On the Keyspace of the Hill Cipher

Jeffrey Overbey[1], William Traves[2], and Jerzy Wojdylo[3]

ADDRESSES: (1) Department of Mathematics, Southeast Missouri State University, One University Plaza, Cape Girardeau, MO 63701, joverbey@actilon.com, (2) Department of Mathematics, U.S. Naval Academy, 572C Holloway Road, Annapolis, MD 21402, traves@usna.edu, and (3) Department of Mathematics, Southeast Missouri State University. jwojdylo@semo.edu.

ABSTRACT: In its most general form, the Hill cipher's keyspace consists of all matrices of a given dimension that are invertible over $\mathbb{Z}_m$. Working from known results over finite fields, we assemble and prove a formula for the number of such matrices. We also compare this result with the total number of matrices and the number of involutory matrices for a given dimension and modulus, identifying the effects of change in dimension and modulus on the order of the keyspace.

KEYWORDS: Hill cipher, invertible matrices, involutory matrices, general linear group, keyspace

1 Introduction

Although the Hill cipher's susceptibility to cryptanalysis has rendered it unusable in practice, it still serves an important pedagogical role in both cryptology and linear algebra. It is this role in linear algebra that raises several interesting questions.

In the general case, the keyspace of the Hill cipher is precisely $\mathrm{GL}(d, \mathbb{Z}_m)$—the group of $d \times d$ matrices that are invertible over $\mathbb{Z}_m$ for a predetermined modulus m. We first present a formula for the order of this group (with proof). We then consider involutory matrices, which eliminate the necessity of computing matrix inverses for Hill decryptions. Finally, we compare the total number of matrices with the number of invertible and involutory matrices, identifying the effects of change in dimension and modulus on the order of the keyspace.

1.1 The Hill Cipher

For the sake of completeness, we provide the following description of the Hill cipher. Though it varies slightly from Hill's original description given in [2] and [3], it is fundamentally equivalent and is consistent with modern texts in cryptography (such as [10]). A plaintext string over an alphabet of order m is rewritten as a vector over $\mathbb{Z}_m$ using a natural correspondence. In either column-major or row-major order, the vector is rewritten as a matrix P with d rows, where d is an arbitrarily chosen positive integer. (To fill the matrix, it may be necessary to pad the string with extra characters to make its length a multiple of d.) A matrix $K \in \mathrm{GL}(d, \mathbb{Z}_m)$ is chosen to be the key matrix. The encryption is performed by computing

$$C = e_K(P) = KP$$

and rewriting the resulting matrix as a string over the same alphabet. Decryption is performed similarly by computing

$$P = d_K(C) = K^{-1}C.$$

1.2 Prerequisites

Before beginning our look into the order of $\mathrm{GL}(d, \mathbb{Z}_m)$, we establish some prerequisite properties of matrices that are componentwise congruent.

Definition Let $A, B \in M_{r\times s}(\mathbb{Z}), d \in \mathbb{Z}^+$. We use the notation $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$ to denote that, for all i and j, $a_{ij} \equiv b_{ij} \pmod{m}$.

Lemma 1.2.1 $\overset{\mathrm{M}}{\equiv}$ *is an equivalence relation.* $\blacksquare$

Lemma 1.2.2 *Let A, B, C, and D be matrices in $M_{r\times s}(\mathbb{Z})$, X and Y be matrices in $M_{s\times t}(\mathbb{Z})$, α and β be integers, and m be a positive integer.*

a) If $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$ and $C \overset{\mathrm{M}}{\equiv} D \pmod{m}$, then $A \pm C \overset{\mathrm{M}}{\equiv} B \pm D \pmod{m}$.

b) If $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$ and $\alpha \equiv \beta \pmod{m}$, then $\alpha A \overset{\mathrm{M}}{\equiv} \beta B \pmod{m}$.

c) If $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$ and $X \overset{\mathrm{M}}{\equiv} Y \pmod{m}$, then $AX \overset{\mathrm{M}}{\equiv} BY \pmod{m}$.

d) If A and B are square matrices such that $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$, then $\det A \equiv \det B \pmod{m}$.

Proof Parts a and b follow directly from properties of congruence. To prove part c, assume that $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$ and $X \overset{\mathrm{M}}{\equiv} Y \pmod{m}$. Then, for all i, j, and k, $a_{ik} \equiv b_{ik} \pmod{m}$ and $x_{kj} \equiv y_{kj} \pmod{m}$. Let $Z = AX$ and $\hat{Z} = BY$. Now $z_{ij} = \sum_{k=1}^{s} a_{ik}x_{kj} \bmod m$, and $\hat{z}_{ij} = \sum_{k=1}^{s} b_{ik}y_{kj} \bmod m$. But since $a_{ik} \equiv b_{ik} \pmod{m}$ and $x_{kj} \equiv y_{kj} \pmod{m}$ for all i, j, and k, it follows that $z_{ij} \bmod m = \hat{z}_{ij} \bmod m$ and, thus, $z_{ij} \equiv \hat{z}_{ij} \pmod{m}$ for all i and j. Thus, $Z \overset{\mathrm{M}}{\equiv} \hat{Z} \pmod{m}$, and thus $\overset{\mathrm{M}}{\equiv}$ is preserved under matrix multiplication.

To prove part d, we assume that $A \overset{\mathrm{M}}{\equiv} B \pmod{m}$, noting then that $a_{ij} \bmod m = b_{ij} \bmod m$ for all i and j. We will use the standard definition

$$\det A = \sum_{\sigma \in S_d} \operatorname{sign}(\sigma) a_{1\sigma_1} a_{2\sigma_2} \ldots a_{s\sigma_s}.$$

Now $\det A \bmod m = \left(\sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \cdot a_{1\sigma_1} \cdot a_{2\sigma_2} \cdot \ldots \cdot a_{s\sigma_s}\right) \bmod m$
$= \left(\sum_{\sigma \in S_d} (\operatorname{sign}(\sigma) \cdot a_{1\sigma_1} \cdot a_{2\sigma_2} \cdot \ldots \cdot a_{s\sigma_s}) \bmod m\right) \bmod m$
$= \left(\sum_{\sigma \in S_d} (\operatorname{sign}(\sigma) \cdot (a_{1\sigma_1} \bmod m) \cdot \ldots \cdot (a_{s\sigma_s} \bmod m)) \bmod m\right) \bmod m$
$= \left(\sum_{\sigma \in S_d} (\operatorname{sign}(\sigma) \cdot (b_{1\sigma_1} \bmod m) \cdot \ldots \cdot (b_{s\sigma_s} \bmod m)) \bmod m\right) \bmod m$
$= \left(\sum_{\sigma \in S_d} (\operatorname{sign}(\sigma) \cdot b_{1\sigma_1} \cdot b_{2\sigma_2} \cdot \ldots \cdot b_{s\sigma_s}) \bmod m\right) \bmod m$
$= \left(\sum_{\sigma \in S_d} \operatorname{sign}(\sigma) \cdot b_{1\sigma_1} \cdot b_{2\sigma_2} \cdot \ldots \cdot b_{s\sigma_s}\right) \bmod m$
$= \det B \bmod m$. Therefore, $\det A \equiv \det B \bmod m$. ∎

2 The Order of GL$(d, \mathbb{Z}_m)$

We now present the formulae necessary for computing the order of the Hill cipher's keyspace. We begin with the case of a prime modulus, then extend that formula to consider prime-power and general moduli.

2.1 $|\mathrm{GL}(d, \mathbb{Z}_p)|$, where p is prime

When p is prime, calculating the order of $\mathrm{GL}(d, \mathbb{Z}_p)$ is fairly straightforward since $\mathbb{Z}_p$ is a field. The following result is fairly well-known.

Theorem 2.1.1 *The number of $d \times d$ invertible matrices over $\mathbb{Z}_p$ for a prime p is*

$$|GL(d, \mathbb{Z}_p)| = \prod_{i=0}^{d-1} (p^d - p^i).$$

Proof The standard proof of this formula [9, Thm 8.13] describes how d column vectors over $\mathbb{Z}_p$ can be chosen such that they will form an invertible matrix. The only restriction on the first vector is that it be nonzero, as this would destroy the linear independence of the columns. Thus, there are $p^d - 1$ possible "first columns." Now suppose that we have i linearly independent columns. The $(i+1)$-st column can be chosen in $p^d - p^i$ ways, avoiding linear combinations of the previous i columns. By induction, the number of invertible matrices is $\prod_{i=0}^{d-1} (p^d - p^i)$. ∎

2.2 $|\mathrm{GL}(d, \mathbb{Z}_{p^n})|$, for a prime p and a natural number n

We now consider the case where the modulus is a power n of a prime p. We will require the following lemmas.

Lemma 2.2.1 *Suppose $m = p^n$, where $n \in \mathbb{Z}^+$ and p is prime. Let $A \in M_{d\times d}(\mathbb{Z}_m)$, where $d \in \mathbb{Z}^+$. Dividing the entries of A by p, we may form a $d \times d$ matrix C of quotients and a $d\times d$ matrix B of remainders such that $A = B+pC$. Then A is invertible mod p^n if and only if B is invertible mod p.*

Proof Let A, B, and C be as indicated. Note that B has entries in $\{0, \ldots, p-1\}$ and C has entries in $\{0, \ldots, p^{n-1} - 1\}$.

Because B is the matrix of quotients upon dividing the entries of A by p, it follows that $A \overset{\text{M}}{\equiv} B \pmod p$ (Definition 1.2). By Lemma 1.2.2d, we know that $\det A \equiv \det B \pmod p$. Thus, $\gcd(\det A, p) = \gcd(\det B, p)$. Now A is invertible mod p^n iff $\gcd(\det A, p^n) = 1$ iff $\gcd(\det A, p) = 1$ iff $\gcd(\det B, p) = 1$ iff B is invertible mod p. ▮

Theorem 2.2.2 *The number of $d \times d$ invertible matrices over $\mathbb{Z}_{p^n}$ for a prime p and natural number n is*

$$|GL(d, \mathbb{Z}_{p^n})| = p^{(n-1)d^2} \prod_{i=0}^{d-1} (p^d - p^i).$$

Proof We can use Lemma 2.2.1 to rewrite any matrix $A \in \mathbb{Z}_{p^n}$ as $B + pC$, where B has entries in $\mathbb{Z}_p$ and C has entries in $\mathbb{Z}_{p^n-1}$, noting that A is invertible mod p^n iff B is invertible mod p. By Theorem 2.1.1, we know that there are $\prod_{i=0}^{d-1}(p^d - p^i)$ possible matrices B. There are $p^{(n-1)d^2}$ possibilities for C since it need not be invertible. Then the number of possibilities for A and, therefore, the total number of invertible matrices mod p^n is $p^{(n-1)d^2} \prod_{i=0}^{d-1}(p^d - p^i)$. ▮

2.3 $|\mathbf{GL}(d, \mathbb{Z}_m)|$, for any integer $m \geq 2$

We now consider the case of a composite modulus m. In Lemma 2.3.1 and Theorem 2.3.3, we will rewrite m as a product of powers of distinct prime numbers, i.e. $m = p_1^{n_1} p_2^{n_2} \ldots p_z^{n_z}$. In Lemma 2.3.1, a matrix A over $\mathbb{Z}_m$ is mapped to a z-tuple of matrices where each component is A modulo a single factor in this decomposition of m. We prove that such a mapping is an isomorphism.

Lemma 2.3.1 *Define $\phi : M_{d\times d}(\mathbb{Z}_m) \longrightarrow \bigoplus_{i=1}^{z} M_{d\times d}(\mathbb{Z}_{p_i^{n_i}})$ by $\phi(A) = \bigoplus \phi_i(A)$, where $\phi_i(A) = A$ mod $p_i^{n_i}$, with ϕ_i being the i-th component of the image tuple. Then ϕ is a ring isomorphism.*

Proof Let A and B be matrices in $M_{d\times d}(\mathbb{Z}_m)$.

Clearly ϕ is well-defined. Bijectivity of ϕ follows directly from the Chinese Remainder Theorem. We prove that ϕ is operation preserving.

Let $C = AB$. Let $\phi(A) = (A^{(1)}, A^{(2)}, \ldots, A^{(z)})$ and $\phi(B) = (B^{(1)}, \ldots, B^{(z)})$. Note that $A^{(i)}B^{(i)} = (AB)^{(i)}$ for all i since matrix multiplication preserves modular equivalence. Now

$$\begin{aligned}\phi(A)\phi(B) &= (A^{(1)}, A^{(2)}, \ldots, A^{(z)})(B^{(1)}, B^{(2)}, \ldots, B^{(z)}) \\ &= (A^{(1)}B^{(1)}, A^{(2)}B^{(2)}, \ldots, A^{(z)}B^{(z)}) \\ &= ((AB)^{(1)}, (AB)^{(2)}, \ldots, (AB)^{(z)}) \\ &= \phi(AB).\end{aligned}$$

Therefore, ϕ is an isomorphism. ▮

Lemma 2.3.2 *Let ϕ be as defined in Lemma 2.3.1. Then a matrix A over $\mathbb{Z}_m$ is invertible iff $\phi(A)$ is invertible.*

Proof If A is invertible mod m, then $\phi(A)\phi(A^{-1}) = \phi(AA^{-1}) = \phi(I)$, which is the identity in the codomain since ϕ is a ring isomorphism. Thus $\phi(A)$ is invertible. The converse follows because ϕ is an isomorphism.

Theorem 2.3.3 *The number of $d \times d$ matrices invertible mod $m = \prod_i p_i^{n_i}$ is*

$$|GL(d, \mathbb{Z}_m)| = \prod_i \left(p_i^{(n_i-1)d^2} \prod_{k=0}^{d-1} (p_i^d - p_i^k) \right).$$

Proof Let ϕ be as in Lemma 2.3.1. For a tuple in the codomain of ϕ to be invertible, each of its matrix components must be invertible. The number of invertible matrices mod $p_i^{n_i}$, for each i, is given by Theorem 2.2.2. The number of invertible tuples, then, is the product of the number of possibilities for each component. From Lemma 2.3.2, we know that a matrix A over $\mathbb{Z}_m$ is invertible iff its tuple representation $\phi(A)$ is invertible. Since ϕ is bijective, the number of matrices invertible mod m is as above. ▮

3 Involutory Matrices

Using Theorem 2.3.3, one can determine that, of the 3×3 matrices over a 26-letter alphabet, 1,634,038,189,056 can serve as key matrices. For most of these matrices, however, the decryptor is left with the nuisance of computing a matrix inverse. Hill [3] recognized this problem and proposed the use of involutory matrices in order to make the cipher symmetric. Since a matrix A is involutory if and only if $A^2 = I$, or, equivalently, $A = A^{-1}$, the Hill decryption function

$$d_K(C) = K^{-1}C$$

can be rewritten as

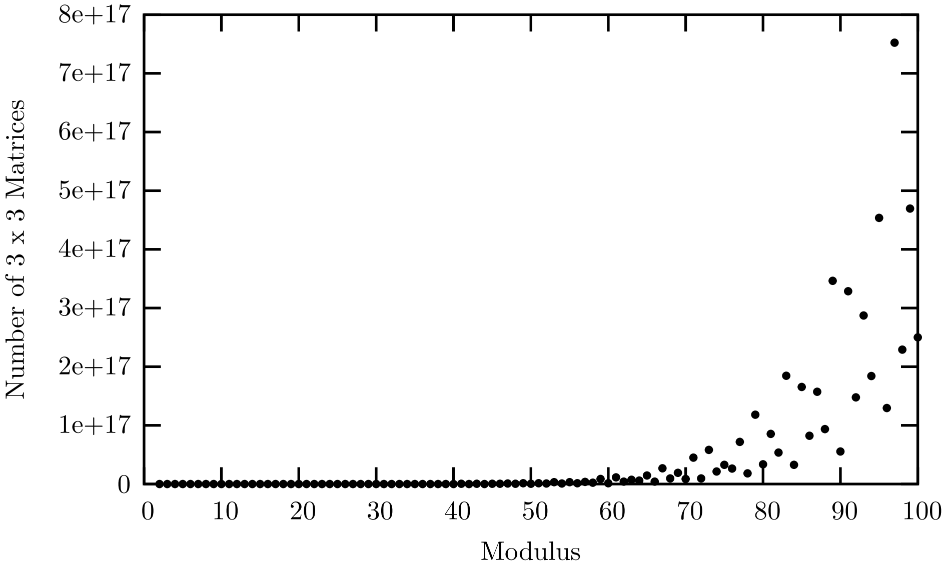
$$d_K(C) = KC,$$

which coincides with the encryption function

$$e_K(P) = KP.$$

In Hill's time, this meant that the same machinery could be used both for encryption and decryption of messages; no additional hardware would be needed to compute inverses before decrypting.

Figure 1: Numbers of 3 × 3 Invertible Matrices



The question arises: if the keyspace is restricted to matrices that are involutory, how does this affect its order? The answer can be determined using results of Reiner, Hodges, Levine, and Korfhage, which we will soon state. In the next section, we will observe exactly how restrictive the involutory keyspace is. (Since our primary focus is on the order of the keyspace, we will not present methods of construction of involutory matrices. The interested reader is referred to [7] or [6].) As in the case of invertible matrices, we begin with the cases of prime and prime-power moduli and then consider a general modulus. The following theorems are restated nearly verbatim, with changes made only for the purpose of clarification. Proofs have been omitted; the interested reader may refer to the references.

3.1 Involutory matrices over $\mathbb{Z}_{p^n}$, for a prime p and a natural number n

When the modulus is an odd prime p or a power of that prime, we may use the following result of Reiner. In order to conform to the literature, we adopt the convention of using g_t to stand for the formula for $\mathrm{GL}(t, \mathbb{Z}_p)$ given in Theorem 2.1.1.

Theorem 3.1.1 *[8, p. 774] For a $d \times d$ integer matrix X, the number of solutions of $X^2 = I$ (mod p^{a+1}) for an odd prime p is given by*

$$\sum_{t=0}^{d} \left(\frac{g_d}{g_t g_{d-t}} \cdot p^{2t(d-t)a} \right),$$

where g_t is given by

$$g_t = p^{t^2} \prod_{i=1}^{t} (1 - p^{-i}) = \prod_{i=0}^{t-1} (p^t - p^i). \tag{3.1.2}$$

for $0 < t \leq d$ and $g_0 = 1$.

When the modulus is, instead, a power of 2, we must use the following result. The case $n = 1$ was derived by Hodges [4, p. 520]; the remainder is the work of Levine and Korfhage [6, p. 644]. The two results have been combined for simplification. The formula for $T(d, 2^2)$ below corrects a minor typesetting error in [6, p. 644].

Theorem 3.1.3 *The number $T(d, 2^n)$ of $d \times d$ involutory matrices mod 2^n is given by the following, where g_t is given by (3.1.2). When $n = 1$, we have*

$$T(d, 2^1) = g_d \sum_{t=0}^{\lfloor d/2 \rfloor} \frac{2^{-t(2d-3t)}}{g_t g_{d-2t}}.$$

If $n = 2$, then

$$T(d, 2^2) = g_d \sum_{h=0}^{\lfloor d/2 \rfloor} \frac{2^{d^2 - 4hd + 5h^2}}{g_h g_{d-2h}}.$$

If $n \geq 3$, then

$$T(d, 2^n) = 2^{d^2} g_d \sum_{h=0}^{d} 2^{2h(d-h)(n-3)} \sum_{r=0}^{\min(h,d-h)} \frac{2^{r(3r-2d)}}{g_r g_{h-r} g_{d-h-r}}.$$

3.2 Involutory matrices over $\mathbb{Z}_m$, for composite m

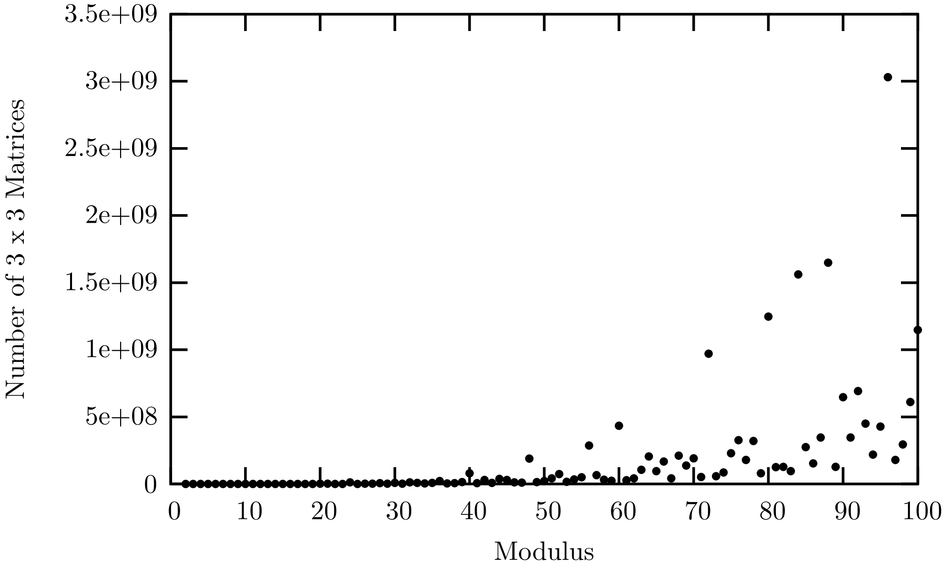
As with invertible matrices, we are able to determine the number of involutory matrices modulo a composite number m by decomposing m into its prime power factorization and computing the number of involutory matrices modulo each prime power.

Theorem 3.2.1 *[6, p. 652] Let $m = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$ be the prime power factorization of m. Then*

$$T(d, m) = \prod_{i=0}^{k} T(d, p_j^{n_j}),$$

where $T(d, m)$ denotes the number of $d \times d$ involutory matrices over $\mathbb{Z}_m$.

Figure 2: Numbers of 3 × 3 Involutory Matrices



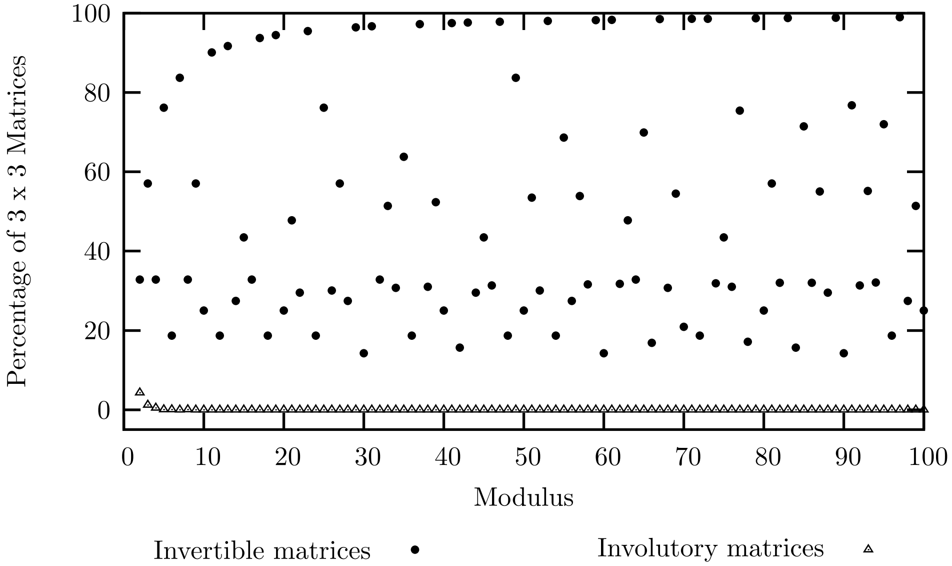
4 A Comparison

We now have the necessary formulae to compute the total number of matrices, the number of invertible matrices, and the number of involutory matrices over $\mathbb{Z}_m$ for any modulus m. This allows us to answer several fundamental questions. For example, of the 3×3 matrices over $\mathbb{Z}_{26}$, what percentage are invertible? Involutory? Does increasing the dimension or modulus necessarily guarantee that the keyspace is increasing?

Figures 1 and 2 show the number of 3×3 invertible and involutory matrices, respectively, for moduli up to 100. (Graphs of these functions for matrices of higher dimension are similar in appearance.) A quick comparison of these figures provides two insights. First, the number of invertible matrices is significantly larger than the number of involutory matrices, no matter what the modulus. Second, an increase in the modulus does not always correspond with an increase in the number of matrices. In terms of the Hill cipher, this indicates that increasing the size of the alphabet will not necessarily increase the size of the keyspace; it may even shrink it. (On the other hand, increasing the dimension of the matrix does increase the number of invertible matrices.)

Given these graphs, one may hypothesize that if a random matrix of given dimension and modulus is chosen, then there is a fair probability that it is invertible but a much lower probability that it is involutory. This is generally true. Figure 3 shows the percentage of all 3×3 matrices that are invertible and involutory for moduli up to 100. Clearly, the values for involutory matrices

Figure 3: Percentage of 3 × 3 Matrices That Are Invertible, Involutory



approach zero quite rapidly. The values for invertible matrices, however, seem rather scattered. Despite this, there is one particularly noticeable trend: The values for prime moduli are much higher than those for nearby composite moduli. In fact, they appear to be approaching 100% as the modulus increases. There are several other patterns, however, that are not immediately obvious. For example, the values for moduli m = 6, 12, 18, 24, 36, 48, 54, 72, and 96 are the same. It can be shown that this is because they all have the same set of prime factors, $\{2, 3\}$. The values at $42 = 2 \cdot 3 \cdot 7$ and $84 = 2^2 \cdot 3 \cdot 7$ are also the same. We will now pursue these notions more rigorously.

Definition Let $m \geq 2$ and d be positive integers. Define

$$f(d, m) = \frac{|\mathrm{GL}(d, \mathbb{Z}_m)|}{|M_{d \times d}(\mathbb{Z}_m)|}. \tag{4.2}$$

In other words, $f(d, m)$ denotes the proportion of $d \times d$ matrices over $\mathbb{Z}_m$ that are invertible.

Above, we observed that the percentage of matrices that were invertible was the same for moduli with the same primes in their factorization, regardless of how many times each prime occurred. We will now prove this result. Before doing so, we establish a simple result that allows us to more easily evaluate $f(d, m)$ when m is prime.

Lemma 4.3 *Let d denote a positive integer and p a prime. Then*

$$f(d,p) = \prod_{j=1}^{d} \left(1 - \frac{1}{p^j}\right).$$

Proof Since p is prime, we can use Theorem 2.1.1 to rewrite

$$f(d,p) = \frac{\prod_{i=0}^{d-1}(p^d - p^i)}{p^{d^2}}. = \prod_{i=0}^{d-1}\left(\frac{p^d - p^i}{p^d}\right) = \prod_{j=1}^{d}\left(1 - \frac{1}{p^j}\right). \quad \blacksquare$$

Theorem 4.4 *Let d denote a positive integer and m a positive integer ≥ 2 with prime factorization $m = \prod_i p_i^{n_i}$. Then*

$$f(d,m) = \prod_i \prod_{j=1}^{d}\left(1 - \frac{1}{p_i^j}\right).$$

Proof We begin by substituting the result of Theorem 2.3.3 into (4.2):

$$f(d,m) = \frac{\prod_i \left(p_i^{(n_i-1)d^2} \prod_{k=0}^{d-1}(p_i^d - p_i^k)\right)}{m^{d^2}}.$$

Rewriting m in its prime factorization and commuting the factors in the denominator, we have

$$\begin{aligned}
f(d,m) &= \frac{\prod_i \left(p_i^{(n_i-1)d^2} \prod_{k=0}^{d-1}(p_i^d - p_i^k)\right)}{\prod_i p_i^{n_i d^2}} \\
&= \frac{\prod_i \left(p_i^{n_i d^2} p_i^{-d^2} \prod_{k=0}^{d-1}(p_i^d - p_i^k)\right)}{\prod_i p_i^{n_i d^2}} \\
&= \frac{\prod_i p_i^{n_i d^2} p_i^{-d^2} \prod_{k=0}^{d-1}(p_i^d - p_i^k)}{\prod_i p_i^{n_i d^2}} \\
&= \prod_i \frac{p_i^{n_i d^2} p_i^{-d^2} \prod_{k=0}^{d-1}(p_i^d - p_i^k)}{p_i^{n_i d^2}} \\
&= \prod_i \left(\frac{\prod_{k=0}^{d-1}(p_i^d - p_i^k)}{p_i^{d^2}}\right).
\end{aligned}$$

Finally, by Lemma 4.3, we have

$$f(d,m) = \prod_i \prod_{j=1}^{d}\left(1 - \frac{1}{p_i^j}\right). \quad \blacksquare$$

In addition to providing a simpler evaluation of f, this proves the following corollary.

Corollary 4.5 *$f(d,m)$ does not depend on the exponents in the prime factorization of m.* $\blacksquare$

Example To illustrate, we consider various moduli that factor into powers of 2 and 3.

$$\begin{array}{lclcl}
f(3,6) & = & f(3,2\cdot 3) & \approx & 0.1872 \\
f(3,72) & = & f(3,2^3\cdot 3^2) & \approx & 0.1872 \\
f(3,96) & = & f(3,2^5\cdot 3) & \approx & 0.1872 \\
f(3,62208) & = & f(3,2^8\cdot 3^5) & \approx & 0.1872
\end{array}$$

We now prove the "particularly noticeable trend" in Figure 3, i.e., that the percentage of matrices that are invertible mod p approaches 100% as p approaches infinity. In terms of f, this is stated as follows.

Theorem 4.6 *Let d denote a positive integer and p a prime. Then*

$$\lim_{p\to\infty} f(d,p) = 1.$$

Proof By Lemma 4.3, we know that we are considering

$$\lim_{p\to\infty} f(d,p) = \lim_{p\to\infty} \prod_{j=1}^{d} \left(1 - \frac{1}{p^j}\right).$$

Since $\lim_{p\to\infty}(1 - 1/p^j) = 1$ for all j, and there are finitely many j's, it follows that

$$\lim_{p\to\infty} f(d,p) = \prod_{j=1}^{d} 1 = 1. \quad \blacksquare$$

Another interpretation of this theorem is as follows: the probability that a randomly-chosen $d \times d$ matrix is invertible mod p is about 1 for any "large" prime p.

While the above theorem showed that we can make f arbitrarily close to 1, it is also possible to make f arbitrarily close to zero: simply increase the number of distinct primes in its factorization. The following theorem presents this more formally.

Theorem 4.7 *Let d denote a positive integer and m a positive integer with k distinct prime factors $m = p_1 p_2 \ldots p_k$. Then*

$$\lim_{k\to\infty} f(d,m) = 0.$$

Before presenting the proof of this theorem, we review two results from number theory.

Definition [5, p. 1694] The Riemann zeta function, $\zeta(s)$, is defined as

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}. \tag{4.8}$$

Theorem 4.9 *(Euler) [5, p. 1694] An alternative representation of* $\zeta(s)$ *is given by*

$$\zeta(s) = \prod_{p-prime} \frac{1}{1-\frac{1}{p^s}}.$$

We are now ready to prove Theorem 4.7.

Proof By Theorem 4.4, we have

$$\begin{aligned} \lim_{k\to\infty} f(d,m) &= \lim_{k\to\infty} \prod_{i=1}^{k}\prod_{j=1}^{d}\left(1-\frac{1}{p_i^j}\right) \\ &= \lim_{k\to\infty} \prod_{j=1}^{d}\prod_{i=1}^{k}\left(1-\frac{1}{p_i^j}\right) \end{aligned}$$

Now $\lim_{k\to\infty}\prod_{i=1}^{k}\left(1-1/p_i^j\right) = 1/\zeta(j)$ by Euler's product (Theorem 4.9). This allows us to simplify

$$\begin{aligned} \lim_{k\to\infty} f(d,m) &= \lim_{k\to\infty} \prod_{j=1}^{d}\prod_{i=1}^{k}\left(1-\frac{1}{p_i^j}\right) \\ &= \left(\lim_{k\to\infty}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)\right)\prod_{j=2}^{d}\frac{1}{\zeta(j)}. \end{aligned}$$

From (4.8), we can quickly determine that $\zeta(1) = \infty$ and so

$$\lim_{k\to\infty}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right) = 0.$$

All other $\zeta(s) > 0$ for $s \geq 2$. Therefore, the product is zero, and the theorem is proved. ▌

As above, this theorem can be restated in terms of probability: the probability that a randomly-chosen $d \times d$ matrix is invertible mod m is almost zero provided m has sufficiently many different prime divisors.

5 Conclusion

We have given formulae for the numbers of $d \times d$ invertible and involutory matrices mod m. We note that, while involutory matrices may save decryption time, requiring that key matrices be involutory significantly reduces the size of the keyspace. We have also observed that, while increasing the dimension of key matrices leads to a larger keyspace, increasing the modulus (i.e., the size of the alphabet) may not. When the keyspace is $\mathrm{GL}(d, \mathbb{Z}_m)$, prime moduli generally produce larger keyspaces than composite moduli. Thus, the largest keyspaces result from a large matrix dimension and an alphabet of prime order.

Acknowledgements

The authors would like to thank Professor Brian Winkel for suggesting the topic of involutory matrices and Project NExT for facilitating discussions that fostered these results. They would also like to acknowledge the invaluable suggestions provided by the referee.

References

[1] Acosto-de-Orozco, M.T., and J. Gomez-Calderon. 1993. On the Matrix Equation $X^n = B$ over Finite Fields. *Internat. J. Math. Math. Sci.* 16(3): 537–544.

[2] Hill, L.S. 1929. Cryptography in an Algebraic Alphabet. *Am. Math. Mon.* 36: 306–312.

[3] Hill, L.S. 1931. Concerning Certain Linear Transformation Apparatus of Cryptography. *Am. Math. Mon.* 38: 135–154.

[4] Hodges, J. 1958. The Matrix Equation $X^2 - I = 0$ over a Finite Field. *Am. Math. Mon.* 65: 518–520.

[5] Itô, K. 1987. *Iwanami Sūgaku Jiten (Encyclopedic Dictionary of Mathematics).* The MIT Press.

[6] Levine, J., and R.R. Korfhage. 1964. Automorphisms of Abelian Groups Induced by Involutory Matrices, General Modulus. *Duke Math J.* 31: 631–654.

[7] Levine, J., and H.M. Nahikian. 1962. On the Construction of Involutory Matrices. *Am. Math. Mon.* 69: 267–272.

[8] Reiner, I. 1960. The Matrix Congruence $X^2 = I(\mod p^a)$. *Am. Math. Mon.* 67: 773–775.

[9] Rotman, J. 1965. *The Theory of Groups.* Boston: Allyn and Bacon.

[10] Stinson, D.R. 2002. *Cryptography: Theory and Practice.* 2nd ed. Boca Raton: Chapman & Hall/CRC Press.

Biographical Sketches

Jeffrey Overbey is a double major in Applied Mathematics and Computer Science at Southeast Missouri State University. His interests include cryptography, programming language and compiler design, and music.

Will Traves received his Ph.D. from the University of Toronto in 1998. He is currently an Associate Professor at the U.S. Naval Academy, where he lectures about Calculus, does research on Algebra and Algebraic Geometry and helps train the Women's Rugby Team. He is co-author of *An Invitation to Algebraic Geometry* (Springer-Verlag, 2000).

Jerzy Wojdylo is an assistant professor in the Department of Mathematics at Southeast Missouri State University. He received his Ph.D. from Iowa State University in 1998 under the supervision of Jonathan D. H. Smith. In 1999 he was selected to be a national Project NExT fellow (Brown Dot). His areas of interest are combinatorics, graph theory, cryptography and card tricks.