Hill Ciphers and Modular Linear Algebra

Murray Eisenberg*

November 3, 1999

Hill ciphers are an application of linear algebra to cryptology (the science of making and breaking codes and ciphers). Below we describe what Hill ciphers are and how they are broken. And we discuss the requisite notions and facts about modular arithmetic, and about linear algebra when the scalars are no longer the real numbers but instead the integers modulo some m.

Our discussion of ciphers is based in part upon passages from David Kahn's 1967 book, *The Codebreakers: The Story of Secret Writing*. Kahn's book is a fascinating non-mathematical account of codes and ciphers in a historical context.

All arithmetic we shall do involves only **integers**, that is, the positive, negative, and zero "whole numbers" $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$.

1 Polygraphic Ciphers

Both codes and ciphers are methods for transforming a given message—the **plaintext**—into a new form that is unintelligible to anyone who does not know the rule—the **key**—for doing the transformation or who, more significantly, does not know the secret rule—the **inverse key**—for reversing the transformation in order to recover the original plaintext.[1] Both codes and ciphers may be used in conjunction with methods for concealing the very existence of the message, such as secret ink or microdots. Or, the messages transformed by codes and ciphers may be communicated in the open, for example, transmitted over shortwave radio or printed in a newspaper ad.

1.1 Codes

In the case of a **code**, the key is a table that associates to each of a specific list of words and phrases (plus syllables to "spell" words not otherwise in the list) a corresponding "codeword" (or "codenumber"). For example, the key for a code used by lobbyists might include such entries as shown in Table 1.

[1] The key itself need not be kept secret! In a *public-key* cipher, such as the RSA system devised in 1978 by Rivest, Shamir, and Adleman, the key can be published openly, because recovering the inverse key from it is extremely difficult. Ordinarily only someone who has the secret inverse key can decipher a message encrypted with such a system.

plaintext	*codeword*
senator	rabbit
soft money	applesauce
next Monday	green
vote	think

Table 1: A code table fragment

1.2 Substitution ciphers

In the case of a **cipher**, by way of contrast, the key transforms individual plaintext letters and other characters, or fixed-length groups of several characters, into new characters—the **ciphertext**. To use the key to transform plaintext into ciphertext is to **encipher** that plaintext; to use the **inverse key** to transform ciphertext back into plaintext is to **decipher** that ciphertext.

One type of cipher simply rearranges the letters of the given plaintext. The ciphers we shall study, however, are **substitutions**, where letters of the plaintext are replaced by different letters of the alphabet.

One example of a substitution cipher is a "cryptogram" such as you may have seen in a newspaper's puzzle section. Here is an example (the ciphertext is arbitrarily grouped into sets of five letters):

Y P Q M K A Y C A Z L F X Y Z

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
F	T	G	S	M	O	N	A	Y	V	P	D	C	Q	Z	R	W	L	E	X	H	B	K	I	U	J

Table 2: A substitution cipher

You can solve such a cryptogram, that is, discover the secret meaning, if you know the key. For the one above, the key is given by Table 2, where below each plaintext letter is the corresponding ciphertext letter. What is the secret plaintext in this example?

Of course, it's no fun or challenge to decipher a message if you know the key. What you really want to be able to do is figure out what the key and its inverse are—as we shall say, to **crack** the cipher (in technical terms, to "cryptanlyze" it). A simple letter-for-letter substitution, such as in the example above, may be fairly easy to crack if you have enough ciphertext. For then you can statistically analyze the ciphertext, by looking for letters (or even letter pairs or triples) appearing frequently or infrequently in it and then making use of known frequencies of letters (or pairs or triples) in typical English text. For example, one standard table of letter frequencies in English lists E T A O N I R S H D L U as the most frequent, in that order.

1.3 Polygraphic substitution ciphers

The basic shortcoming of such a simple letter-for-letter substitution is that the same plaintext letters always get replaced by the same ciphertext letters (until the key is changed, of course), and that's what makes the statistical analysis of letter frequencies applicable. A more challenging type to analyze is a **polygraphic** cipher, where the plaintext is divided into groups of adjacent letters of the same fixed length n, and then each such group is transformed into a different group of n letters. If n is not too small, then such a polygraphic substitution can render letter frequency analysis useless.

Many kinds of polygraphic ciphers have been devised. One of the more famous ones, for example, is the Playfair cipher, invented in 1854 by Charles Wheatstone, which uses digraphs (two letters per group). The first systematic yet simple polygraphic ciphers using more than two letters per group are the ones we shall study below—the **Hill ciphers**. These were first described in 1929 by their inventor, the mathematician Lester S. Hill, in the journal *The American Mathematical Monthly*. Hill ciphers constitute the first general method for successfully applying algebra—specifically, linear algebra—to polygraphic ciphers, and for applying it in a way that is, in fact, practical.

For a polygraphic substitution, changing just one or two plaintext letters can completely change the corresponding ciphertext! For one Hill cipher discussed later (where our "alphabet" includes . and ?), the two plaintexts

DESTROY PLANE and DEPLOY PLANES

correspond to ciphertexts:

PPAJZTXVTATHZERI and GHDCDKQOW.?ILSEH

That illustrates why Hill ciphers are so difficult to crack—unless you happen to be so lucky as to have "captured" some pieces of plaintext along with the corresponding pieces of ciphertext. And, fortunately, the latter situation is the one you will be faced with in the computer work concerning Hill ciphers!

1.4 Our alphabet

In all the examples below, and in the computer work with Hill ciphers, *our alphabet consists of the 26 upper-case letters of the English alphabet followed by the period (.), the question mark (?), and the blank space,* in that order. In plaintext, the period and question mark have their usual meaning, and the blank space is employed, as usual, to separate words. So we can see it, *we shall denote the blank space below by the character*(␣). (In your work on the computer, use the blank space itself, however, and not this special character!)

When enciphering or deciphering, we shall represent the 29 characters in our alphabet in order by the nonnegative integers $0, 1, \ldots, 28$, as shown in Table 3.

Whenever we do not need to refer to the specific alphabet above, we shall denote the length of the alphabet by m. Moreover, we shall often refer to any

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

Table 3: Numerical representation of 29-letter alphabet

character in such an alphabet as a "letter" even when it is a punctuation symbol (as in our 29-character alphabet) or some other character that is not actually a letter in the ordinary sense.

Note that, when working with Hill ciphers on the computer, you should make your functions as general as possible to handle various alphabets, you should be using an arbitrary integer $m > 1$ (and sometimes an arbitrary prime integer m) as the alphabet length.

There is nothing magical about numbering the letters in our alphabet in ascending order (starting with 0). In practice, one would undoubtedly scramble the numbers in some arbitrary order (known to both the sender and the receiver of an enciphered message) so as to make cracking the cipher a little more difficult. For simplicity's sake, we shall stick with the numbering scheme shown.

2 Modular Arithmetic

To understand Hill ciphers, you first have to understand "modular arithmetic". This section explains what that is and why Hill ciphers require it.

2.1 A Hill 2-cipher

An arbitrary Hill n-cipher has as its key a given $n \times n$ matrix whose entries are nonnegative integers from among $0, 1, \ldots, m-1$, where m is the length of the alphabet.

Here is an example of a Hill 2-cipher for our alphabet above (where $m = 29$). The key is the 2×2 matrix:

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

We use A to encipher groups of 2 consecutive characters—**digraphs**—at a time. Let us first encipher the digraph BE. In our alphabet, the letters B and E are numbered 1 and 4, respectively, so we represent BE by the column vector

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

To encipher BE, we multiply this column vector by the key matrix A:

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 14 \\ 24 \end{bmatrix}$$

Finally, we use in reverse the correspondence between alphabet characters and numbers to see that 14 and 24 represent the characters O and Y, respectively. Thus the ciphertext corresponding to BE is OY.

Suppose, now, we wish to apply the same Hill 2-cipher instead to the plaintext AN. This is represented by the vector $[0\ 13]^T$, and we find:

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 13 \end{bmatrix} = \begin{bmatrix} 39 \\ 65 \end{bmatrix}$$

Oops! What letters do 39 and 65 represent? Our 29-letter alphabet is numbered from 0 up to only 28. What we do is simply "wrap around" the numbers from 29 through 57 to represent the 29 letters again, then the numbers from 58 through 86, etc., as in Table 4.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	.	?	␣
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86

Table 4: Wrapping numerical representations modulo 29

Thus 39 represents the same letter as $39 - 29 = 10$, namely, K; and 65 represents the same letter as $65 - 2 \cdot 29 = 7$, namely, H. Finally, we get from AN the corresponding ciphertext KH.

2.2 Equivalence and residues modulo an integer

The procedure of "wrapping" used in the Hill 2-cipher above is quite general. It is the same procedure you are accustomed to using when each midnight and each noon you begin again to number the hours 1, 2, etc. Here is the formal definition:

Definition 1. We are given an integer $m > 1$, called the **modulus**. Then we say that two integers a and b **congruent** to one another **modulo** m (congruent "mod m" for short), and we write

$$a \equiv b \pmod{m},$$

to mean that the difference $a - b$ is an integral multiple of m. In other words, $a \equiv b \pmod{m}$ when $a = b + k \cdot m$ for some integer k (positive, negative, or zero).

For our Hill 2-cipher above, we had

$$39 \equiv 10 \pmod{29} \quad \text{and} \quad 65 \equiv 7 \pmod{29}$$

because

$$39 - 10 = 1 \cdot 29 \quad \text{and} \quad 65 - 7 = 2 \cdot 29.$$

Here are some more examples, when the modulus is small, namely, when $m = 6$:

$$\begin{aligned} 9 &\equiv 3 \pmod 6 \\ 23 &\equiv 5 \pmod 6 \\ -8 &\equiv 4 \pmod 6 \\ 12 &\equiv 0 \pmod 6 \end{aligned}$$

Evidently $a \equiv b \pmod m$ whenever $a = b$ (for then $a - b = 0 = 0 \cdot m$), but not conversely. Thus the idea of "sameness" expressed by equivalence mod m generalizes the usual idea of equality. The most basic properties of equality also carry over to equivalence modulo m:

$$\begin{aligned} a &\equiv a \pmod m; \\ \text{if } a &\equiv b \pmod m, \text{ then } b \equiv a \pmod m; \\ \text{if } a &\equiv b \pmod m \text{ and } b \equiv c \pmod m, \text{ then } a \equiv c \pmod m. \end{aligned}$$

Observe that no two integers a and b differing by less than m can possibly be congruent modulo m. In particular, no two of the integers $0, 1, \ldots, m-1$ are congruent to one another modulo m.

Observe also that an arbitrary integer a can be divided by m to get a quotient q and a remainder r, that is, $a = q \cdot m + r$, with $0 \le r < m$. (For example, $23 = 3 \cdot 6 + 5$.) In that case, of course, $a \equiv r \pmod m$. (In the example just given, $23 \equiv 5 \pmod 6$.)

From the preceding two observations it now follows that *each integer is congruent modulo m to exactly one of the integers* $0, 1, \ldots, m-1$. This justifies the following definition:

Definition 2. Let m be a integer with $m > 1$. For an arbitrary integer a, the **residue of a modulo m** is the unique integer r among $0, 1, \ldots, m-1$ to which a is congruent modulo m.

For example, 5 is the residue of 23 modulo 6. And 5 is also the residue of -7 modulo 6.

To indicate that we replace a given integer by its residue modulo m, we sometimes say that we **reduce** the integer modulo m.

2.3 Modular arithmetic

For our purposes, what is most important is that *equivalence modulo m preserves sums*, that is, adding one pair of integers that are congruent modulo m to a second pair of integers gives a sum of the first pair that is congruent to the sum of the second pair:

$$\text{if } a \equiv c \pmod m \text{ and } b \equiv d \pmod m, \text{ then } a + b \equiv c + d \pmod m$$

For example, from $23 \equiv 5 \pmod 6$ and $-8 \equiv 4 \pmod 6$, we conclude $15 = 23 + (-8) \equiv 5 + 4 = 9 \pmod 6$.

Similarly, *equivalence modulo m preserves products*, that is, multiplying one pair of integers that are congruent modulo m to a second pair of integers gives a product of the first pair that is congruent to the product of the second pair:

$$\text{if } a \equiv c \pmod{m} \text{ and } b \equiv d \pmod{m}, \text{ then } a \cdot b \equiv c \cdot d \pmod{m}$$

For example, from $23 \equiv 5 \pmod 6$ and $-8 \equiv 4 \pmod 6$, we conclude $-184 = 23 \cdot (-8) \equiv 5 \cdot 4 = 20 \pmod 6$.

Because equivalence modulo m preservers both sums and products, we can now do **arithmetic modulo** m as follows: after adding or multiplying two integers, replace their sum or product, respectively, by its residue modulo m. If we do any combination of additions and multiplications of a number of integers, and then replace the final answer by its residue modulo m, we will get the same result as replacing the original integers by their residues modulo m (if they are not already between 0 and $m-1$) and then replacing each sum and product along the way by its residue modulo m.

For example, let us work with modulus $m = 29$ again. Then we can compute the residue of $4 \cdot 14 + 5 \cdot 17$ either by doing all the arithmetic in the usual way and then taking the residue—

$$4 \cdot 14 + 5 \cdot 17 = 56 + 85 = 141 \equiv 25 \pmod{29}$$

—or instead by taking the residues at every step along the way:

$$4 \cdot 14 + 5 \cdot 17 = 56 + 85 \equiv 27 + 27 = 54 \equiv 25 \pmod{29}$$

Similarly, we can find $2 \cdot 14 + 3 \cdot 17 \equiv 21 \pmod{29}$ in either of two ways (try it!). But these are precisely the computations one needs in order to encipher the plaintext OR by means of the Hill 2-cipher with the same key matrix

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

used in the example in Section 2.1.

The numerical representation of the digraph OR for our 29-letter alphabet is the column vector $[14\ 17]^T$. Now do all arithmetic modulo 29. Either work modulo 29 all along the way—

$$\begin{aligned} \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 14 \\ 17 \end{bmatrix} &= \begin{bmatrix} 28+51 \\ 56+85 \end{bmatrix} \\ &\equiv \begin{bmatrix} 28+22 \\ 27+27 \end{bmatrix} \pmod{29} \\ &= \begin{bmatrix} 50 \\ 54 \end{bmatrix} \\ &\equiv \begin{bmatrix} 21 \\ 25 \end{bmatrix} \pmod{29} \end{aligned}$$

—or else wait until the end:

$$\begin{aligned}\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 14 \\ 17 \end{bmatrix} &= \begin{bmatrix} 28+51 \\ 56+85 \end{bmatrix} \\ &= \begin{bmatrix} 79 \\ 141 \end{bmatrix} \\ &\equiv \begin{bmatrix} 21 \\ 25 \end{bmatrix} \pmod{29}\end{aligned}$$

Doing the arithmetic modulo 29 above has the same effect as "wrapping" of numbers greater than 28 back around our 29-letter alphabet, as we described when we first introduced this Hill cipher. And that is why we use modular arithmetic for Hill ciphers.

As in the example just worked, we may use arithmetic modulo m whenever we are doing matrix addition and multiplication, and we can freely take residues modulo m any step along the way without affecting the values modulo m in the final matrix answer.

2.4 The number system $\mathbb{Z}_m$

Let m be a given modulus. If we add or multiply two integers from among $0, 1, \ldots, m-1$ and take the residue of the sum or product, respectively, then we get another integer from among $0, 1, \ldots, m-1$. Thus the set $\{0,\ 1,\ \ldots,\ m-1\}$ is *closed* under the two new operations of "addition" and "multiplication" so defined. We use the usual symbols $+$ and $\cdot$ (or just juxtaposition) to denote these two operations on $\{0,\ 1,\ \ldots,\ m-1\}$.

Definition 3. The set

$$\mathbb{Z}_m = \{0,\ 1,\ \ldots,\ m-1\}$$

together with the two operations of addition and multiplication described above is called the number system of the **integers modulo** m.

For example, take $m = 4$. Then in $\mathbb{Z}_4$ we have

$$2 + 3 = 1 \quad \text{and} \quad 2 \cdot 3 = 2$$

because, of course,

$$2 + 3 = 5 \equiv 1 \pmod{4} \quad \text{and} \quad 2 \cdot 3 = 6 \equiv 2 \pmod{4}.$$

For such a low modulus as $m = 4$, we can display all possible sums and products in $\mathbb{Z}_m$ via an addition table and a multiplication table, as shown in Table 5.

In general, for any modulus m, addition and multiplication modulo m have many—but not quite all—of the familiar properties of addition and multiplication of ordinary real numbers:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Table 5: Addition and multiplication tables for $\mathbb{Z}_4$

Proposition 1. *Let m be an integer with $m > 1$. Then in $\mathbb{Z}_m$:*

1. *For all $a, b, c \in \mathbb{Z}_m$, $(a + b) + c = a + (b + c)$.*
2. *For all $a, b \in \mathbb{Z}_m$, $a + b = b + a$.*
3. *For each $a \in \mathbb{Z}_m$, $a + 0 = a = 0 + a$.*
4. *For each $a \in \mathbb{Z}_m$, there is a unique $x \in \mathbb{Z}_m$, called the* additive inverse *of a, such that $a + x = 0 = x + a$.*
5. *For all $a, b, c \in \mathbb{Z}_m$, $(ab)c = a(bc)$.*
6. *For all $a, b \in \mathbb{Z}_m$, $ab = ba$.*
7. *For each $a \in \mathbb{Z}_m$, $1 \cdot a = a = a \cdot 1$.*
8. *(This property is intentionally omitted!)*
9. *For all $a, b, c \in \mathbb{Z}_m$, $a(b + c) = ab + ac$.*
10. *In $\mathbb{Z}_m$, $1 \neq 0$.*

For a given $a \in \mathbb{Z}_m$, its additive inverse $x \in \mathbb{Z}_m$ as given by property 4, above, is *not* the negative integer $-a$ (except in the trivial case that $a = 0$). Rather, the additive inverse of a is the residue of $-a$ modulo m. For example, as the addition table (Table 5) for $\mathbb{Z}_4$ shows, the additive inverse of 3 in $\mathbb{Z}_4$ is 1, because $3 + 1 \equiv 0 \pmod 4$, and that just restates the equivalence $-3 \equiv 1 \pmod 4$.

Among the properties listed in Proposition 1, you may have expected to see the following one, which was omitted:

8. For each $a \in \mathbb{Z}_m$ with $a \neq 0$, there is a unique $y \in \mathbb{Z}_m$, called the **multiplicative inverse** or **reciprocal** of a, such that $a \cdot y = 1 = y \cdot a$.

The reason that Property 8 was omitted from Proposition 1 is that it simply is not true in general! For example, it is not true in $\mathbb{Z}_4$. Look again at the multiplication table for $\mathbb{Z}_4$ (see Table 5). We see that 2 does *not* have a reciprocal in $\mathbb{Z}_4$, since 1 is not among its products with each of the elements in $\mathbb{Z}_4$. By way of contrast, 3 *does* have a reciprocal in $\mathbb{Z}_4$, namely 3 itself, because in $\mathbb{Z}_4$ we

have $3 \cdot 3 = 1$. Thus, some nonzero elements of $\mathbb{Z}_4$ have reciprocals but others do not.[2]

When working with Hill ciphers, we shall need to do elementary row operations on matrices in order to put them into reduced echelon form, and we shall need to do all the arithmetic modulo a given alphabet length m. Row reduction will all go smoothly until we need to scale a row to make a pivot 1. Then we shall need to "divide" each entry in the row by the pivot modulo m, that is, multiply each entry by the reciprocal of the pivot. Unfortunately, *in general, the pivot need not have a reciprocal!* This will happen, for example, if the pivot is 2 and we are working modulo 4.

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Table 6: Multiplication table for $\mathbb{Z}_5$

Take a look, next, at the multiplication table for $\mathbb{Z}_5$ (see Table 6). Except for the row showing multiplication by 0, each row in the multiplication table includes a 1, and that means that every nonzero element of $\mathbb{Z}_5$ has a multiplicative inverse. In other words, Property 8, above, does hold in $\mathbb{Z}_5$.

The difference between $\mathbb{Z}_5$ and $\mathbb{Z}_4$—in so far as the existence of multiplicative inverses is concerned—is that the modulus 5 is a prime and 4 is not. Recall the definition:

Definition 4. An integer greater than 1 is said to be **prime** when it cannot be written as a product of two positive integers other than 1 and itself.

For example, 2, 3, 5, 7, 11, and 13 are primes, whereas 4, 6, 8, 9, 10, and 12 are not.)

The following definition is convenient:

Definition 5. A number system closed under two operations of addition and multiplication is called a **field** when Properties 1–10 all hold in it.

The examples of $\mathbb{Z}_5$ and $\mathbb{Z}_4$ discussed above illustrate the following result:

Proposition 2. *Let m be an integer with $m > 1$. Then $\mathbb{Z}_m$ is a field if and only if m is prime.*

In particular, $\mathbb{Z}_5$ is a field whereas $\mathbb{Z}_4$ is not. For every prime m, the field $\mathbb{Z}_m$ is finite. The real number system $\mathbb{R}$ is an example of an infinite field. So is the number system $\mathbb{Q}$ consisting of just the rational numbers.

[2]The general theorem relevant here is the following: a has a multiplicative inverse modulo m if and only if a is relatively prime to m, that is, 1 is the only positive integer that divides both a and m (in the ordinary sense of division). Proposition 2, below, follows from that.

The number system $\mathbb{Z}_{29}$ is a field because 29 is prime. For convenience, we list in Table 7 the reciprocals (multiplicative inverses) of the nonzero elements of $\mathbb{Z}_{29}$.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	15	10	22	6	5	25	11	13	3	8	17	9	27	2	20	12	21	26	16	18	4	24	23	7	19	14	28

Table 7: Multiplicative inverses modulo 29

You should try the very easy exercise of constructing the addition and multiplication tables for the field $\mathbb{Z}_2$.

2.5 Modular linear algebra

What we are doing, in effect, when working with Hill ciphers is using matrices and vectors whose entries belong to $\mathbb{Z}_m$ for some integer $m > 1$, and doing all arithmetic modulo m. Then much of what you have learned and will learn about linear algebra still makes sense in the context of $\mathbb{Z}_m$: matrix addition and multiplication; elementary row operations; linear independence; and linear transformations.

However, within $\mathbb{Z}_m$ we cannot, in general, always put a matrix into a row-equivalent *reduced*-echelon form, unless we are guaranteed that each nonzero element of $\mathbb{Z}_m$ has a multiplicative inverse, in other words, unless m is prime.

It should now be obvious why we adjoined three extra "letters" to the normal 26-letter English alphabet: the number system $\mathbb{Z}_{29}$ is a field, whereas $\mathbb{Z}_{26}$ is not. Of course it is always possible to use as a Hill cipher's key a matrix with entries in $\mathbb{Z}_m$ for an arbitrary $m > 1$. But unless the matrix entries are chosen especially carefully, it will not be possible to compute a matrix inverse modulo m, in other words, to get an inverse key for decipherment.

Hill ciphers for English ordinarily use $m = 26$. Such a non-prime modulus just makes Hill ciphers technically more difficult to work with. So for simplicity's sake, in the computer work we are usually sticking with a prime modulus, most often with 29.

When m is a prime such as 29, $\mathbb{Z}_m$ is a field. For *any* field of scalars—whether the reals $\mathbb{R}$ or $\mathbb{Z}_m$ or something else—virtually all the familiar theory and computations about vector spaces and their linear transformations carry over, including virtually everything about matrices.

In the next section we shall illustrate row reduction and computation of matrix inverses over $\mathbb{Z}_{29}$.

3 Hill Ciphers

At last we ready to look in detail at Hill ciphers and how to crack them.

3.1 Encipherment with a Hill cipher

Suppose we are given an alphabet of length $m > 1$ and an integer $n > 1$. Then a **Hill n-cipher** is given by an n-by-n matrix A with entries in $\mathbb{Z}_m$. That matrix prescribes the key for the cipher.

For such a key matrix A given, Hill's **algorithm to encipher a given plaintext** is as follows:

1. Separate the plaintext from left to right into some number k of groups (polygraphs) of n letters each. If you run out of letters when forming the final group, repeat the last plaintext letter as many times as needed to fill out that final group to n letters.
2. Replace each letter by the corresponding number of its position (from 0 through $m - 1$) in the alphabet to get k groups of n integers each.
3. Reshape each of the k groups of integers into an n-row column vector and in turn multiply A by each of those k column vectors modulo m.
4. After arranging all k of the resulting product n-row column vectors in order into a single $(k \cdot n)$-vector (with entries in $\mathbb{Z}_m$), replace each of these $k \cdot n$ entries with the corresponding letter of the alphabet.

   The result is the ciphertext corresponding to the original plaintext.

Note: When you implement the preceding algorithm on the computer, it may be more convenient to reverse steps 1 and 2, that is, first replace the letters with numbers, and only then do the grouping (and repeating the final number as needed to fill out the final group).

The n-row column vectors formed in step 3, above, to represent groups of plaintext letters are called **plaintext vectors**, even though these vectors are composed of numbers, not letters. Likewise, the n-row column vectors obtained by multiplying them by A modulo m in step 3 are called **ciphertext vectors**; these ciphertext vectors numerically represent groups of ciphertext letters.

Plaintext could, in principle, be any jumble of letters whatsoever from our alphabet, whether it makes sense or not, and so could ciphertext. Then the set of all plaintext vectors—and likewise the set of all ciphertext vectors—is nothing other than the set we shall denote by $\mathbb{Z}_m^n$ and that consists of all n-row column vectors with entries in $\mathbb{Z}_m$. And it is reasonable to call elements of $\mathbb{Z}_m^n$ "vectors", because we may add them and multiply them by scalars (that is, by elements of $\mathbb{Z}_m$) to obtain again elements of $\mathbb{Z}_m^n$—provided, as always, that we reduce results modulo m. In view of Proposition 1, the eight fundamental properties of vectors in $\mathbb{R}^n$ also hold for $\mathbb{Z}_m^n$. Hence we may speak of $\mathbb{Z}_m^n$ as a

"linear space" or "vector space".[3]

At this point we can say what a Hill n-cipher with key matrix A "really" is: the linear transformation from $\mathbb{Z}_m^n$ to $\mathbb{Z}_m^n$ that has standard matrix A. Even though the domain and codomain of this linear transformation are the same, it is suggestive to refer to the domain as consisting of all plaintext vectors, and the codomain as consisting of all ciphertext vectors.

In the example of a Hill 2-cipher worked out in Section 2.1, we did not have to cope with the complication mentioned in step 1 of the preceding algorithm, namely, a plaintext whose length is not an integral multiple of the key matrix size n. We shall illustrate how to do that below.

In step 3 of the algorithm, instead of separately multiplying A by each of the k column matrices, we can, of course, multiply A by the single n-by-k matrix formed from those columns. We shall also illustrate that in the following example.

Throughout all our examples, we shall continue to use our 29-letter alphabet, but the key matrix size n will be fairly small. (In a "real" Hill cipher, n might be fairly large, to make cracking the cipher harder.)

Example 1. The key matrix size here is $n = 3$, and the key matrix is

$$A = \begin{bmatrix} 17 & 5 & 20 \\ 23 & 9 & 3 \\ 11 & 2 & 12 \end{bmatrix}.$$

Our plaintext is the 10-letter message

WANT␣HELP.

(recall that we are displaying a blank space in text as an underscore).

First, we group the plaintext into polygraphs of length 3, repeating the final letter (a period) twice to fill out the fourth group:

WAN T␣H ELP ...

Second, we replace the letters there by the corresponding numbers (see Table 3).

22 0 13 19 28 7 4 11 15 26 26 26

[3]Technically speaking, we should refer to $\mathbb{Z}_m^n$ as a vector space only when the set $\mathbb{Z}_m$ of scalars is a field, that is, when m is prime. In the general case, when the scalars need not form a field, it is more proper to call $\mathbb{Z}_m^n$ a *module*.

Third, we apply the key:

$$A \begin{bmatrix} 22 & 19 & 4 & 26 \\ 0 & 28 & 11 & 26 \\ 13 & 7 & 15 & 26 \end{bmatrix} = \begin{bmatrix} 17 & 5 & 20 \\ 23 & 9 & 3 \\ 11 & 2 & 12 \end{bmatrix} \begin{bmatrix} 22 & 19 & 4 & 26 \\ 0 & 28 & 11 & 26 \\ 13 & 7 & 15 & 26 \end{bmatrix}$$
$$= \begin{bmatrix} 634 & 603 & 423 & 1092 \\ 545 & 710 & 236 & 910 \\ 398 & 349 & 246 & 650 \end{bmatrix}$$
$$\equiv \begin{bmatrix} 25 & 23 & 17 & 19 \\ 23 & 14 & 4 & 11 \\ 21 & 1 & 14 & 12 \end{bmatrix} \pmod{29}$$

Fourth, the numbers

$$25 \;\; 23 \;\; 21 \;\; 23 \;\; 14 \;\; 1 \;\; 17 \;\; 4 \;\; 14 \;\; 19 \;\; 11 \;\; 12$$

from the "strung-out" columns just obtained represent the letters:

Z X V X O B R E O T L M

For this Hill 3-cipher, the ciphertext corresponding to the plaintext SEND␣HELP. is thus ZXVXOBREOTLM.

Earlier, in Section 1.3, we illustrated how a polygraphic substitution such as a Hill cipher can translate two quite similar plaintexts into completely different ciphertexts. Here is that example again:

Example 2. This is a Hill 4-cipher. The key matrix is:

$$A = \begin{bmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{bmatrix}$$

For this A, the plaintexts DESTROY␣PLANE and DEPLOY␣PLANES encipher to PPAJZTXVTATHZERI and GHDCDKQOW.?ILSEH, respectively. We leave the arithmetic details of the encipherment as an exercise.

3.2 Deciphering ciphertext with given Hill key

For a Hill cipher, the transformation from ciphertext back to plaintext is just the inverse of the original transformation from plaintext to ciphertext. In other words, **if a Hill cipher has key matrix A, then the inverse transformation is the Hill cipher whose key matrix is A^{-1}.**

If we already have the inverse of the key matrix A, then we can use it to decipher any ciphertext.

Example 3. Again, as in Section 2.1, consider the Hill 2-cipher with key matrix

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}.$$

The inverse of A over our scalar field $\mathbb{Z}_{29}$ is

$$A^{-1} = \begin{bmatrix} 12 & 16 \\ 2 & 28 \end{bmatrix},$$

as we can check by the computations

$$A \begin{bmatrix} 12 & 16 \\ 2 & 28 \end{bmatrix} = \begin{bmatrix} 30 & 116 \\ 58 & 204 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{29}.$$

Suppose that we have received the cipher text

OYKHWBZI

that we want to decipher and that we possess the inverse key. Follow the four-step algorithm outlined in Section 3.1, but using A^{-1} instead of A, and interchanging the words "ciphertext" and "plaintext" there:

First, group the letters into digraphs (there is no need here to "pad out" the final group, since the text's length is a multiple of the key size):

OY KH WB ZI

Second, replace the letters by the corresponding numbers:

14 24 10 7 22 1 25 8

Third, multiply column vectors by A^{-1}:

$$A^{-1} \begin{bmatrix} 14 & 10 & 22 & 25 \\ 24 & 7 & 1 & 8 \end{bmatrix} = \begin{bmatrix} 552 & 232 & 280 & 428 \\ 700 & 216 & 72 & 274 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 & 19 & 22 \\ 4 & 13 & 14 & 13 \end{bmatrix} \pmod{29}$$

Fourth, the columns here, when strung out into

1 4 0 13 19 14 22 13

give the numerical representation of the original plaintext. It is left as an exercise for you to use Table 3 to complete the decipherment.

If you concocted the key matrix A to a Hill cipher (or if you did not but captured it), then you can certainly construct the inverse key by inverting A in the usual way through row reduction, but using always, of course, arithmetic modulo m.

Example 4. Suppose we know the key matrix

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

for a Hill 2-cipher (and, as usual, our 29-letter alphabet). We want to compute the inverse key.

To do that, we augment A with the identity matrix to its right and proceed to apply elementary row operations. (To keep the numbers small, we take residues modulo 29 after each row operation rather than wait until the end.) As usual, we write $\leftrightarrow$ for matrix equivalence under elementary row operations.[4] Here goes:

$$[A \,|\, I] = \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right] \tag{1}$$

$$\leftrightarrow \left[\begin{array}{cc|cc} 30 & 45 & 15 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right] \tag{2}$$

$$\equiv \left[\begin{array}{cc|cc} 1 & 16 & 15 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right] \pmod{29} \tag{3}$$

$$\leftrightarrow \left[\begin{array}{cc|cc} 1 & 16 & 15 & 0 \\ 0 & -59 & -60 & 1 \end{array}\right] \tag{4}$$

$$\equiv \left[\begin{array}{cc|cc} 1 & 16 & 15 & 0 \\ 0 & 28 & 27 & 1 \end{array}\right] \pmod{29} \tag{5}$$

$$\leftrightarrow \left[\begin{array}{cc|cc} 1 & 16 & 15 & 0 \\ 0 & 784 & 756 & 28 \end{array}\right] \tag{6}$$

$$\equiv \left[\begin{array}{cc|cc} 1 & 16 & 15 & 0 \\ 0 & 1 & 2 & 28 \end{array}\right] \pmod{29} \tag{7}$$

$$\leftrightarrow \left[\begin{array}{cc|cc} 1 & 0 & -17 & -448 \\ 0 & 1 & 2 & 28 \end{array}\right] \tag{8}$$

$$\equiv \left[\begin{array}{cc|cc} 1 & 0 & 12 & 16 \\ 0 & 1 & 2 & 28 \end{array}\right] \pmod{29} \tag{9}$$

Did you follow all the steps in the reduction? In step (2) we multiplied the top row by 15, which is the reciprocal modulo 29 of the first pivot 2 (see Table 7). And in step (6) we multiplied the bottom row by 28, which is the reciprocal modulo 29 of the second pivot 28.

From the final reduced-echelon matrix above, we read off the inverse as usual:

$$A^{-1} = \begin{bmatrix} 12 & 16 \\ 2 & 28 \end{bmatrix}$$

[4] Note that two matrices that are congruent modulo m are, in fact, row-equivalent with respect to the scalars $\mathbb{Z}_m$.

3.3 Cracking a Hill cipher

How do you crack a Hill cipher? In other words, how do you discover the inverse key when you do not know the key? If you have "captured" enough plaintext along with the corresponding ciphertext, then you may be able to use the following theorem.

Theorem 1 (Cracking Theorem). *Suppose the length m of the alphabet is a prime. Let $\vec{p_1}, \vec{p_2}, \ldots, \vec{p_n}$ be n plaintext vectors for a Hill n-cipher having (unknown) key matrix A, and let $\vec{c_1}, \vec{c_2}, \ldots, \vec{c_n}$ be the corresponding ciphertext vectors. Suppose these plaintext vectors are* linearly independent *over $\mathbb{Z}_m$. Form the matrix*

$$P = \left[\begin{array}{c|c|c|c} \vec{p_1} & \vec{p_2} & \ldots & \vec{p_n} \end{array} \right]$$

having the plaintext vectors as its columns, and the matrix

$$C = \left[\begin{array}{c|c|c|c} \vec{c_1} & \vec{c_2} & \ldots & \vec{c_n} \end{array} \right]$$

having the ciphertext vectors as its columns.

Then the same sequence of elementary row operations that reduces C^T to the identity matrix I reduces P^T to the transpose *$(A^{-1})^T$ of the inverse key matrix A^{-1}.*

Our proof of the Cracking Theorem will involve "elementary" matrices. An $n \times n$ matrix is said to be *elementary* when it can be obtained from the $n \times n$ identity matrix I_n by performing a single elementary row operation on I_n. Here are the facts you need to know about elementary matrices in order to follow the proof of the Cracking Theorem:

- Performing a single elementary row operation on an arbitrary $n \times k$ matrix M yields the same result as multiplying M on its left by the elementary matrix E formed by performing that same elementary row operation on I_n.
- Each elementary matrix is invertible, and its inverse is also an elementary matrix.

The second of these facts holds, of course, provided that nonzero scalars have multiplicative inverses (otherwise, scaling a row need not be a reversible operation). In our situation, this means that $\mathbb{Z}_m$ is a field, in other words, that m is prime.

In view of the first of these facts, the reduced row-echelon form H of a matrix M can be obtained in the form $E_k \cdots E_2 E_1 M = H$ for some elementary matrices E_1, E_2, ... , E_k (these elementary matrices are just the ones corresponding to the successive elementary row operations needed to row-reduce M to H). In particular, for an invertible matrix M, there are elementary matrices E_1, E_2, ... , E_k for which $E_k \cdots E_2 E_1 M = I$.

Proof of Cracking Theorem. First, note that C^T really is row-equivalent to I, because it is the transpose of C, and the latter is invertible because, by hypothesis, its columns are independent.

Let $E_1, E_2, \ldots, E_k$ be, for some k, the elementary matrices corresponding to the elementary row operations that reduce C^T to I. Then

$$E_k \cdots E_2 E_1 C^T = I.$$

Now $A\vec{p_j} = \vec{c_j}$ for each $j = 1, 2, \ldots, n$, so

$$AP = C.$$

Take transposes there to get

$$P^T A^T = C^T.$$

Multiply both sides by $E_k \cdots E_2 E_1$ and group the results as shown:

$$(E_k \cdots E_2 E_1 P^T) A^T = E_k \cdots E_2 E_1 C^T.$$

But the right-hand side is just I, so:

$$(E_k \cdots E_2 E_1 P^T) A^T = I$$

This shows two things: First, A^T is invertible, and hence A itself—which is $(A^T)^T$—is invertible. Second, multiplication of both sides of the last-displayed equation by $(A^T)^{-1}$—which equals $(A^{-1})^T$—gives

$$E_k \cdots E_2 E_1 P^T = (A^{-1})^T.$$

This means that the elementary row operations corresponding to $E_1, E_2, \ldots, E_k$ do reduce P^T to $(A^{-1})^T$. □

For this theorem to be applicable, you must know—or be able to determine—the size n of the unknown key matrix (as well as the alphabet length m). If you do, and if you have n plaintext vectors and the corresponding ciphertext vectors, you can proceed to apply the theorem. With the notation used there, **the procedure for cracking a Hill cipher** is as follows:

> Use row reduction (modulo m, of course) on the n-by-$2n$ matrix
>
> $$\left[C^T \mid P^T\right].$$
>
> If the reduction can be completed to put that into reduced-echelon form, and if that form is
>
> $$[I \mid X]$$
>
> for some X, then the ciphertext vectors—which are transposes of the rows of C^T—are, in fact, linearly independent, the key matrix A is invertible, and the right half X of the reduced-echelon form $[I \mid X]$ is $(A^{-1})^T$. Then A^{-1} is the transpose of X.

Here is a simple example:

Example 5. Suppose you know both the plaintext BEAN and the corresponding ciphertext OYKH for a Hill 2-cipher. You want to determine the inverse key, so that you can decipher any further enciphered messages you might intercept.

Separate the letters in both plaintext and ciphertext as usual into groups of 2, get the numbers 1 4 0 13 and 14 24 10 7 corresponding to the letters of the plaintext and ciphertext, respectively, and form them into 2-row columns. The plaintext vectors are

$$\vec{p_1} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \vec{p_2} = \begin{bmatrix} 0 \\ 13 \end{bmatrix},$$

and the corresponding ciphertext vectors are

$$\vec{c_1} = \begin{bmatrix} 14 \\ 24 \end{bmatrix}, \quad \vec{c_2} = \begin{bmatrix} 10 \\ 7 \end{bmatrix}.$$

Thus the matrices P and C of the theorem are:

$$P = [\vec{p_1} \mid \vec{p_2}] = \begin{bmatrix} 1 & 0 \\ 4 & 13 \end{bmatrix},$$
$$C = [\vec{c_1} \mid \vec{c_2}] = \begin{bmatrix} 14 & 10 \\ 24 & 7 \end{bmatrix}$$

Then

$$\left[C^T \mid P^T\right] = \left[\begin{array}{cc|cc} 14 & 24 & 1 & 4 \\ 10 & 7 & 0 & 13 \end{array}\right].$$

As in Section 3.2, row-reduce this modulo 29 to get the reduced-echelon form

$$\left[\begin{array}{cc|cc} 1 & 0 & 12 & 2 \\ 0 & 1 & 16 & 28 \end{array}\right].$$

Since the left half of this is I, the right half is $(A^{-1})^T$:

$$(A^{-1})^T = \begin{bmatrix} 12 & 2 \\ 16 & 28 \end{bmatrix}$$

Hence

$$A^{-1} = \begin{bmatrix} 12 & 16 \\ 2 & 28 \end{bmatrix}.$$

(This is the same inverse we found in Example 4.)

Once you have computed the inverse key A^{-1} for an unknown Hill key A, then of course you can compute the original key as the inverse $A = (A^{-1})^{-1}$ of the inverse key. And you might very well want to do that, so you could deceive the "enemy" by using the key yourself to create fake ciphered messages to send to them.