transmission errors, and represents one of the early attempts to combine error correction with cryptography. Eventually, the *ADFGX* cipher was replaced by the *ADFGVX* cipher, which used a 6×6 initial matrix. This allowed all 26 letters plus 10 digits to be used.

For more on the cryptanalysis of the ADFGX cipher, see [Kahn].

2.7 Block Ciphers

In many of the aforementioned cryptosystems, changing one letter in the plaintext changes exactly one letter in the ciphertext. In the shift, affine, and substitution ciphers, a given letter in the ciphertext always comes from exactly one letter in the plaintext. This greatly facilitates finding the key using frequency analysis. In the Vigenère system, the use of blocks of letters, corresponding to the length of the key, made the frequency analysis more difficult, but still possible, since there was no interaction among the various letters in each block. Block ciphers avoid these problems by encrypting blocks of several letters or numbers simultaneously. A change of one character in a plaintext block should change potentially all the characters in the corresponding ciphertext block.

The Playfair cipher in Section 2.6 is a simple example of a block cipher, since it takes two-letter blocks and encrypts them to two-letter blocks. A change of one letter of a plaintext pair will always change at least one letter, and usually both letters, of the ciphertext pair. However, blocks of two letters are too small to be secure, and frequency analysis, for example, is usually successful.

Many of the modern cryptosystems that will be treated later in this book are block ciphers. For example, DES operates on blocks of 64 bits. AES uses blocks of 128 bits. RSA uses blocks several hundred bits long, depending on the modulus used. All of these block lengths are long enough to be secure against attacks such as frequency analysis.

The standard way of using a block cipher is to convert blocks of plaintext to blocks of ciphertext, independently and one at a time. This is called the electronic codebook (ECB) mode. However, there are ways to use feedback from the blocks of ciphertext in the encryption of subsequent blocks of plaintext. This leads to the cipher block chaining (CBC) mode and cipher feedback (CFB) mode of operation. These are discussed in Section 4.5.

In this section, we discuss the Hill cipher, which is a block cipher invented in 1929 by Lester Hill. It seems never to have been used much in practice. Its significance is that it was perhaps the first time that algebraic methods (linear algebra, modular arithmetic) were used in cryptography in an essential way. As we'll see in later chapters, algebraic methods now occupy a central position in the subject.

Choose an integer n, for example $n = 3$. The key is an $n \times n$ matrix M whose entries are integers mod 26. For example, let

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 11 & 9 & 8 \end{pmatrix}.$$

The message is written as a series of row vectors. For example, if the message is *abc*, we change this to the single row vector $(0, 1, 2)$. To encrypt, multiply the vector by the matrix (traditionally, the matrix appears on the right in the multiplication; multiplying on the left would yield a similar theory) and reduce mod 26:

$$(0,\ 1,\ 2)\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 11 & 9 & 8 \end{pmatrix} \equiv (0,\ 23,\ 22) \pmod{26}.$$

Therefore, the ciphertext is AXW. (The fact that the first letter a remained unchanged is a random occurrence; it is not a defect of the method.)

In order to decrypt, we need the determinant of M to satisfy

$$\gcd(\det(M),\ 26) = 1.$$

This means that there is a matrix N with integer entries such that $MN \equiv I \pmod{26}$, where I is the $n \times n$ identity matrix.

In our example, $\det(M) = -3$. The inverse of M is

$$\frac{-1}{3}\begin{pmatrix} -14 & 11 & -3 \\ 34 & -25 & 6 \\ -19 & 13 & -3 \end{pmatrix}.$$

Since 17 is the inverse of -3 mod 26, we replace $-1/3$ by 17 and reduce mod 26 to obtain

$$N = \begin{pmatrix} 22 & 5 & 1 \\ 6 & 17 & 24 \\ 15 & 13 & 1 \end{pmatrix}.$$

The reader can check that $MN \equiv I \pmod{26}$.

For more on finding inverses of matrices mod n, see Section 3.8.

The decryption is accomplished by multiplying by N, as follows:

$$(0,\ 23,\ 22)\begin{pmatrix} 22 & 5 & 1 \\ 6 & 17 & 24 \\ 15 & 13 & 1 \end{pmatrix} \equiv (0,\ 1,\ 2) \pmod{26}.$$

In the general method with an $n \times n$ matrix, break the plaintext into blocks of n characters and change each block to a vector of n integers between 0 and 25 using $a = 0, b = 1, \ldots, z = 25$. For example, with the matrix M as above, suppose our plaintext is

blockcipher.

This becomes (we add an x to fill the last space)

1 11 14 2 10 2 8 15 7 4 17 23.

Now multiply each vector by M, reduce the answer mod 26, and change back to letters:

$$(1, 11, 14)M = (199, 183, 181) \equiv (17, 1, 25) \pmod{26} = RBZ$$

$$(2, 10, 2)M = (64, 72, 82) \equiv (12, 20, 4) \pmod{26} = MUE,$$

etc.

In our case, the ciphertext is

RBZMUEPYONOM.

It is easy to see that changing one letter of plaintext will usually change n letters of ciphertext. For example, if *block* is changed to *clock*, the first three letters of ciphertext change from RBZ to SDC. This makes frequency counts less effective, though they are not impossible when n is small. The frequencies of two-letter combinations, called **digrams**, and three-letter combinations, **trigrams**, have been computed. Beyond that, the number of combinations becomes too large (though tabulating the results for certain common combinations would not be difficult). Also, the frequencies of combinations are so low that it is hard to get meaningful data without a very large amount of text.

Now that we have the ciphertext, how do we decrypt? Simply break the ciphertext into blocks of length n, change each to a vector, and multiply on the right by the inverse matrix N. In our example, we have

$$RBZ = (17, 1, 25) \mapsto (17, 1, 25)N = (755, 427, 66) \equiv (1, 11, 14) = blo,$$

and similarly for the remainder of the ciphertext.

The Hill cipher is difficult to decrypt using only the ciphertext, but it succumbs easily to a known plaintext attack. If we do not know n, we can try various values until we find the right one. So suppose n is known. If we have n of the blocks of plaintext of size n, then we can use the plaintext

and the corresponding ciphertext to obtain a matrix equation for M (or for N, which might be more useful). For example, suppose we know that $n = 2$ and we have the plaintext

$$\begin{array}{cccccccccccccccc} & & & & & & & & & howareyoutoday = & & & & & & \\ 7 & 14 & 22 & 0 & 17 & 4 & 24 & 14 & 20 & 19 & 14 & 3 & 0 & 24 \end{array}$$

corresponding to the ciphertext

$$\begin{array}{cccccccccccccccc} & & & & & & & & & ZWSENIUSPLJVEU = & & & & & & \\ 25 & 22 & 18 & 4 & 13 & 8 & 20 & 18 & 15 & 11 & 9 & 21 & 4 & 20 \end{array}$$

The first two blocks yield the matrix equation

$$\begin{pmatrix} 7 & 14 \\ 22 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 25 & 22 \\ 18 & 4 \end{pmatrix} \pmod{26}.$$

Unfortunately, the matrix $\begin{pmatrix} 7 & 14 \\ 22 & 0 \end{pmatrix}$ has determinant -308, which is not invertible mod 26 (though this matrix could be used to reduce greatly the number of choices for the encryption matrix). Therefore, we replace the last row of the equation, for example, by the fifth block to obtain

$$\begin{pmatrix} 7 & 14 \\ 20 & 19 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 25 & 22 \\ 15 & 11 \end{pmatrix} \pmod{26}.$$

In this case, the matrix $\begin{pmatrix} 7 & 14 \\ 20 & 19 \end{pmatrix}$ is invertible mod 26:

$$\begin{pmatrix} 7 & 14 \\ 20 & 19 \end{pmatrix}^{-1} \equiv \begin{pmatrix} 5 & 10 \\ 18 & 21 \end{pmatrix} \pmod{26}.$$

We obtain

$$M \equiv \begin{pmatrix} 5 & 10 \\ 18 & 21 \end{pmatrix}\begin{pmatrix} 25 & 22 \\ 15 & 11 \end{pmatrix} \equiv \begin{pmatrix} 15 & 12 \\ 11 & 3 \end{pmatrix} \pmod{26}.$$

Because the Hill cipher is vulnerable to this attack, it cannot be regarded as being very strong.

A chosen plaintext attack proceeds by the same strategy, but is a little faster. Again, if you do not know n, try various possibilities until one works. So suppose n is known. Choose the first block of plaintext to be $baaa\cdots = 1000\ldots$, the second to be $abaa\cdots = 0100\ldots$, and continue through the nth