 26 and a fixed Hill matrix acts on those points. The resulting isomorphism to $\mathbb{Z}_{26}$ permits exact attacks on the concrete mapping and key considered in this study.

4. Conclusion

This study analyzes a deterministic Hill-type cipher whose symbols are represented by all 26 elements of the elliptic-curve group defined by $y^2 = x^3+9x+1$ over $\mathbb{F}_{29}$. The main theoretical result is that $(0, 1)$ has order 26, so the group is cyclic and isomorphic to $\mathbb{Z}_{26}$. Under the corresponding public generator-coordinate map, the elliptic-curve point transformation is conjugate to ordinary Hill matrix multiplication modulo 26. The presence of elliptic-curve arithmetic therefore does not, for this parameterization, create the large-group ECDLP hardness associated with standard ECC.

Two concrete cryptanalytic consequences were established under explicit assumptions. First, when the same fixed matrix key is reused and the public mapping exposes symbols for both the identity and the generator, an $m \times m$ key can be recovered with m chosen plaintext blocks. Second, in the known-plaintext setting, key recovery follows whenever m aligned plaintext-ciphertext blocks produce an invertible plaintext coordinate matrix modulo 26; the examined 4×4 case satisfies this condition and its key is recovered from the first four block pairs.

The revised computational experiment applies the same 400 sampled keys to three public plaintext datasets of equal cleaned length, yielding 1200 encryption-decryption evaluations, all of which recover the padded plaintext exactly. Across all three datasets, larger tested block sizes generally move entropy and index-of-coincidence values toward their uniform-distribution references, although the exact values remain plaintext-dependent. The chosen-plaintext recovery succeeds for all 400 keys as predicted by the theorem. Thus the computation serves both as implementation verification and as a multi-text illustration that favorable ciphertext statistics can coexist with exact algebraic key recovery; the statistics do not, by themselves, establish resistance to structural cryptanalysis.

The conclusions are intentionally limited to the fixed deterministic construction analyzed here. They should not be generalized automatically to elliptic-curve/Hill-type schemes that employ randomized encryption, nonce-dependent or secret mappings, changing keys, nonlinear layers, or other mechanisms that remove the exposed public linear relation. Future work on

secure elliptic-curve/matrix constructions should evaluate such mechanisms under an explicit modern security model rather than infer security from elliptic-curve arithmetic or histogram uniformity alone.

CRediT Authorship Contribution Statement

Yanuar Bhakti Wira Tama: Conceptualization, Methodology, Formal Analysis, Software, Investigation, Writing – Original Draft, Supervision.
Sindi Lukmaini: Investigation, Software, Validation, Visualization, Writing – Review & Editing.
Muhammad Firdhausi Fahmi: Conceptualization, Validation, Formal Analysis, Writing – Review & Editing.

Declaration of Generative AI and AI-assisted Technologies

During the preparation of this manuscript, the authors used OpenAI ChatGPT to assist with language refinement, manuscript restructuring, algebraic consistency checking, and the drafting and debugging of code used for reproducible computational verification. All mathematical derivations, proofs, computational procedures, numerical results, and scientific interpretations were independently reviewed and verified by the authors. The authors take full responsibility for the accuracy, integrity, and final content of this manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Funding and Acknowledgments

This research was supported by the Institute for Research and Community Service (LPPM), Institut Teknologi Kalimantan (ITK). The authors gratefully acknowledge LPPM ITK for the financial support provided for this research.

Ethical Considerations

Ethical approval was not required because the study uses mathematical analysis, computational experiments, and publicly available legal and historical texts and does not involve human participants, animals, personal data, or sensitive records.

Data and Code Availability

The complete reproducibility package is supplied with the OJS revision as Supplementary Files S1–S5: S1 (Python implementation), S2 (1200 row-level encryption–decryption results), S3 (CPA key-recovery verification for 400 sampled keys), S4 (summary statistics), and S5 (execution log). A permanent public copy of the same reproducibility package is archived on Zenodo, version 1.0.0, at doi:10.5281/zenodo.23232833. The public plaintext sources used in the experiment are directly accessible at the following official links: the Preamble of the 1945 Constitution of Indonesia, `https://berkas.dpr.go.id/jdih/UUD45-Naskah-ASLI.pdf`; the U.S. Declaration of Independence transcription, `https://www.archives.gov/founding-docs/declaration-transcript`; and the U.S. Bill of Rights transcription, `https://www.archives.gov/founding-docs/bill-of-rights-transcript`. References [22–24] provide the corresponding bibliographic records.

References

[1] Lester S. Hill. “Cryptography in an Algebraic Alphabet”. In: *The American Mathematical Monthly* 36.6 (1929), pp. 306–312. DOI: 10.1080/00029890.1929.11986963.

[2] Kalika Prasad and Hrishikesh Mahato. "Cryptography Using Generalized Fibonacci Matrices with Affine-Hill Cipher". In: *Journal of Discrete Mathematical Sciences and Cryptography* 25.8 (2022), pp. 2341–2352. DOI: 10.1080/09720529.2020.1838744.

[3] Vaishali Billore and Naresh Patel. "Cryptography Utilizing the Affine-Hill Cipher and Extended Generalized Fibonacci Matrices". In: *Electronic Journal of Mathematical Analysis and Applications* 11.2 (2023), pp. 1–11. DOI: 10.21608/ejmaa.2023.295792.

[4] Munesh Kumari and Jagmohan Tanti. "Cryptography Using Multinacci Block Matrices". In: *International Journal of Nonlinear Analysis and Applications* 14.10 (2023), pp. 57–65. DOI: 10.22075/ijnaa.2023.29918.4295.

[5] Yuehong Chen, Rong Xie, Haotong Zhang, Dongdong Li, and Weiwei Lin. "Generation of High-Order Random Key Matrix for Hill Cipher Encryption Using the Modular Multiplicative Inverse of Triangular Matrices". In: *Wireless Networks* 30.6 (2024), pp. 5697–5707. DOI: 10.1007/s11276-023-03330-8.

[6] Şeyda Dalkılıç, Ahmet Eren Kepekçi, and Muhammed Zekeriya Baştürk. "Hill Cipher with r-Circulant Matrices". In: *Journal of Engineering Technology and Applied Sciences* 10.1 (2025), pp. 21–27. DOI: 10.30931/jetas.1516337.

[7] Melih Göcen and Sümeyye Hacer Büyükakyol. "On Matrix-Based Cryptography Using Matrix Norm and Special Integer Sequences". In: *Mathematische Nachrichten* 299.8 (2026), e70180. DOI: 10.1002/mana.70180.

[8] Porter E. III Coggins. "Multidimensional Hill Cipher Substitution–Permutation Network". In: *Journal of Cybersecurity and Privacy* 6.3 (2026), p. 104. DOI: 10.3390/jcp6030104.

[9] Neal Koblitz. "Elliptic Curve Cryptosystems". In: *Mathematics of Computation* 48.177 (1987), pp. 203–209. DOI: 10.1090/S0025-5718-1987-0866109-5.

[10] Parveiz Nazir Lone, Deep Singh, Veronika Stoffová, Deep Chandra Mishra, and Umar Hussain Mir. "Cryptanalysis and Improved Image Encryption Scheme Using Elliptic Curve and Affine Hill Cipher". In: *Mathematics* 10.20 (2022), p. 3878. DOI: 10.3390/math10203878.

[11] Nur Hidayah Mohammad Ismail and Md Yushalify Misro. "An Improved Image Encryption Algorithm Based on Bézier Coefficients Matrix". In: *Journal of King Saud University - Computer and Information Sciences* 34.10 (2022), pp. 10056–10067. DOI: 10.1016/j.jksuci.2022.10.005.

[12] Mohammad Mazyad Hazzazi, Hafeez Ur Rehman, Tariq Shah, and Hajra Younas. "Asymmetric Key Cryptosystem for Image Encryption by Elliptic Curve over Galois Field $GF(2^n)$". In: *Computers, Materials & Continua* 76.2 (2023), pp. 2033–2060. DOI: 10.32604/cmc.2023.040629.

[13] Hiba Hilal Hadi and Ammar Ali Neamah. "An Image Encryption Method Based on Modified Elliptic Curve Diffie-Hellman Key Exchange Protocol and Hill Cipher". In: *Open Engineering* 14.1 (2024). DOI: 10.1515/eng-2022-0552.

[14] Heping Wen, Yiting Lin, Lincheng Yang, and Ruiting Chen. "Cryptanalysis of an Image Encryption Scheme Using Variant Hill Cipher and Chaos". In: *Expert Systems with Applications* 250 (2024), p. 123748. DOI: 10.1016/j.eswa.2024.123748.

[15] Rania Saeed, Hatem M. Bahig, and Dieaa I. Nassr. "Cryptanalysis of Augmented Hill Cipher". In: *Contemporary Mathematics* 6.6 (2025), pp. 8822–8841. DOI: 10.37256/cm.6620258475.

[16] Bety Hayat Susanti, Aditya Sukhoi Lean Sumule, and Mareta Wahyu Ardyani. "Security Analysis of Modified ESRKGS-RSA Using Lenstra's Elliptic Curve Method". In: *CAUCHY: Jurnal Matematika Murni dan Aplikasi* 10.2 (2025), pp. 805–820. DOI: 10.18860/cauchy.v10i2.32189.

[17] Ersen Akinci and Sukran Uygun. "A Block Encryption Construction Based on Bi-Periodic Pell Matrices and Elliptic Curve Group Operations". In: *Axioms* 15.7 (2026), p. 527. DOI: 10.3390/axioms15070527.

[18] Yanuar Bhakti Wira Tama and Muhammad Firdhausi Fahmi. "Sistem Kriptografi Klasik Dengan Memanfaatkan Orde Dari Grup Titik Pada Kurva Eliptik Bentuk Montgomery". In: *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi* 11.2 (2023), pp. 361–371. DOI: 10.37905/euler.v11i2.23009.

[19] Sindi Lukmaini, Yanuar Bhakti Wira Tama, and Kartika Nugraheni. "Penerapan Algoritma Hill Cipher dan Permutasi pada Sistem Kriptografi Polyalphabetic". In: *Epsilon: Jurnal Matematika Murni dan Terapan* 18.2 (2024), pp. 149–164. DOI: 10.20527/epsilon.v18i2.12845.

[20] Yanuar Bhakti Wira Tama and Syamsul Mujahidin. "Implementation of the Elliptic Curve Cryptography Method in Digital Image Security in Medical Images". In: *SPECTA Journal of Technology* 8.3 (2024), pp. 233–241. DOI: 10.35718/specta.v8i3.1253.

[21] Yanuar Bhakti Wira Tama and Syamsul Mujahidin. "Grayscale Image Cryptosystem Using Sub-Matrix Multiplications and Gauss Jordan Elimination in $\mathbb{Z}/256\mathbb{Z}$". In: *AIP Conference Proceedings*. Vol. 3326. 1. AIP Publishing, 2026, p. 020001. DOI: 10.1063/5.0308775.

[22] Sekretariat Jenderal DPR RI. *Undang-Undang Dasar Negara Republik Indonesia 1945: Naskah Asli*. 2025. https://berkas.dpr.go.id/jdih/UUD45-Naskah-ASLI.pdf (visited on 09/02/2026).

[23] U.S. National Archives and Records Administration. *Declaration of Independence: A Transcription*. 2026. https://www.archives.gov/founding-docs/declaration-transcript (visited on 09/13/2026).

[24] U.S. National Archives and Records Administration. *The Bill of Rights: A Transcription*. 2026. https://www.archives.gov/founding-docs/bill-of-rights-transcript (visited on 09/13/2026).

[25] Mahdianta Pandia, Poltak Sihombing, Mohammad Andri Budiman, and Erna Budhiarti Nababan. "Development of a Hybrid Probabilistic Key Generation Method Using Multi-Prime Rivest-Shamir-Adleman and Learning with Errors to Enhance Hill Cipher Security". In: *Eastern-European Journal of Enterprise Technologies* 1.9(139) (2026), pp. 31–39. DOI: 10.15587/1729-4061.2026.348890.